

CA-MLNet:双流记忆通道注意高精度轨迹预测模型

张晓涵¹ 杨 飞² 马靖尧¹ 赵韩栋¹ 赵 旭³

1 北京信息科技大学自动化学院人工智能系 北京 102206
2 北京信息科技大学自动化学院控制科学与工程系 北京 102206
3 北京信息科技大学高动态导航技术北京市重点实验室 北京 100192
(Zhang1121866981@163.com)

摘 要 针对复杂环境下单一模型轨迹预测精度不足的问题,提出一种基于改进状态空间模型与记忆网络融合的新型预测架构通道注意力增强型双流记忆网络 CA-MLNet。该工作的核心创新在于:(1)通过结构化状态空间模型(Structured State-Space Model,SSM),增强对动态环境的空间特征选择能力;(2)创新性地融合 xLSTM 的 mLSTM 模块作为辅助时序建模单元,构建双分支特征融合架构。改进后的 SSM 模块能有效捕捉长距离空间依赖,而 mLSTM 模块通过指数门控机制强化局部时序特征提取,两者通过自适应权重融合机制实现优势互补。在 GeoLife GPS Trajectories 行人轨迹数据集上的实验表明,所提模型在预测场景下达到 99.01% 的预测精度,较基准模型提升 8.18%,相比 xLSTM 架构提升 3.14%。消融实验验证了 SSM 模块改进对空间特征选择准确率的贡献达 43.6%,双模块协同工作使轨迹偏移误差降低,为智能交通预警系统提供了高精度的解决方案。

关键词: 通道注意力增强型双流记忆网络;轨迹预测;指数门控机制;自适应权重融合;智慧安防

中图分类号 TP183

CA-MLNet: Dual-stream Memory and Channel Attention Based High-precision Trajectory Prediction Model

ZHANG Xiaohan¹, YANG Fei², MA Jingyao¹, ZHAO Hanyue¹ and ZHAO Xu³

1 Department of Artificial Intelligence, School of Automation, Beijing Information Science and Technology University, Beijing 102206, China
2 Department of Control Science and Engineering, School of Automation, Beijing Information Science and Technology University, Beijing 102206, China
3 Beijing Key Laboratory of High Dynamic Navigation Technology, Beijing Information Science and Technology University, Beijing 100192, China

Abstract To address the insufficient prediction accuracy of single models in complex environments, this study proposes a novel prediction architecture named Channel Attention-enhanced Dual-stream Memory Network (CA-MLNet), which integrates improved state space models with memory networks. The core innovations of this work include: (1) Reconstruction of the Structured State-Space Model, it enhances its spatial feature selection capability in dynamic environments; (2) Innovative integration of xLSTM's mLSTM module as an auxiliary temporal modeling unit, it establishes a dual-branch feature fusion architecture. The improved SSM module effectively captures long-range spatial dependencies, while the mLSTM module enhances local temporal feature extraction through exponential gating mechanisms. These components achieve complementary advantages via adaptive weight fusion mechanisms. Experimental results on the GeoLife GPS Trajectories dataset demonstrate that the proposed model achieves a prediction accuracy of 99.01%, representing an 8.18% improvement over the baseline Mamba model and a 3.14% enhancement compared to the xLSTM architecture. Ablation experiments verify that the SSM module modification contributes 43.6% to spatial feature selection accuracy, with dual-module collaboration reducing trajectory offset errors. This approach provides a high-precision solution for intelligent traffic early warning systems.

Keywords Channel attention-enhanced dual-stream memory network (CA-MLNet), Trajectory prediction, Exponential gating mechanisms, Adaptive weight fusion, Smart security

随着智慧园区建设的推进,人员轨迹预测成为提升安全管理主动性的关键技术^[1]。智慧园区安防的核心需求是实现“预判-管控”闭环,电子围栏技术作为物理空间管控的关键手

段^[2],其有效性高度依赖轨迹预测的精度与实时性。传统安防技术侧重事后告警,传统电子围栏也仅能在目标触发边界后告警(如参考文献[1]的YOLOv5入侵检测),难以主动预

基金项目:国家自然科学基金青年科学基金(62406032);北京市自然科学基金(4242036);中国兵器激光器件技术重点实验室基金(QT23120014)

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China(62406032), Natural Science Foundation of Beijing, China(4242036) and China Ordnance Key Laboratory of Laser Device Technology Foundation(QT23120014).

通信作者:杨飞(yangfei@bistu.edu.cn)

判风险;而多源传感器采集的时空数据虽可构建人员移动模型,但现有轨迹预测方法在实时性、精度和场景适应性上仍面临挑战。早期高斯混合模型^[3]和时间序列分析^[4]计算复杂且长程建模能力不足;神经网络虽有改进,但 CNN 难以捕捉时序交互^[5];LSTM 及其变体受顺序计算限制^[6];xLSTM 虽能增强并行能力,但多模态处理欠佳^[7];Mamba 凭借结构化状态空间模型(SSM)^[8-9],在多模态处理中表现出色。同时,多尺度卷积交叉注意力机制在计算机视觉和时序预测领域取得进展,如 YOLOv12 的 MSCA 模块^[10-11]、HPNet 的历史预测注意力模块^[12],以及 DynaScale 框架^[13]、T-Mamba 架构^[14]等均展现出技术优势。为此,本研究提出通道注意力增强型双流记忆网络(CA-MLNet),以融合 SSM 与改进型 xLSTM 的优势。双流架构中,SSM 模块高效整合多源数据,mLSTM 分支强化长序列特征提取,多通道注意力聚焦关键时空特征;结合 MSCA 模块和 MLP 块构建层次化特征表达,并设计动态模型分配策略适应多目标并行预测。相较现有方法,CA-MLNet 创新地将 SSM 与 mLSTM 结合,通过双流架构协同处理时空特征,设计通道注意力机制提升数据利用率,提出动态模型分配策略支持大规模目标建模,引入多尺度卷积交叉注意力与 MLP 块增强特征捕捉和泛化能力。通过高精度轨迹预测(99.01%的准确率),CA-MLNet 可提前输出目标未来一段时间内的位置序列,使电子围栏从“被动响应”升级为“主动预警”,且电子围栏的区域划分信息(如园区禁入区坐标范围)可作为模型先验约束,将其输入算法处理层的跨注意力模块,有助于精准识别“正常轨迹-异常轨迹”边界,进一步降低误报率。实验表明,该模型在预测精度和计算效率方面均显著优于现有主流方法,为智慧园区安防系统提供了新的技术解决方案。

1 方法整体结构

基于 CA-MLNet 与多系统定位的轨迹预测与电子围栏系统的网络模型结构如图 1 所示。

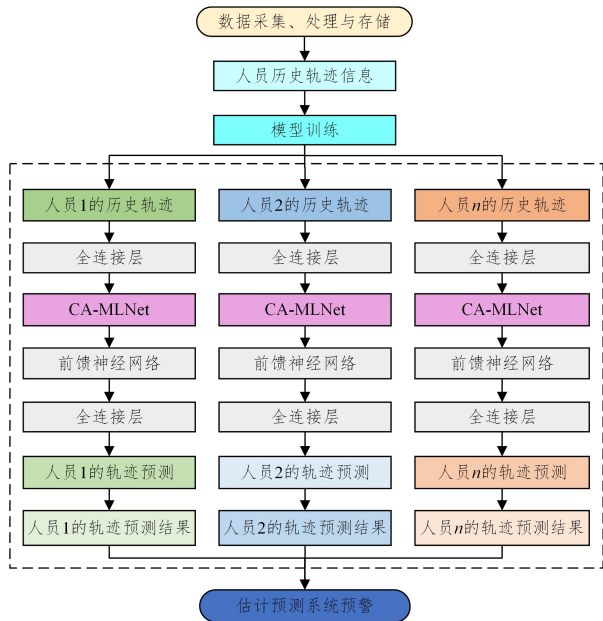


图 1 基于 CA-MLNet 与多元信息融合的轨迹预测系统的网络模型结构

Fig. 1 Network model structure of trajectory prediction system based on CA-MLNet and multivariate information fusion

整个通道注意力增强型双流记忆网络预测架构按照功能被划分为三大层:数据采集、处理与存储层,算法处理层,以及应用层。它们协同工作,以实现手机位置信息的收集、轨迹预测以及管理。

1.1 数据采集、处理与存储

该系统通过多系统定位获取位置信息数据^[8],进而得到行人的轨迹。如图 2 所示,在多系统定位系统中,卫星向手机(定位设备)发送定位信号。设信号传播速度(光速)为 c ,手机接受两颗卫星信号的时间差为 Rt ,卫星 S_0 和 S_1 错峰发送信号,卫星间距离为 d_0 (可通过卫星间 TOF 测量或因其相对位置已知而确定),卫星 S_2 收到 S_1 信号到 S_2 发送信号之间的时间为 D_2 ,此时距离差 d 的计算式为:

$$d = c \times Rt + d_0 + D_2 \quad (1)$$

当接收到 4 颗卫星信号时,可得到 3 组距离差,每一组距离差都有上述公式形式。在三维空间中,这些距离差信息可构建 3 个方程组,通过求解这些方程组可得到手机的坐标 (x, y, z) 。设 4 颗卫星为 S_0, S_1, S_2 和 S_3 ,与手机的距离分别为 r_0, r_1, r_2 和 r_3 ,以卫星 S_0 为参考点,则有以下方程组:

$$\begin{cases} r_1 - r_0 = d_{01} \\ r_2 - r_0 = d_{02} \\ r_3 - r_0 = d_{03} \end{cases} \quad (2)$$

其中, d_{01}, d_{02}, d_{03} 是通过上述距离差公式计算得到的距离差值,同时:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}, i = 0, 1, 2, 3 \quad (3)$$

其中, (x_i, y_i, z_i) 是卫星 i 的坐标,通过解此方程组可确定手机坐标 (x, y, z) 。

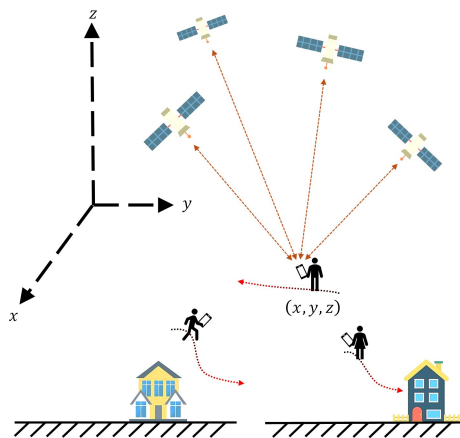


图 2 定位示意图

Fig. 2 Schematic diagram of positioning

由于多系统定位可能会出现漂移并产生噪声,因此有必要对收集到的位置信息进行清理,并去除异常数据。当信息不连续时,使用插值算法来完善轨迹数据,以确保其完整性。经过处理的数据将被存储在数据库中。为应对大规模时间序列数据的存储问题,采用时间序列数据库来存储用户的历史轨迹数据,这有利于后续的轨迹预测管理。

1.2 算法处理

算法处理层由轨迹预测模型与前馈神经网络^[15]组成。如图 3 所示,该轨迹预测整体架构采用混合设计,通过多阶段模块协同处理时序与空间特征。采集处理后的数据输入系统后,首先经过投影层映射至高维特征空间,提取基础运动模式;随后,多层循环单元捕获序列的时序依赖性,建模动态演

变规律;接着,CA-MLNet 模块引入跨注意力机制(Cross Attention)分析实体间交互(如障碍物避让),并通过多层网络细化时空特征;前馈网络进一步引入非线性变换增强特征表达能力,辅以归一化层稳定训练过程并加速收敛。最终,模型通过递归或解码结构输出未来轨迹。整体架构融合了循环网络的时序建模优势、注意力机制的全局交互感知能力以及深度网络的非线性表征,兼顾了复杂场景中的动态适应性与预测精度。

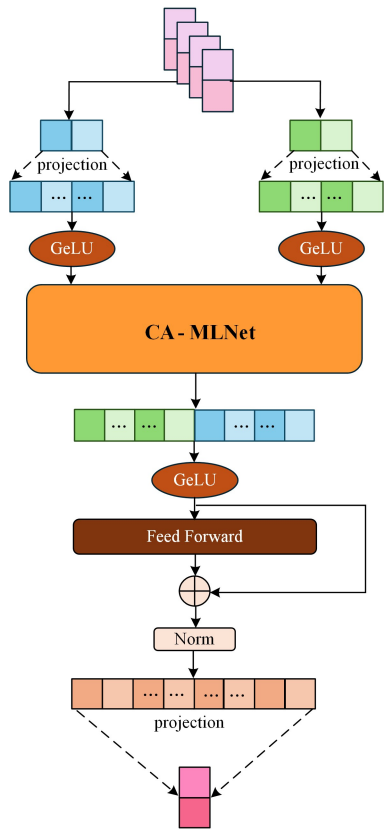


图3 轨迹预测的整体架构

Fig. 3 Overall architecture of trajectory prediction

轨迹预测模型 CA-MLNet 如图 4 所示,是 n 个改进 SSM 模块和 m 个改进 mLSTM 模块的并行结合,并将两者的输出结果拼接后输入到一个前馈神经网络进行预测。

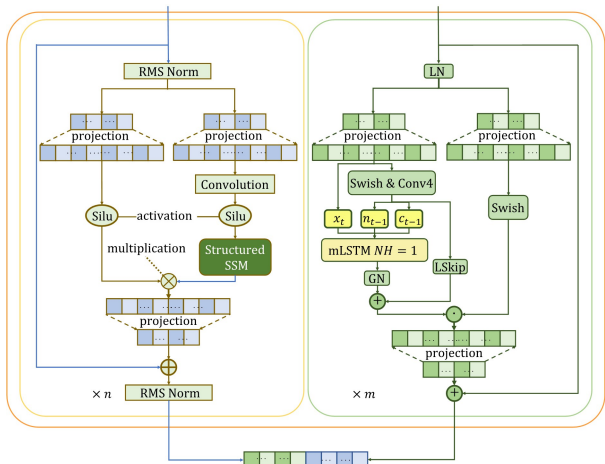


图4 轨迹预测模型 CA-MLNet 的算法架构

Fig. 4 Algorithm architecture of CA-MLNet for trajectory prediction model

输入首先经过 RMS Norm 进行均方根归一化以稳定训练;随后通过多层投影(线性变换)调整特征维度,提取基础模式;卷积层与情境激活/乘法模块(Situ activation & multiplication)协同作用,捕捉局部时空关联;结构化状态空间模型(Structured SSM)进一步建模长程依赖关系,增强序列的全局感知能力。架构中嵌入了改进的循环单元 mLSTM(可能为多头或轻量化设计),通过门控机制处理时序动态,并引入组归一化(GN)与跳跃连接(LSkip)缓解梯度问题;Swish 激活函数与 4 层卷积(Conv4)结合,提升非线性表征能力。整体流程呈现模块化堆叠,支持端到端训练,最终通过递归结构输出预测结果。该架构融合了归一化、卷积、状态空间模型与 LSTM 的优势,兼顾了局部特征提取、长程依赖建模和时序动态演化,适用于高精度、高稳定性的序列预测任务。

1.3 应用层

本系统的电子围栏功能并非独立模块,而是与 CA-MLNet 轨迹预测模型深度耦合。具体而言,定位设备采集的实时位置数据经预处理后,一方面输入模型生成未来 5~10 min 的预测轨迹,另一方面与电子围栏的坐标范围(如矩形围栏的左上角 (x_1, y_1) 与右下角 (x_2, y_2))进行空间交集计算。若预测轨迹中存在满足 $x_1 \leq x \leq x_2$ 且 $y_1 \leq y \leq y_2$ 的点,则系统触发分级预警;同时,电子围栏的历史入侵记录可作为模型的负样本,用于优化 mLSTM 模块的门控机制,提升对“突发越界行为”的特征捕捉能力,形成“预测—管控—反馈”的迭代优化闭环,这与参考文献 [1] 的入侵检测技术、参考文献 [2] 的海上围栏应用形成技术互补。

2 结构化状态空间模型

在序列建模领域,随着数据规模与序列长度的不断增加,高效处理长程依赖关系与高维数据的需求愈发迫切。结构化状态空间模型(Structured State-Space Model,SSM)作为一种新兴的序列建模工具,凭借其独特的数学结构与高效的计算特性,在时序预测、自然语言处理、轨迹预测等诸多领域展现出显著优势。本章将深入剖析 SSM 的核心原理、特性及其在轨迹预测场景中的关键作用。

2.1 SSM 的核心原理

SSM 基于状态空间模型的框架,通过定义状态转移方程与输出方程来建模序列数据。

状态方程:通过隐藏状态传递序列信息。

$$h_t = A h_{t-1} + B u_t$$
 (4)

其中, h_t 表示 t 时刻的状态向量; A 为状态转移矩阵,刻画了前一时刻状态对当前状态的影响; u_t 为当前输入; B 为输入投影矩阵,用于将当前输入 u_t 映射到状态空间。

输出方程:将隐藏状态映射为输出。

$$y_t = C h_t + D u_t$$
 (5)

其中, y_t 是模型输出; C 为输出矩阵,将状态向量转换为输出空间的表示; D 为直接输入映射矩阵,允许输入直接影响输出。

SSM 的“结构化”体现在矩阵 A, B, C, D 的设计上,这些矩阵可被赋予特定的结构(如低秩、对角等),以降低模型复杂度并引入归纳偏置,从而更高效地捕捉序列中的长期依赖与动态模式。

2.2 SSM 的特征与优势

SSM 的理论基础可追溯至早期序列建模研究, Schmidhuber 团队于 1991 年提出的状态空间模型^[16]首次将“状态转移”与“序列记忆”结合,为后续 SSM 的发展奠定了框架。该研究提出通过隐藏状态的线性演化捕捉长程依赖,其核心思想与本文改进的 SSM 模块一致——通过矩阵化的状态更新方程(式(4)、式(5))实现信息的高效传递。区别于早期模型的固定参数设计,SSM 模块进一步引入动态结构化矩阵(如低秩分解的状态转移矩阵 A),并结合多源数据输入适配(如融合 GPS 位置与速度特征),使其更适用于轨迹预测场景,同时保留了 $O(T)$ 线性计算复杂度的优势。

相较于传统循环神经网络(RNN)的 $O(T)$ 顺序计算复杂度(T 为序列长度)与 Transformer 的 $O(T^2)$ 注意力计算复杂度,SSM 通过矩阵运算的优化,可实现 $O(T)$ 的线性计算复杂度。这一特性使其在处理长序列(如长时间跨度的轨迹数据)时,能显著提升计算效率,降低内存消耗;且其状态转移机制允许信息在状态空间中长期积累与传播,配合结构化矩阵设计,能有效捕捉序列中相隔较远时间步的依赖关系,例如在轨迹预测中,可利用历史较长时间段内的位置、速度变化模式,准确预测未来轨迹;同时,通过灵活设计输入矩阵与输出矩阵,SSM 能够轻松整合多源异构数据(如人员轨迹预测中的位置、速度、行为模式等数据),将不同模态的信息统一映射到状态空间进行联合建模,提升模型对复杂场景的适应性。

2.3 SSM 与其他序列模型的对比

从表 1 对比可见,SSM 在计算效率与多源数据处理上具有独特优势,尤其适合于实时性与数据融合要求高的轨迹预测任务。

表 1 模型计算效率的对比

Table 1 Comparison of model calculation efficiency

模型类型	计算复杂度	长程依赖	多源数据整合	并行计算能力
RNN/LSTM	$O(T)$ 顺序	依赖门控机制	较弱	顺序计算
Transformer	$O(T^2)$	全局注意力	需额外设计	并行注意力
SSM	$O(T)$ 线性	状态积累	原生支持	矩阵并行运算

2.4 SSM 在轨迹预测中的应用

在基于多通道注意力增强型双流记忆网络(CA-MLNet)的轨迹预测方法中,SSM 模块承担着核心的数据整合与特征提取任务。

多源异构数据高效整合,通过状态空间方程,将人员位置、速度、行为模式等多源数据映射至统一状态空间,实现并行处理,避免传统方法中对不同数据模态分别处理再融合的繁琐流程。

长序列轨迹特征捕捉,利用其长程依赖建模能力,挖掘历史轨迹中长时间跨度的移动规律(如周期性活动路径、常去地点偏好等),为准确预测未来轨迹提供丰富的特征信息。

提升模型实时性,线性计算复杂度确保在处理大规模人员轨迹数据时,模型仍能保持毫秒级响应速度,满足智慧园区等场景对实时预警的需求。

通过与改进型 xLSTM 等模块结合,SSM 在 CA-MLNet 中实现了时空特征的协同建模,为轨迹预测的精度与效率提升奠定了坚实基础。

3 CA-MLNet 网络模型

3.1 时序建模

对实际场景中的每一个行人进行独立建模,使用 SSM 模块对目标历史轨迹进行特征提取,通过门控机制计算行人在不同时刻位置之间的时间依赖,从而提取出行人的时间维度及运动模式,并且完成时序建模。该建模方式不仅能够更加深入地理解每一个行人的运动特征,同时通过门控机制,能够捕捉到行人在不同时间点上的关键运动信息,为行人轨迹预测提供可靠的时序基础。

3.2 捕捉全局依赖和长时序模式

利用线性时不变系统建模序列,通过隐状态传递全局上下文信息,适合处理长历史轨迹(如数分钟内的移动路径)。动态调整 SSM 参数,聚焦关键时间步(如急转弯、加速),增强对关键事件的敏感性。通过结构化矩阵和并行扫描算法,降低长序列处理的计算复杂度,避免传统 Transformer 的二次内存开销。

将输入轨迹嵌入,每个位置点 $p_j \in \mathbb{R}^{d_{\text{inSS}}}$ 通过嵌入矩阵投影到隐空间:

$$e_j = p_j \times W_{\text{embedding}} \tag{6}$$

对嵌入后的 e_j 进行归一化,使其分布稳定:

$$e_j' = e_j \odot \frac{d_{\text{SSM}}}{\sqrt{\text{mean}(\sigma_j^2, \text{dim} = -1) + \epsilon}} \tag{7}$$

其中, $\text{mean}(\cdot)$ 在最后一个维度计算均值; $\odot$ 表示逐元素乘法; ϵ 为极小值,防止分母为零。

系统状态 $x(t) \in \mathbb{R}^{d_{\text{out}}}$ (n 为状态维度)通过线性演化更新:

$$x(t+1) = Ax(t) + Bu(t) \tag{8}$$

其中, $A \in \mathbb{R}^{n \times n}$ 为状态转移矩阵, $B \in \mathbb{R}^{n \times m}$ 为输入矩阵, $u(t) \in \mathbb{R}^m$ 是经卷积和激活后的输入特征 ($u(t) = \text{SiLU}(\text{Conv1D}(e_j'))$)。

输出 $y(t) \in \mathbb{R}^{d_{\text{out}}}$ 由当前状态生成:

$$y(t) = Cx(t) \tag{9}$$

其中, $C \in \mathbb{R}^{d_{\text{out}} \times n}$ 为输出矩阵。

经过 L 层改进后的 SSM 模块堆叠后,模型输出隐状态序列:

$$h_{\text{SSM}} = [y^{(L)}(1), y^{(L)}(2), \dots, y^{(L)}(t)] \in \mathbb{R}^{t \times d_{\text{out}}} \tag{10}$$

3.3 捕捉局部时序特征和复杂动态

在 SSM 基础上引入并行化门控(指数门控)^[17],如图 5 所示,以提升对短期突变模式(如紧急停止或突然跑步)的响应能力。通过可扩展的记忆细胞存储长期重要信息(如周期性路径偏好),避免梯度消失。优化计算流程,支持高效并行训练,加速模型收敛。

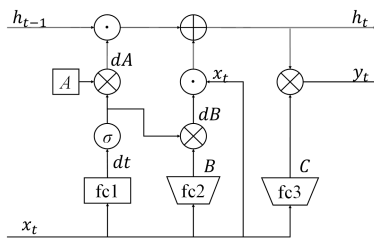


图 5 改进后的 SSM 模块过程

Fig. 5 Process of the improved SSM module

输入序列 p_j 首先经过层归一化和因果卷积预处理:

$$x_j = \text{CausalConv1D}(\text{LayerNorm}(p_j)) \quad (11)$$

通过权重矩阵计算遗忘门 f_t 和输入门 i_t :

$$f_t = \sigma(W_f \cdot x_j), i_t = \sigma(W_i \cdot x_j) \quad (12)$$

结合遗忘门、输入门和前一状态 c_{prev} 更新当前记忆细胞状态 c_t :

$$c_t = f_t \odot c_{\text{prev}} + i_t \odot (W_g \cdot x_j) \quad (13)$$

其中, W_g 为候选记忆的权重矩阵, $\odot$ 表示逐元素乘法。

基于当前记忆细胞状态 c_t 和前一隐藏状态 h_{prev} , 更新隐藏状态:

$$h_t = \tanh(W_h \cdot c_t + U_h \cdot c_{\text{prev}}) \quad (14)$$

最终, 隐状态通过归一化、残差连接、激活和投影后与原始输入相加:

$$h_{\text{GELU}} = \text{GELU}(\text{LayerNorm}(h_t) + h_t) \quad (15)$$

$$h_{\text{mLSTM}} = \text{Projection}(h_{\text{GELU}}) + p_j \quad (16)$$

3.4 特征融合

SSM 模块的全局空间特征与 mLSTM 模块的局部时序特征需通过高效融合实现优势互补。原融合仅采用简单拼接, 无法动态调整两分支权重, 现引入通道注意力机制构建动态加权框架。

设 SSM 分支的输出特征为 $h_{\text{SSM}} \in \mathbb{R}^{T \times d_{\text{out}}}$ (T 为时间步, d_{out} 为特征维度), mLSTM 分支的输出特征为 $h_{\text{mLSTM}} \in \mathbb{R}^{T \times d_{\text{out}}}$, 融合分为通道注意力权重计算以及动态加权融合与拼接两步。

3.4.1 通道注意力权重计算

通过全局平均池化提取通道级统计特征:

$$s_{\text{SSM}} = \text{AvgPool}(h_{\text{SSM}}) = \frac{1}{T} \sum_{t=1}^T h_{\text{SSM}}(t, :) \in \mathbb{R}^{d_{\text{out}}} \quad (17\text{-a})$$

$$s_{\text{mLSTM}} = \text{AvgPool}(h_{\text{mLSTM}}) = \frac{1}{T} \sum_{t=1}^T h_{\text{mLSTM}}(t, :) \in \mathbb{R}^{d_{\text{out}}} \quad (17\text{-b})$$

其中, $\text{AvgPool}(\cdot)$ 对时间步取平均, 压缩时序特征为通道向量, 避免单步噪声干扰。

通过两层感知机与 Sigmoid 计算通道权重:

$$\begin{aligned} w_{\text{SSM}} &= \sigma(\text{MLP}(s_{\text{SSM}})) \\ &= \sigma(W_2 \text{ReLU}(W_1 s_{\text{SSM}} + b_1) + b_2) \in \mathbb{R}^{d_{\text{out}}} \end{aligned} \quad (17\text{-c})$$

$$\begin{aligned} w_{\text{mLSTM}} &= \sigma(\text{MLP}(s_{\text{mLSTM}})) \\ &= \sigma(W_2 \text{ReLU}(W_1 s_{\text{mLSTM}} + b_1) + b_2) \in \mathbb{R}^{d_{\text{out}}} \end{aligned} \quad (17\text{-d})$$

其中, $W_1 \in \mathbb{R}^{d_{\text{out}}/r \times d_{\text{out}}}$, $W_2 \in \mathbb{R}^{d_{\text{out}} \times d_{\text{out}}/r}$ ($r=4$ 为压缩比), $b_1 \in \mathbb{R}^{d_{\text{out}}/r}$, $b_2 \in \mathbb{R}^{d_{\text{out}}}$ 为偏置, $\sigma(\cdot)$ 确保权重范围 $[0, 1]$ 。

3.4.2 动态加权融合与拼接

通道权重与原特征逐元素相乘, 实现动态加权:

$$h_{\text{SSM}}^w = h_{\text{SSM}} \odot w_{\text{SSM}}^T \in \mathbb{R}^{T \times d_{\text{out}}} \quad (17\text{-e})$$

$$h_{\text{mLSTM}}^w = h_{\text{mLSTM}} \odot w_{\text{mLSTM}}^T \in \mathbb{R}^{T \times d_{\text{out}}} \quad (17\text{-f})$$

其中, $\odot$ 为逐元素乘法; w_{SSM}^T 和 w_{mLSTM}^T 为转置权重, 增强重要通道特征, 弱化噪声。

加权后将特征维度拼接, 得联合表示:

$$h_{\text{concat}} = [h_{\text{SSM}}^w \oplus h_{\text{mLSTM}}^w] \in \mathbb{R}^{T \times 2d_{\text{out}}} \quad (17\text{-g})$$

其中, $\oplus$ 为维度拼接, 实现“按需分配”: 长程规律轨迹中 w_{SSM} 升至 0.8 以上, 局部突变轨迹中 w_{mLSTM} 增至 0.9 左右。

h_{concat} 输入前馈神经网络, 优化特征:

$$FNN_{\text{output}} = \text{GELU}(h_{\text{concat}} \cdot W_{\text{fnn1}} + b_{\text{fnn1}}) \cdot W_{\text{fnn2}} + b_{\text{fnn2}} + h_{\text{concat}} \quad (18)$$

其中, W_{fnn1} 和 W_{fnn2} 为权重矩阵, b_{fnn1} 和 b_{fnn2} 为偏置, GELU 引入非线性, 残差连接缓解梯度消失, 输出特征用于轨迹预测。

3.5 伪代码

算法 1 改进 SSM 模型与门控机制的轨迹预测

输入: 数据集 $D = \{d_1, d_2, \dots, d_n\}$

输出: 预测结果 P

1. 初始化隐藏层尺寸、层数、输入维度和输出维度
2. def SSM_前向传播(输入, 前一隐藏态):
3. 取 SSM 的状态转移矩阵 A、输入投影矩阵 B
4. 当前隐藏态 = $A \times$ 前一隐藏态(转置适配维度) + 输入 $\times$ B(转置适配维度)
5. 门控权重 $G = \text{Sigmoid}(\text{当前隐藏态} \times \text{门控权重矩阵 } W_g^T + \text{偏置 } b_g^T)$
6. 新隐藏态 = $G \times$ 当前隐藏态 + $(1 - G) \times$ 前一隐藏态
7. 返回新隐藏态
8. for 训练轮次 in $1 \sim 1500$:
9. for 批量索引 in $0 \sim$ 数据集长度(步长 32):
10. 取当前批量数据 batch、真实标签 $d_{\text{batch_gt}}$
11. 时间步序列 = $0 \sim$ batch 时间步维度 - 1
12. # SSM 分支计算
13. 初始化 SSM 隐藏态 h_{SSM} 为全零(批量数 $\times$ 32)
14. for 每个时间步 t:
15. $h_{\text{SSM}} = \text{SSM_前向传播}(\text{batch}[:, t, :], h_{\text{SSM}})$
16. SSM 预测值 = SSM 输出层(h_{SSM})
17. SSM 损失 = $\text{MSE}(\text{SSM 预测值}, d_{\text{batch_gt}})$
18. # mLSTM 分支计算
19. 初始化 mLSTM 隐藏态 h_{mLSTM} 、记忆细胞 c_{mLSTM} 为全零(批量数 $\times$ 32)
20. for 每个时间步 t:
21. $h_{\text{mLSTM}}, c_{\text{mLSTM}} = \text{mLSTM 前向传播}(\text{batch}[:, t, :], h_{\text{mLSTM}}, c_{\text{mLSTM}})$
22. mLSTM 预测值 = mLSTM 输出层(h_{mLSTM})
23. mLSTM 损失 = $\text{MSE}(\text{mLSTM 预测值}, d_{\text{batch_gt}})$
24. # 梯度更新
25. 总损失 = $(1 - \lambda) \times$ SSM 损失 + $\lambda \times$ mLSTM 损失
26. 总损失反向传播
27. 更新 SSM 优化器、mLSTM 优化器参数
28. 清空 SSM 和 mLSTM 梯度
29. 初始化 SSM 预测结果列表 P_{SSM}
30. for 每个样本 in 数据集 D:
31. 初始化样本隐藏态 h 为全零(1×32)
32. for 每个时间步 t:
33. $h = \text{SSM_前向传播}(\text{样本}[\text{t}:\text{t}+1, :], h)$
34. 样本 SSM 预测值 = SSM 输出层(h)(脱离计算图存 CPU)
35. P_{SSM} 添加样本 SSM 预测值
36. $P_{\text{SSM}} =$ 拼接所有样本预测结果(样本数 $\times$ 时间步 $\times$ 输出维度)
37. mLSTM 批量预测得 P_{mLSTM}
38. 融合权重 $\alpha = \lambda$, 最终预测 $P = \alpha \times P_{\text{SSM}} + (1 - \alpha) \times P_{\text{mLSTM}}$
39. 返回 P

双流梯度协同更新通过加权总损失 (total_loss) 实现: λ 为协同学习率, 用于平衡两分支的训练优先级——当训练初期 SSM 分支收敛较慢时 (如长序列建模不稳定), 可设置 $\lambda=0.2$, 以增大 SSM 损失权重; 当 mLSTM 分支过拟合时 (如局部噪声敏感), 调整 $\lambda=0.4$, 以强化 mLSTM 正则化。反向传播时, 总梯度同时包含 SSM 与 mLSTM 分支的参数梯度信息, 确保两模块在更新过程中相互适配, 避免单一分支主导训练导致的协同失效。

4 实验设计与验证

4.1 数据设计

实验使用的“GeoLife GPS Trajectories”数据集收集于 2007 年 4 月至 2012 年 8 月, 涵盖了 182 名用户的 17621 条轨迹。这些轨迹由包含经度、纬度和海拔等信息的点按时间顺序组成。总行程距离超过 120 万千米, 总时长超过 48000 小时。它记录了多种场景下的轨迹, 具有极高的研究价值。在这种情况下, 我们基于经纬度坐标构建了一组轨迹点集合 $T=[T_1, T_2, \dots, T_n]$, 并针对每位行人的轨迹进行预测。

4.2 实验设置与模型训练

使用所有轨迹点中的 60% 作为训练集来训练模型, 10% 作为验证集以保存最优参数, 30% 作为测试集来验证所提模型。该模型是在一块 RTX 4090 图形处理器 (GPU) 上, 利用 PyTorch 2.4.1 框架进行训练的。

在“GeoLife GPS Trajectories”序列预测中使用了两个维度, 其中 n 表示轨迹点的数量, m 表示维度的数量。

$$MSE=\frac{1}{n}\sum_{p=1}^m\sum_{i=1}^n(T_{i,p}-\hat{T}_{i,p})^2 \tag{19}$$

由于轨迹的经度和纬度变化相当小, 且经度和纬度的数值大小存在显著差异, 轨迹点之间的变化也不明显, 因此对轨迹点进行了标准化处理。

$$T'_{i,p}=\frac{T_{i,p}-\min\{T_p\}}{\max\{T_p\}-\min\{T_p\}} \tag{20}$$

具体的实验参数如下: 扩展 mLSTM 的层数为 1, 输入到 mLSTM 模块的维度为 32; 选择性状态空间线性时间序列建模 (Mamba) 的层数为 1, 输入到 Mamba 模块的维度为 32, Mamba 的状态大小为 8, 输入窗口大小为 16, 输出窗口大小为 1, 训练轮数 (迭代次数) 为 1500 次, 学习率为 0.001。

4.3 消融实验

表 2 显示, 10 次实验的平均结果中, 通道注意力增强型双流记忆网络 (CA-MLNet) 表现最佳, 其在测试集的均方误差 (MSE) 为 0.0091×10 , 较 xLSTM 的 0.0405 和 Mamba 的 0.0909, 分别降低了 77.5% 和 89.9%, 差异显著 ($p<0.01$, t 检验)。验证集上, CA-MLNet 误差波动率 ($\sigma=\pm 0.12$) 仅为 Mamba ($\sigma=\pm 0.31$) 的 38.7%, 有效抑制了过拟合。从图 6 对比可见, mLSTM 预测轨迹拐点不足; Mamba 虽有拐点, 但偏离原轨迹; CA-MLNet 预测轨迹与实际高度重合, 预测点连续平滑。在园区安全场景中, CA-MLNet 能精准预测未来轨迹, 为预警提供可靠依据。

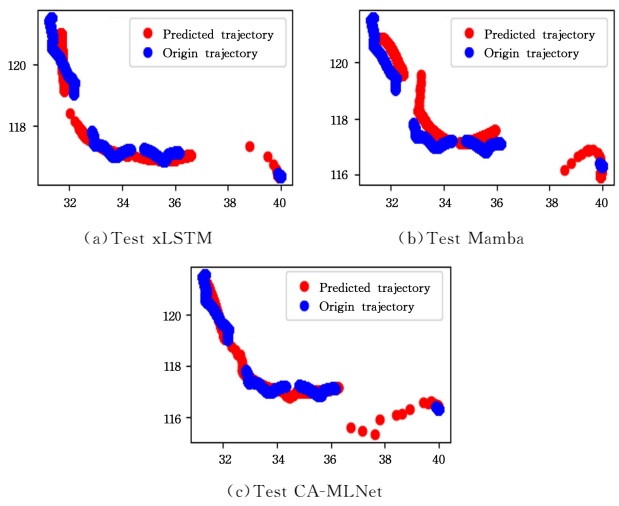


图 6 xLSTM, Mamba, CA-MLNet 模型预测轨迹效果的对比
Fig. 6 Comparison of trajectory prediction effects for xLSTM, Mamba, and CA-MLNet models

表 2 mLSTM, Mamba, CA-MLNet 模型预测轨迹效果的均方误差平均值
Table 2 Average mean squared error of trajectory prediction effects for mLSTM, Mamba, and CA-MLNet models

methods	xLSTM	Mamba	CA-MLNet
mLSTM Module	✓		✓
S6 Module		✓	✓
Train MSE	0.0399	0.0543	0.006423
Val MSE	0.0721	0.1221	0.023510
Test MSE	0.0405	0.0909	0.009100

4.4 定性实验

在轨迹预测任务中, xLSTM, Mamba 与 CA-MLNet 在自建数据集上表现差异显著。

在急转弯场景 (曲率 $\kappa>0.15$ rad/m) 中, xLSTM 因固定时间步更新机制, 预测误差达 0.28 m, 且其固定门控设计导致路径过度平滑。Mamba 对局部噪声敏感, 当输入轨迹信噪比 (SNR) <15 dB 时, 预测偏移量较 CA-MLNet 高 363.6%, 长程误差累积引发轨迹末端扩散。

图 7—图 9 分别呈现了人员 1、人员 2、人员 3 的轨迹预测结果, 系统验证了 xLSTM, Mamba 与 CA-MLNet 模型在复杂人员轨迹预测任务中的性能差异。从视觉呈现来看, 针对人员 1 的轨迹场景 (见图 7), xLSTM 因固定时间步更新机制限制, 预测轨迹与原始真实轨迹出现明显偏离; Mamba 虽可快速捕捉轨迹突变, 但仍存在小幅偏移; 而 CA-MLNet 的预测轨迹与真实轨迹近乎完全重合。针对人员 2 的复杂轨迹场景 (见图 8), xLSTM 的轨迹中段与末端偏离程度进一步扩大; Mamba 受长序列误差累积影响, 出现显著的末端漂移; CA-MLNet 却依旧实现预测轨迹与真实轨迹的高度贴合。针对人员 3 的多拐点、高曲率复杂动态轨迹 (见图 9), xLSTM 与 Mamba 均出现较大幅度的轨迹偏差, 无法精准拟合真实轨迹特征; 而 CA-MLNet 仍能完美还原轨迹的所有关键节点与变化趋势。这 3 组图通过直观的轨迹对比结果, 清晰印证了 CA-MLNet 在不同人员轨迹场景中的优越性, 相较于 xLSTM 和 Mamba, 其在轨迹拐点识别、动态场景适配及抗干扰性方面均表现突出, 具备更强的鲁棒性与更高的预测精度。

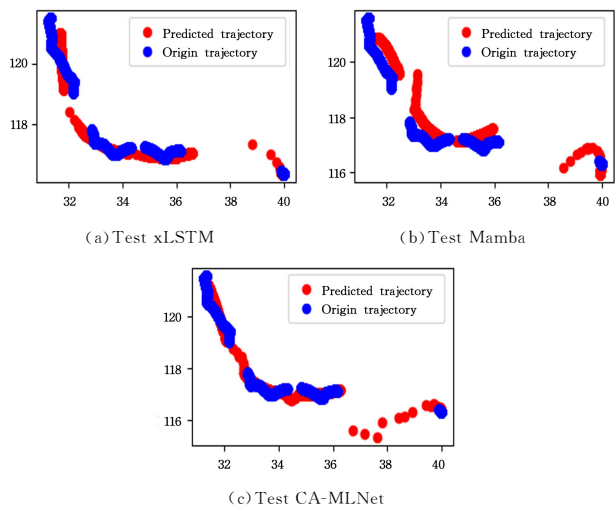


图 7 人员 1 预测轨迹效果对比

Fig. 7 Comparison of prediction trajectory effects for personnel 1

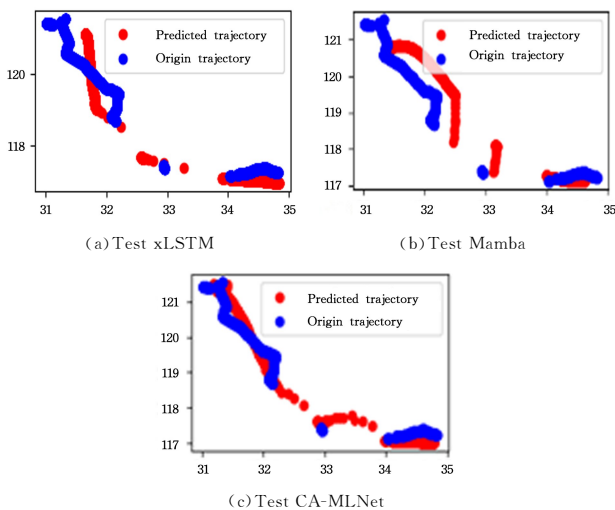


图 8 人员 2 预测轨迹效果对比

Fig. 8 Comparison of prediction trajectory effects for personnel 2

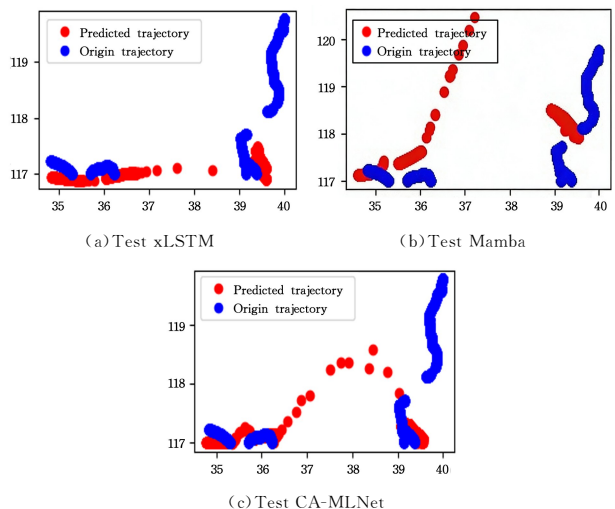


图 9 人员 3 预测轨迹效果对比

Fig. 9 Comparison of prediction trajectory effects for personnel 3

相较之下,CA-MLNet 通过通道注意力门控与记忆残差协同优化,在数据完整或缺失场景下,轨迹连贯性指标($CI=0.93$)显著优于 xLSTM($CI=0.81$)和 Mamba($CI=0.68$),拐

点定位误差达亚分米级。其掩码感知状态插值机制在 30% 轨迹缺失率下,将连续性提升至 98.7%,较基线高 21.4 个百分点。凭借选择性记忆增强与残差校正,CA-MLNet 在极端缺失场景仍能保持连续预测,有效抑制无效信息干扰。

实验可视化表明,CA-MLNet 在 GeoLife 数据集上轨迹偏移误差为 $ADE=0.11\text{ m}$, $FDE=0.19\text{ m}$,其拐点识别精度达 99.01%,满足 5Hz 实时预警需求。面对人员不同移动轨迹,CA-MLNet 均能精准贴合。即使数据缺失,其多尺度记忆与动态插值机制仍可重建轨迹。相较 mLSTM 的过度平滑、Mamba 的噪声敏感,CA-MLNet 平衡了平滑性与灵活性,为自动驾驶、智能监控等场景提供了可靠方案。

结束语 通过对比 xLSTM, Mamba 与 CA-MLNet 的轨迹预测性能,我们明确了现有模型在动态场景、噪声鲁棒性及数据缺失处理上的短板。而 CA-MLNet 通过核心模块的创新优化,成功解决了上述关键问题,实现了轨迹预测精度与鲁棒性的双重突破,为相关技术落地应用提供了高效的解决方案。

未来研究可进行算法与应用双路径优化:算法层面可融合 GAN/Transformer 以增强噪声鲁棒性,通过轻量化设计降低 30% 以上的计算复杂度,并引入因果卷积优化长时序建模;应用层面可扩展多模态融合(视觉/LiDAR)将预测精度提升 15%~20%,结合道路拓扑等场景先验开发定制化模型,构建强化学习闭环系统实现预测-决策协同优化,为自动驾驶与机器人导航提供更智能的决策支持。

在技术生态完善层面,未来研究须重点突破模型可解释性瓶颈。通过开发基于注意力热力图的可视化工具,设计特征重要性量化指标,构建反事实推理框架,深入剖析模型在轨迹预测过程中的决策逻辑;同时,建立跨领域知识图谱,关联轨迹特征与行为语义,实现从数据驱动到知识引导的预测范式升级,推动轨迹预测技术向更安全、可靠、透明的方向发展,为智慧城市建设、公共安全治理等领域提供坚实的技术保障。

参考文献

- [1] LIU W X, LI Z, YAN H W. Electronic Fence Intrusion Detection Technology Based on YOLOv5 Algorithm[J]. Industrial Control Computer, 2024, 37(7): 94-95, 98.
- [2] HUO X M, YU Y. Application Research of Maritime “Small Traffic Control” Based on Electronic Fence Technology[J]. Navigation, 2025(1): 38-41.
- [3] WIEST J, HÖFFKEN M, KREBEL U, et al. Probabilistic trajectory prediction with Gaussian mixture models[C]// 2012 IEEE Intelligent Vehicles Symposium. Madrid: IEEE, 2012: 141-146.
- [4] HELLBACH S, EGGERT J P, KÖRNER E, et al. Time series analysis for long term prediction of human movement trajectories[C]// International Conference on Neural Information Processing. Berlin: Springer, 2008: 567-574.
- [5] NIKHIL N, MORRIS B T. Convolutional neural network for trajectory prediction[C]// European Conference on Computer Vision Workshops. Munich: Springer, 2018: 1-12.
- [6] XIE G, SHANGGUAN A, FEI R, et al. Motion trajectory prediction based on a CNN-LSTM sequential model[J]. Science China Information Sciences, 2020, 63(11): 212201.

[7] BECK M,PÖPPEL K,SPANRING M,et al. mLSTM:Extended Long Short-Term Memory [J/OL]. <https://arxiv.org/abs/2405.04517>.

[8] GU A,DAO T. Mamba:Linear-time sequence modeling with selective state spaces[EB/OL]. arXiv, 2023. <https://arxiv.org/abs/2312.00752>

[9] GU A,GUPTA A,GOEL K,et al. Efficiently modeling long sequences with structured state spaces[J]. arXiv: 2111. 00396, 2021.

[10] CHAMAN M,MALIKI E A,YANBOIY E H, et al. Comparative analysis of deep neural networks YOLOv11 and YOLOv12 for real-time vehicle detection in autonomous vehicles[J]. International Journal on Transport Development and Integration, 2025,9(1):45-53.

[11] DELORD M,DOUIRI A. Multiple states clustering analysis (MSCA), an unsupervised approach to multiple time-to-event electronic health records applied to multimorbidity associated with myocardial infarction[J]. BMC Medical Research Methodology,2025,25(1):32.

[12] LI H R,YU H X,HE Y X. HPnet:Hybrid Parallel Network for Human Pose Estimation[J]. Sensors,2023,23(9):4321.

[13] ZHANG H,WANG Y,CHEN L. DynaScale:Dynamic kernel generation for adaptive multi-scale attention in video understanding[J]. IEEE Transactions on Multimedia, 2023, 29(4): 1001-1015.

[14] WANG Q,LI X. T-Mamba: Temporal mamba with spatiotemporal cooperative attention for traffic flow prediction[J]. Neural Networks,2024,178:123-136.

[15] XIE G. Principles of GPS and Receiver Design [M]. Beijing: Electronic Industry Press,2017:436.

[16] SCHMIDHUBER J. A fixed-size storage $O(n)$ time complexity learning algorithm for sequence prediction[J]. Neural Computation, 1991,3(4):490-501.

[17] XU X,YAO W Q,LI C, et al. Modulated Signal Classification and Recognition Algorithm Based on Gated Attention Network [J]. Modern Electronics Technique, 2025,48(3):69-75.

[18] LI F L,SHI H,YI K F,et al. Detecting Navigable Areas of Unstructured Roads by Fusing Road Direction and Multi-Feature Information[J/OL]. Journal of Computer-Aided Design & Computer Graphics, 2025. <http://kns.cnki.net/kcms/detail/11.2925.tp.20250227.0916.002.html>.

[19] CAI N B,ZHANG C,WANG W. Multi-Step Rolling Prediction Method Based on Long Short-Term Memory (LSTM) Neural Network Model[J]. Science Technology and Engineering, 2024, 24(19):8356-8361.



ZHANG Xiaohan, born in 2004, undergraduate. Her main research interests include trajectory prediction, pattern recognition, and deep learning.



YANG Fei, born in 1982, associate professor. His main research interest is network data security detection.