

基于 EAC-GWO-Attention-LSTM 模型的短时交通流量预测模型探究——以嘉兴市为例

摘 要

在大数据时代背景下，大数据与前沿技术的深度融合在短时交通流量预测的应用，对优化交通信号控制、规划交通基础设施建设等方面具有重要意义。为提升短时交通流量预测的准确性，本文建立 EAC-GWO-Attention-LSTM 模型，从多模型对比视角出发，以嘉兴市交通流数据为研究对象，探究该模型的性能，同时对比得到最适合该市短期交通流量预测的模型。

首先，本研究对嘉兴市智慧交通研究中心所提供的开源交通流量数据集做预处理，将同一方向的车道数据合并。考虑交通流量随时间逐渐变化，少量异常值的替换对预测结果影响不大，因此本文将异常值替换为异常值附件两个正常值的平均值。

接着，考虑时间序列特征信息的重要性和自相关误差的影响，本文在 LSTM 模型中，引入注意力机制，利用自相关误差补偿法，同时结合 GWO 算法对注意力机制的初始权重参数进行优化，建立 EAC-GWO-Attention-LSTM 模型。借助 python 软件，将该模型与 ARIMA、RNN、LSTM、Attention-LSTM、GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 进行模拟对比，验证模型的性能。

通过实验数据的模拟分析，验证了 EAC-GWO-Attention-LSTM 的 MAE、RMSE 的预测误差值明显低于另外六个模型。在 12 个朝向的路线中，与其他模型相比，该模型的 MAE 值(辆)分别平均降低了 15.95、6.93、4.67、4.56、2.58、3.73；RMSE 值(辆)分别平均降低了 17.21、11.10、7.49、7.96、3.81、5.73。同时，该模型能够准确地预测实际交通流量时间序列的趋势，模型的预测值折线图与实际值的折线图基本一致、预测点与几个实际点的峰值高度接近，可见该模型具有良好的非线性拟合能力预测准确性。此外，通过模型优化分析证明，EAC-GWO-Attention-LSTM 更易于训练。

因此，将 EAC-GWO-Attention-LSTM 应用于当地的短时交通流量预测，能够在一定程度上提高该城市交通管理的智能化水平，为出行者提供实施可靠的交通信息，从而有效缓解交通拥堵、优化出行路径以及减少汽车尾气的排放。

关键词：多模型对比 GWO 算法 EAC 短时交通流量预测 EAC-GWO-Attention-LSTM

目录

摘 要	I
表格与插图清单	IV
一、 绪论	1
(一) 研究背景及意义	1
(二) 国内外研究现状	1
(三) 研究内容	2
(四) 研究路线	3
(五) 研究创新	4
1. 多模型融合优化短时交通流量预测模型	4
2. 多模型对比视角下 EAC-GWO-Attention-LSTM 模型短时交通流量预 测效果探究	4
二、 数据来源与预处理	5
(一) 交通数据来源	5
(二) 交通数据预处理	6
三、 预备知识	7
(一) ARIMA 模型	7
(二) RNN 模型	8
(三) LSTM 模型	9
(四) Attention-LSTM 模型	9
(五) 自相关误差	11
四、 基于 GWO-Attention-LSTM 的短时交通流量预测模型	12
(一) GWO 算法	12
(二) GWO-Attention-LSTM 模型	13
五、 基于 EAC-Attention-LSTM 和 GWO-Attention-LSTM 的短时交通流量预测模 型	15
(一) 自相关误差补偿方法	15
(二) EAC-Attention-LSTM 模型	16

(三) EAC-GWO-Attention-LSTM 模型	17
六、 模型模拟与分析	18
(一) 预测模型评估指标	18
1. 平均绝对误差 MAE	18
2. 均方根误差 RMSE	18
(二) 预测模型参数设置	18
(三) 预测结果分析	20
(四) 优化分析	22
1. GWO 优化分析	22
2. EAC 优化分析	23
3. Epochs 和损失函数曲线对比	24
七、 研究结论与反思	26
(一) 研究结论	26
(二) 研究反思	26
八、 参考文献	28
附录 1 各模型具有参数设置表	30
附录 2 各模型不同方向的 MAE、RMSE 误差值	32
附录 3 EAC-GWO-Attention-LSTM 与其他各模型预测误差的对比	34
附录 4 GWO 优化分析	36
附录 5 EAC 优化分析	37
附录 6 Epochs 和损失函数曲线对比	38

表格与插图清单

表 1 合并后的数据部分展示

表 2 部分清理后的相关输入数据示例

表 3 ARIMA 各朝向数据的参数设置

表 4 各朝向收敛迭代次数及稳定值

表 5 RNN、LSTM、Attention-LSTM 模型参数设置

表 6 GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM
模型参数设置

表 7 各模型不同方向的 MAE 误差 (辆)

表 8 各模型不同方向的 RMSE 误差 (辆)

表 9 EAC-GWO-Attention-LSTM 与其他模型不同方向的 MAE 误差 (辆) 减少值

表 10 EAC-GWO-Attention-LSTM 与其他模型不同方向 RMSE 误差 (辆) 减少值

图 1 研究路线图

图 2 三个相邻交叉口渠化简图

图 3 ARIMA 原理图

图 4 RNN 逻辑结构

图 5 LSTM 细胞结构

图 6 Attention-LSTM 模型结果图

图 7 灰狼的社会等级

图 8 GWO-Attention-LSTM 流程图

图 9 EAC-Attention-LSTM 流程图

图 10 EAC-GWO-Attention-LSTM 流程图

图 11 所有模型预测误差比较(MAE、RMSE)

图 12 预测结果与实际值对比(依次为 crossing_1_east/north/south/west)

图 13 预测结果与实际值对比(依次为 crossing_2_east/north/south/west)

图 14 预测结果与实际值对比(依次为 crossing_3_east/north/south/west)

图 15 迭代次数与 RMSE 变化过程(依次为 crossing_1_south gwo_rmse、crossing_2_west_gwo_rmse、crossing_3_south_gwo_rmse)

图 16 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

图 17 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

图 18 Attention-LSTM(依次为 crossing_3_east/north/south/west)

图 19 GWO-Attention-LSTM(依次为 crossing_3_east/north/south/west)

图 20 EAC-Attention-LSTM 依次为(crossing_3_east/north/south/west)

图 21 EAC-GWO-Attention-LSTM 依次为(crossing_1_east/north/south/west)

图 22 迭代次数与 RMSE 变化过程(依次为 crossing_1_east/north/south/west_gwo_rmse)

图 23 迭代次数与 RMSE 变化过程(依次为 crossing_2_east/north/south/west_gwo_rmse)

图 24 迭代次数与 RMSE 变化过程(依次为 crossing_3_east/north/south/west_gwo_rmse)

图 25 EAC 随 Epoch 收敛变化曲线依次为(crossing_1_east/north/south/west)

图 26 EAC 随 Epoch 收敛变化曲线依次为(crossing_2_east/north/south/west)

图 27 EAC 随 Epoch 收敛变化曲线依次为(crossing_3_east/north/south/west)

图 28-30 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_1、crossing_2、crossing_3 的四个朝向)

图 31-33 LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_1、crossing_2、crossing_3 的四个朝向)

图 34-36 Attention-LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_1、crossing_2 和 crossing_3 的四个朝向)

图 37-39 GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线(crossing_1、crossing_2 和 crossing_3 的四个朝向)

图 40-42 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线(crossing_1、crossing_2 和 crossing_3 的四个朝向)

图 43-45 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线(crossing_1、crossing_2 和 crossing_3 的四个朝向)

基于 EAC-GWO-Attention-LSTM 模型的短时交通流量预测模型探究——以嘉兴市为例

一、绪论

（一）研究背景及意义

在智慧城市的众多组成部分中，交通系统作为城市运行的基础，提高其智能化水平对缓解交通拥堵、减少交通事故、提高出行效率具有至关重要的作用。《数字交通“十四五”发展规划》目标下，我国通过交通新基建，提升数字化、网络化、智能化水平，支撑交通运输行业高质量发展和交通强国建设^[1]。

城市交通系统是支撑城市快速发展的^[10]重要基础，对城市的经济发展有着显著影响。交通问题关系到人们的日常出行。嘉兴市作为连接长三角南北的“黄金走廊”，其公路、航道密度均跃居浙江省首位。其积极响应国家促进交通高质量发展的需求，以先进信息技术赋能交通运输发展，强化交通数字治理，逐步提升大数据运用能力，构建智能交通系统，提升城市运行效率、支撑区域协同发展，提升交通运输决策分析水平。

然而，无论是对于交通控制还是交通诱导系统来说，实时准确的交通流量预测是这些系统实现的前提及关键，交通流量预测结果的好坏直接关系到交通控制与诱导的效果。在当今大数据和人工智能技术迅猛发展的时代背景下，大数据与前沿技术的深度融合在短时交通流量预测的应用，对优化交通信号控制、制定交通管理策略、规划交通基础设施建设等方面具有重要意义。

（二）国内外研究现状

在大数据时代背景下，机器学习模型能够有效地处理交通流量预测中的多组道路数据，但缺乏对交通流量时序特征的深度挖掘，难以捕捉时间序列中的隐藏模型。随着神经网络的发展，基于深度学习算法的预测模型应运而生^[11]。

深度学习算法擅长深度挖掘数据中的特征信息，并具有更强的非线性拟合能力。其中循环神经网络（RNN）主要用于处理序列数据，神经元在某时刻的输出可以作为输入再次输入到神经元的串联网结构可以保持数据中的依赖关系，但大量的实践也

表明标准的 RNN 难以实现信息的长期保存^[2]。LSTM 作为 RNN 的变体，其具有记忆功能，能够获取历史信息的时间特征，Ma et al.(2015)^[12]、Cao et al.(2018)^[13]和 Zhong et al.(2019)^[14]都使用 LSTM 构建了短期交通流预测模型。但由于长时间序列的短子序列中特征重要程度存在差异，重要的显著特征往往会包含更多信息量，对实际需求量的趋势影响程度更大。假若赋予模型更关注高重要度特征的能力，可以更好实现短期模型的有效提取和 LSTM 输入信息的优化^[3]，提升模型动态结构输出质量。

然而，注意力机制的初始权重参数都是随机选取的，在模型迭代过程中，权重值会根据设定的参数和输入信息进行调整，因此存在不确定性。张晓静^[15]等人将注意力机制与 LSTM 相结合，将高维空间中灰狼优化算法(GWO)对注意力机制的初始权重参数进行优化，建立基于 GWO-Attention-LSTM 的短时交通流预测模型，设置了多组对比模型进行预测分析，结果表明该模型相较于其他模型具有明显优势。

（三）研究内容

在交通流量预测的实际应用中，通常难以获取如天气、出行兴趣点等全面的数据，在大多数情况下可以获取单个输入因素，因此有必要强化预测模型，在模型中加入注意力机制。同时通过权重分配的方法，对从每个单元提取的特征信息的重要性进行评分，以提升模型对重要信息的关注度。

然而，时间序列中预测误差各项之间存在于时间序列数据中，因为相邻的时间点之间可能存在一些未知的共同因素，从而导致误差之间存在一定的相关性。同时交通流残差天然地具有拖尾效应，并且传感器漂移、信号配时延迟、上下游耦合等因素，会使得相邻时刻预测误差往往呈自相关（EAC）。

为此，本文以嘉兴市的交通流量数据为研究对象，考虑时间序列特征信息的重要性和自相关误差的影响，在传统的 LSTM 模型中，利用自相关误差补偿法，同时结合 GWO 算法对注意力机制的初始权重参数进行优化，建立 EAC-GWO-Attention-LSTM 模型。

为验证该模型的性能，本研究设置了多组对比模型进行预测分析，模拟分析对比 ARIMA、RNN、LSTM、Attention-LSTM、GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 这七个模型的预测性能，探究 EAC-GWO-Attention-

LSTM 在该市短期交通流量预测的预测情况，对比得到最适合该市短期交通流量预测的模型，进而帮助提高该城市交通管理的智能化水平，为出行者提供实施可靠的交通信息，从而有效缓解交通拥堵、优化出行路径以及减少汽车尾气的排放。

（四）研究路线

本文的研究路线见图 1。

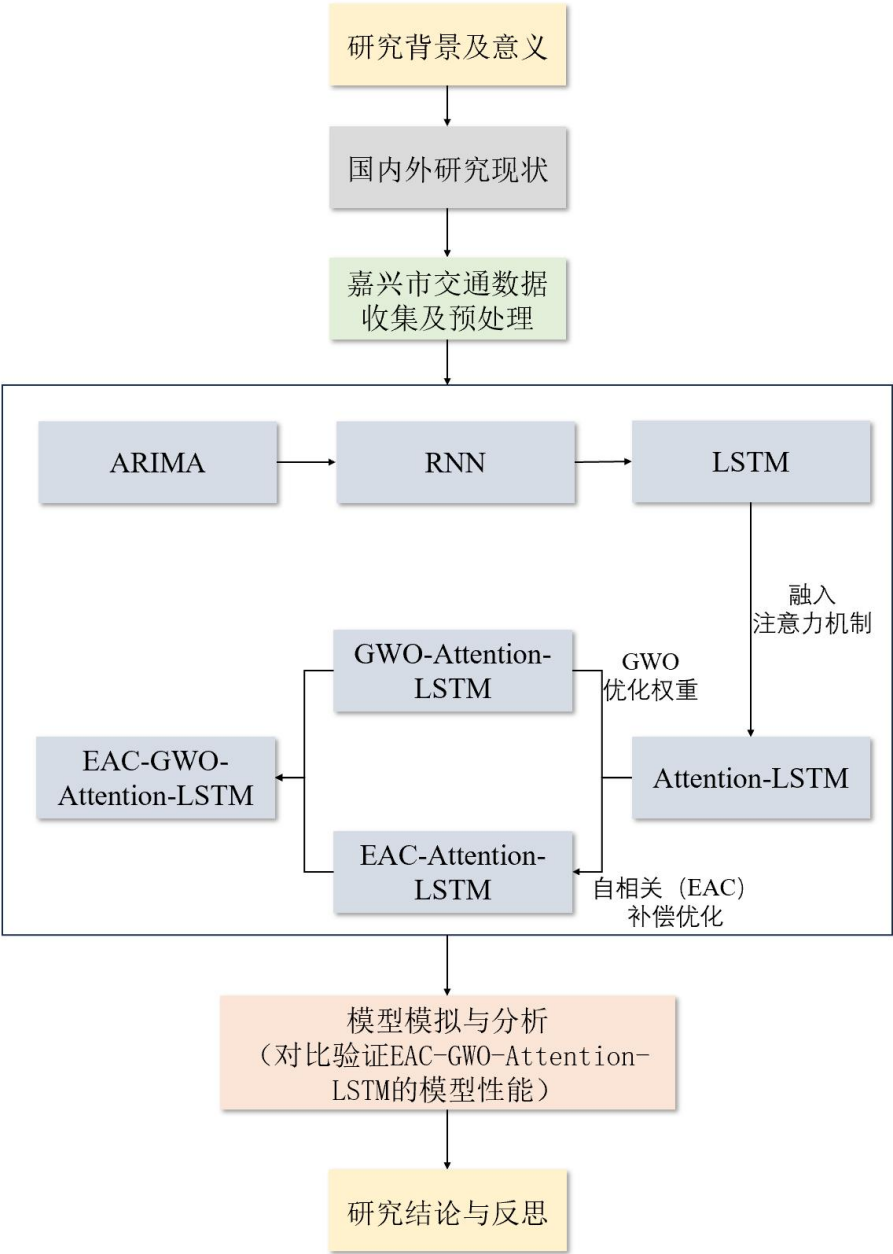


图 1 研究路线图

（五）研究创新

1. 多模型融合优化短时交通流量预测模型

由于单个深度学习模型的存在难以实现长信息的保存、特征重要程度存在差异等问题，因此本文考虑多模型融合的方式优化传统的短时交通流量预测模型。在传统的 LSTM 模型中，利用自相关误差补偿法，缓解交通流残差的拖尾效应，减少相邻时刻预测误差自相关的，同时结合 GWO 算法对注意力机制的初始权重参数进行优化，建立 EAC-GWO-Attention-LSTM 模型，探究该模型预测嘉兴市短时交通流量的预测性能。

2. 多模型对比视角下 EAC-GWO-Attention-LSTM 模型短时交通流量预测效果探究

本文通过分析现有的国内外研究，总结现有常用的短时交通流量预测模型。从多模型对比视角出发，结合嘉兴市实际交通拥堵状况，借助 python，对比 ARIMA、RNN、LSTM、Attention-LSTM、GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 这七个模型的预测性能，探究 EAC-GWO-Attention-LSTM 在该市短期交通流量预测的预测情况，对比得到最适合该市短期交通流量预测的模型，进而帮助提高该城市交通管理的智能化水平，为出行者提供实施可靠的交通信息。

二、数据来源与预处理

（一）交通数据来源

本研究数据来源于嘉兴市智慧交通研究中心所提供的开源交通流量数据集，包含三个连续路口在两个月内每周三（一周中相对车流稳定的时间点）各车道级别的不同车型统计车流数据，数据包括车辆进口方向、统计时间戳（统计时间单位为 15 分钟）、车道号、大、小型车辆数。采集时间依次为 2023 年 1 月 4、11、18、25 日，2 月 1、8、15、22 日，共 8 天，28416 条的交通流量数据。

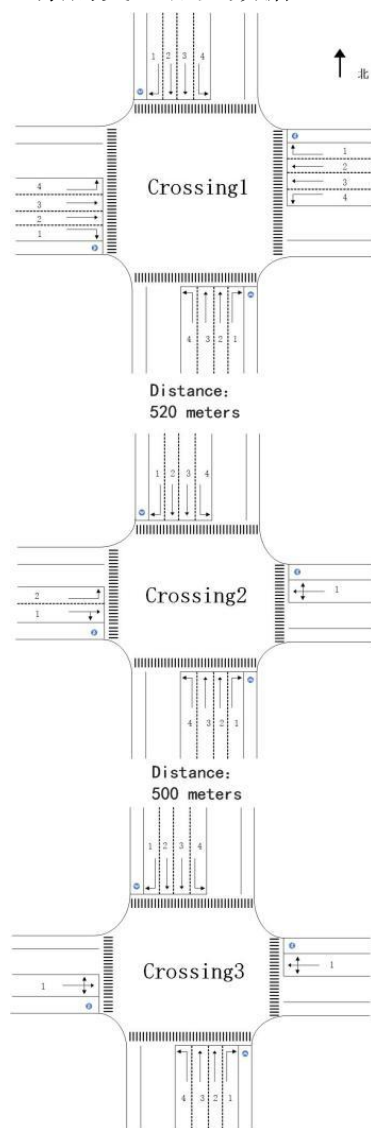


图 2 三个相邻交叉口渠化简图

注：本次数据集来源网址为 <https://strc.zjxu.edu.cn/info/1142/1104.htm>

（二）交通数据预处理

本文先对数据进行调整，将同一方向的车道数据合并，比如 crossing1_west 的 1，2 车道合并。

表 1 合并后的数据部分展示

方向	时间	小车车流	大车车流	车流量
crossing_1_south	2023/1/11 0:00	6	0	6
crossing_2_east	2023/1/11 0:00	1	0	1
crossing_1_west	2023/1/11 0:05	10	0	10
crossing_3_north	2023/1/11 0:07	3	0	3
crossing_3_south	2023/1/11 0:07	3	0	3
crossing_3_east	2023/1/11 0:08	0	0	0

设定有效数据范围，对合并的后数据使用四分位法筛选异常数据。由于本文所用的数据量足够大，考虑交通流量随时间逐渐变化，少量异常值的替换对预测结果影响不大，因此本文将异常值替换为异常值附件两个正常值的平均值。借助 python 共筛除异常数据 900 条。

表 2 部分清理后的相关输入数据示例

方向	时间	小车车流	大车车流	车流量
crossing_1_east	2023-01-04 17:56:00	48	5	53
crossing_1_east	2023-01-04 18:11:00	38	3.5	41.5
crossing_1_east	2023-01-04 18:26:00	28	2	30
crossing_2_north	2023-02-01 10:14:00	43	2	45
crossing_2_north	2023-02-01 10:29:00	42.5	2	44.5
crossing_2_north	2023-02-01 10:44:00	42	2	44
crossing_3_south	2023-01-04 08:08:00	73	1	74
crossing_3_south	2023-01-04 08:23:00	72.5	1	73.5
crossing_3_south	2023-01-04 08:38:00	72	1	73

三、预备知识

(一) ARIMA 模型

ARMA 是一种随机时间序列模型，能够识别自然界中时间序列结构，并以最小化协方差矩阵的方式寻找最优的预测值^[4]。

设时间序列中 X_t 是一个依赖相邻数据和随机项的函数，相关表达式为公式(1)：

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{j=0}^q \theta_j a_{t-j} \quad (1)$$

其中 p 、 q 分别为 AR、MA 的阶数； φ_i 和 θ_j 是模型的自回归参数和移动平均系数； a_{t-j} 为随机差项。

在时间序列分析中，ARIMA 是 ARMA 模型的延申，可用来预测非平稳时间序列，ARIMA 模型通常定义为 ARIMA(p,d,q)，如公式(2)-(4)：

$$\varphi(B)\nabla^d X_t = \theta(B)a_t \quad (2)$$

$$\begin{cases} \varphi(B) = 1 - \sum_{i=1}^p \varphi_i B^i \\ \theta(B) = 1 - \sum_{j=1}^q \theta_j B^j \end{cases} \quad (3)$$

$$(1 - B)^d X_t = X_t - X_{t-d} = \nabla^d X_t \quad (4)$$

其中 d 为平稳化过程中的差分阶数， B 为滞后算子。

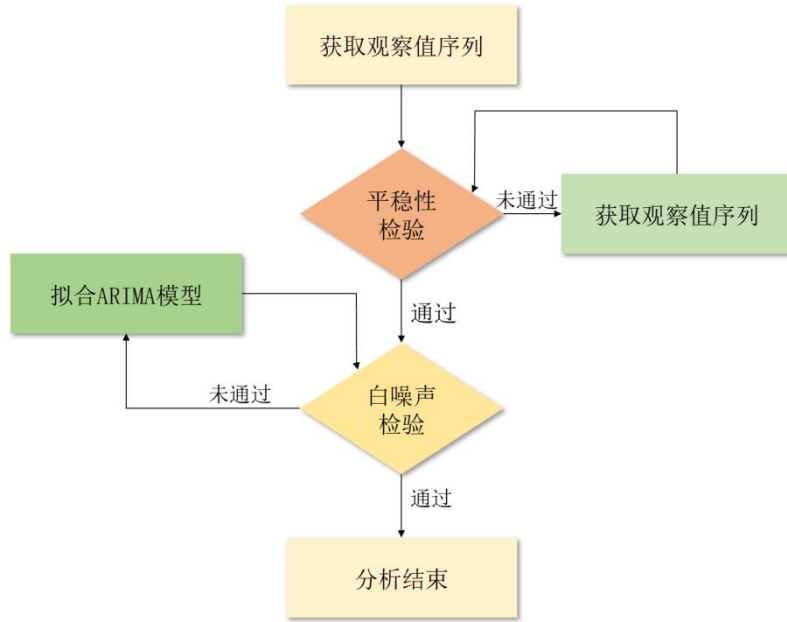


图 3 ARIMA 原理图

(二) RNN 模型

RNN 由输入层、输出层、隐藏层三部分构成，参数 U 、 V 、 W 在该循环结构的每一次输入中都会被共享。其主要工作由隐藏层实行，用于对序列的信息捕获，工作原理^[5]如公式(5)-(6)：

$$s_t = f_s(W_{s_{t-1}} + W_{in}x_t) \quad (5)$$

$$o_t = f_v(W_{out}s_t) \quad (6)$$

其中矩阵 $W_{s_{t-1}}$ 、 W_{in} 、 W_{out} 主要起连接作用，分别用在隐藏层节点的自连接，输入层到隐藏层和权值连接以及输出层到隐藏层的权值连接； x_t 、 o_t 分别表示为第 t 步的输入和输出，第 $t-1$ 步和第 t 步隐藏层的输出则分别表示为 $s_{(t-1)}$ 、 s_t 、 f_s 、 f_o 都为激活函数，分别属于隐藏层和输出层。

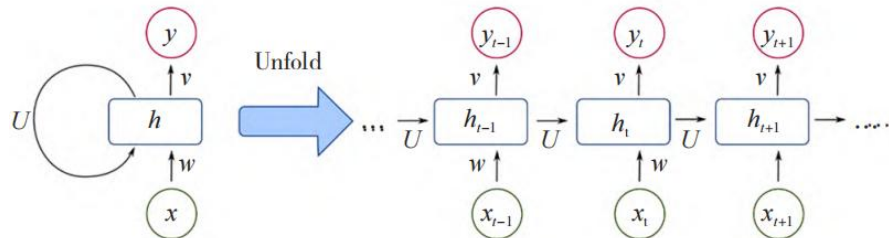


图 4 RNN 逻辑结构

将时间序列数据 $(x_1, x_2, \dots, x_6)$ 输入到 RNN 中，可以得到输出序列 $(y_1, y_2, \dots, y_6)$ ，其中 y_6 的值可以看成 x_7 时刻的预测值。再下一时刻可以输入 $(x_2, x_3, \dots, x_7)$ 时间序列数据，同理输出 y_7 作为 x_8 时刻的预测值，依次类推^[6]。

(三) LSTM 模型

相较于传统的 RNN，LSTM 的隐藏层单元被记忆块(细胞结构)代替，这些记忆块中的记忆单元能够存储前面的序列信息，同时增加 3 种门：输入门 i 、输出门 o 、遗忘门 f 用于控制记忆的序列信息，分别实现信息的流入、流出及对以前状态的更新，使 LSTM 能够解决 RNN 存在的梯度消失或梯度爆炸问题。各个门计算公式^[7]如公式(7)-(12)：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (7)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (8)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t * \tanh(C_t) \quad (10)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (11)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{C}_t \quad (12)$$

其中“ $\circ$ ”表示两个向量对应位置上的元素相乘；“ $\sigma(x) = 1/(1 + \exp(-x))$ ”表示 sigmoid 函数，为非线性激活函数； $\tilde{C}$ 为当前输入的单元状态； h 为细胞输出。

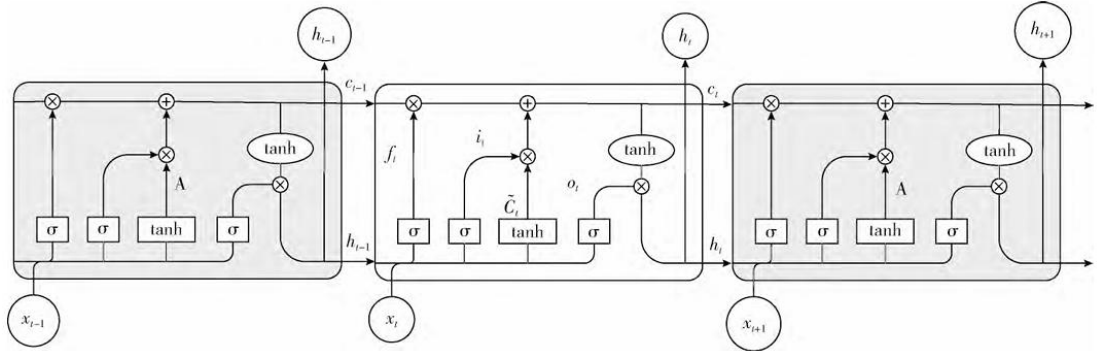


图 5 LSTM 细胞结构

(四) Attention-LSTM 模型

注意力机制通过计算注意力的概率分布，来突出某个关键输入对输出的影响。在 LSTM 中引入注意力机制，在不同时刻的分布不同的注意力权重，以此提高对输出更

为重要不同状态时刻数据的影响，从而提升模型的预测效果^[8]。

注意力机制模块的输入 $X = \{x^1, ..., x^k, ..., x^n\}$ ，表示 n 个输入时间序列： $x^k = \{x_1^k, ..., x_t^k, ..., x_T^k\}$ ，表示第 k 个输入序列，时间窗口为 T ； $h_1, h_2, ..., h_t$ 表示对应输入数量的隐藏层的状态值。通过加权平均来计算时刻 t 、特征 k 的概率。第一部分使用评分函数计算每个输入特征的分值，采用加性模型对输入时刻 t 、特征 k 用 $\tanh$ 函数进行评分，具体见公式(13)：

$$u_t^k = V_u \tanh(W_u * x_t^k + b_u) \quad (13)$$

第二部分通过对特征得分进行加权平均得到特征的概率向量，即注意力分布 a_t^k ，具体见公式(14)：

$$a_t^k = \frac{\exp(u_t^k)}{\sum_k \exp(u_t^k)} \quad (14)$$

u_t^k 表示第 k 个特征在 t 时刻的得分，即最值输出的最后节点， a_t^k 表示 u_t^k 经 Softmax 函数归一化后的概率向量； V_u, W_u, b_u 是可以被自行学习的虚拟网络物理参数。

$$a_t^k = \frac{\exp(u_t^k)}{\sum_k \exp(u_t^k)} \quad (15)$$

将注意力分布 a_t^k 与输入特征 x_t^k 相乘，便得到 t 时刻输入特征经 a_t^k 加权后的向量 $\tilde{x}_t = (a_t^1 x_t^1, a_t^2 x_t^2, ..., a_t^k x_t^k, ..., a_t^n x_t^n)$ 。并将 $\tilde{x}_t$ 通过 LSTM 模型进行预测。

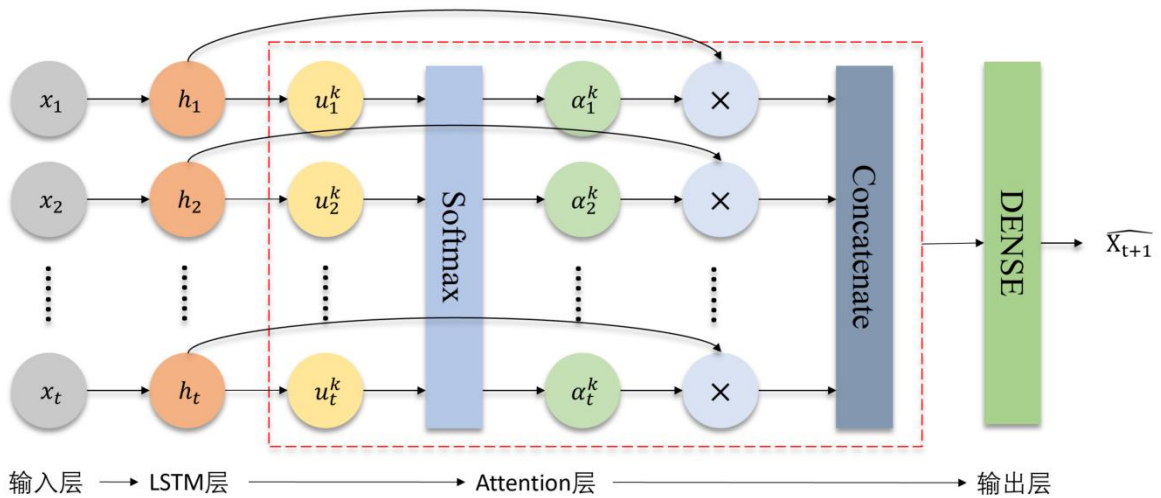


图 6 Attention-LSTM 模型结果图

其中，考虑车流量是非负的，本文在输出层（Dense 层）采用 ReLU 激活函数，即公式(16)-(17)：

$$Z = W_0[h_i; C] + b_0 \quad (16)$$

$$\widehat{x_{t+1}} = ReLU(Z) + W_0 \in R^{1 \times 2k}, b_0 \in R \quad (17)$$

其中 k 为 LSTM 隐藏层的维度，其中上下文向量 $C = \sum_k^T a_t^k h_t$ 。

（五）自相关误差

在时间预测系统^[9]中，输入矩阵 $X = [X_1, X_2, \dots, X_t, \dots, X_T] \in R^{T \times N}$ 表示在 T 个时间步长上均匀采样的 N 个变量。时序预测的目标是通过历史数据 $[X_{t-w}, X_{t-w+1}, \dots, X_{t-1}]$ 来预测 X_t ，即通过设计预测模型 f 实现：

$$\widehat{X_t} = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) + e_t \quad (18)$$

其中 k 为模型分为 k 个时间步数预测， θ 为模型参数。通常情况下，时序预测模型假设模型的预测误差 e_t 是独立同分布的，即 e_t 满足公式(19)-(20)：

$$cov(e_{t-\Delta t}, e_t) = 0, \forall \Delta t \neq 0 \quad (19)$$

$$e_t \sim N(0, \sigma^2) \quad (20)$$

但实际上时序数据会有自相关误差，一个具有 p 阶自相关误差的系统在形式上表现为模型的预测误差满足公式(21)：

$$e_t = \rho_1 e_{t-1} + \dots + \rho_p e_{t-p} + \epsilon_t, |\rho_i| < 1 \quad (21)$$

其中， ρ_i 为模型的 p 阶自相关系数， ϵ_t 是非自相关误差且满足(2)式。

四、基于 GWO-Attention-LSTM 的短时交通流量预测模型

为应对注意力机制模型迭代过程中权重值的不确定性，本文采用灰狼优化算法 (GWO) 随注意力机制的初始权重参数进行优化，将高维空间中灰狼的最佳捕猎位置作为初始权重参数的最优解，建立基于 GWO-Attention-LSTM 的短时交通预测模型。

(一) GWO 算法

GWO 算法的核心思想为模拟狼的社会关系和捕猎行为，通过不断调整参数来找到不同等级灰狼的最优捕猎位置。狼的社会关系分为 4 个等级(图 7)，依次为狼 a 、 b 、 c 、 u ，头狼 a 负责带领狼群寻找猎物、追逐和压制猎物，其余灰狼必须服从 a 的决定；狼 b 负责协助狼 a 并指挥除狼 a 以外的其余狼；狼 c 需要服从狼 a 、 b 的管理，并管理更低的灰狼；狼 u 为狼群中地位最低的灰狼，听从所有上级灰狼的指挥。

根据上述社会等级制度，狼群在最高级别的灰狼指挥下进行捕猎。利用 GWO 算法模拟的捕猎过程如下：

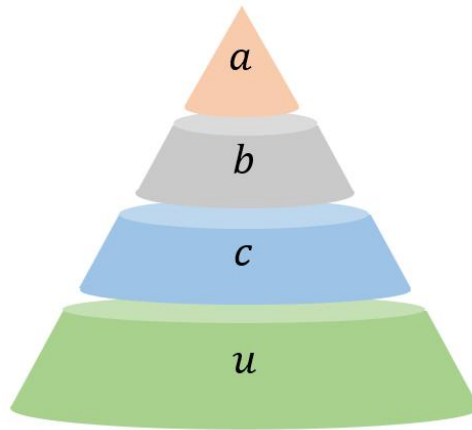


图 7 灰狼的社会等级

假设狼的数量为 l ，灰狼与其猎物之间的距离为 H 。狼寻找猎物并不断接近猎物的过程见方程(22)：

$$\begin{cases} H = |C \cdot X_R(k) - X_i(k)| \\ X_i(k+1) = X_R(k) - A \cdot H \end{cases} \quad (22)$$

其中 k 为迭代次数； $X_i(k)$ 为第 k 迭代后灰狼 i 的位置； $X_R(k)$ 为第 k 迭代后猎物的位置； $X_i(k+1)$ 表示灰狼再生后的位置； A 、 C 为系数向量，计算见方程(23)

$$\begin{cases} A = 2\alpha(r_1 - 1) \\ C = 2r_2 \end{cases} \quad (23)$$

其中, r_1 、 r_2 为0到1之间的随机数; α 为收敛因子, 从2逐渐减少至0.

在 d 维空间中, 设狼 a 的位置为 $X_a(X_{a,1}, X_{a,2}, \dots, X_{a,d})$, 狼 b 的位置为 $X_b(X_{b,1}, X_{b,2}, \dots, X_{b,d})$, 狼 c 的位置为 $X_c(X_{c,1}, X_{c,2}, \dots, X_{c,d})$ 。在狼 a 、 b 、 c 的带领下, 其他灰狼能够找到并包围猎物。其他灰狼的位置更新过程见方程(24)-(26)

$$\begin{cases} H_a = |C_1 \cdot X_a(k) - X_i(k)| \\ H_b = |C_2 \cdot X_b(k) - X_i(k)| \\ H_c = |C_3 \cdot X_c(k) - X_i(k)| \end{cases} \quad (24)$$

$$\begin{cases} X_{i,a}(k+1) = X_a(k) - A_1 H_a \\ X_{i,b}(k+1) = X_b(k) - A_2 H_b \\ X_{i,c}(k+1) = X_c(k) - A_3 H_c \end{cases} \quad (25)$$

$$X_i(k+1) = \frac{X_{i,a} + X_{i,b} + X_{i,c}}{3} \quad (26)$$

其中 H_a 、 H_b 、 H_c 分布为当前狼 $X_i(k)$ 与狼 a 、 b 、 c 之间的距离; $X_{i,a}(k+1)$ 、 $X_{i,b}(k+1)$ 、 $X_{i,c}(k+1)$ 为狼 i 依次在狼 a 、 b 、 c 带领下更新后的位置; $X_i(k+1)$ 为狼 i 在狼 a 、 b 、 c 共同带领下更新后的位置。

(二) GWO-Attention-LSTM 模型

生成灰色种群, 初始化每只灰狼的位置向量 $\beta = \{x_1, x_2, \dots, x_{N_w}\}$, N_w 为群体规模, $X_i = [q_1, q_2, \dots, q_\alpha] = [q_1, q_2, \dots, q_t] \in [-1, 1]^T$, α 为搜索维度, t 为时间步长。其中 α 、 t 与 GWO 中的 d 维空间维度相同。

计算每只灰狼的适应度, 保留适应前三的狼 a 、 b 、 c 。将 X_i 代入初始注意力权重 q , 通过 $\text{Softmax}([q_1, q_2, \dots, q_t] \cdot h_t)$ 计算 α_i , 将得到的权重与隐藏层相乘后得出系数向量 C , 再输入 Dense 层, 得出预测结果。将 X_i 代入 Attention-LSTM 模型计算适应度 $RMSE$, 并按大小依次排序得出最高的狼 a 、 b 、 c :

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (27)$$

其中, m 为样本数量; y_i 、 $\hat{y}_i$ 分别表示时刻 i 的实际交通流量值和交通流量预测值。

接着利用方程(24)-(26)和方程(23)分别更新每只狼的位置和狼 a 、系数向量 A 、 C

的值。

重复以上步骤，直至迭代次数 $> t_{max}$ ，输出最优灰狼位置 $X_{a,t} = [X_{a,1}, X_{a,2}, \dots, X_{a,t}]$ 与适应度 $RMSE$ 。

将最优的 $X_{a,t}$ 输入 Attention-LATM 中，重复公式(18)-(21)的步骤，得到最终预测结果。具体计算流程见图 8

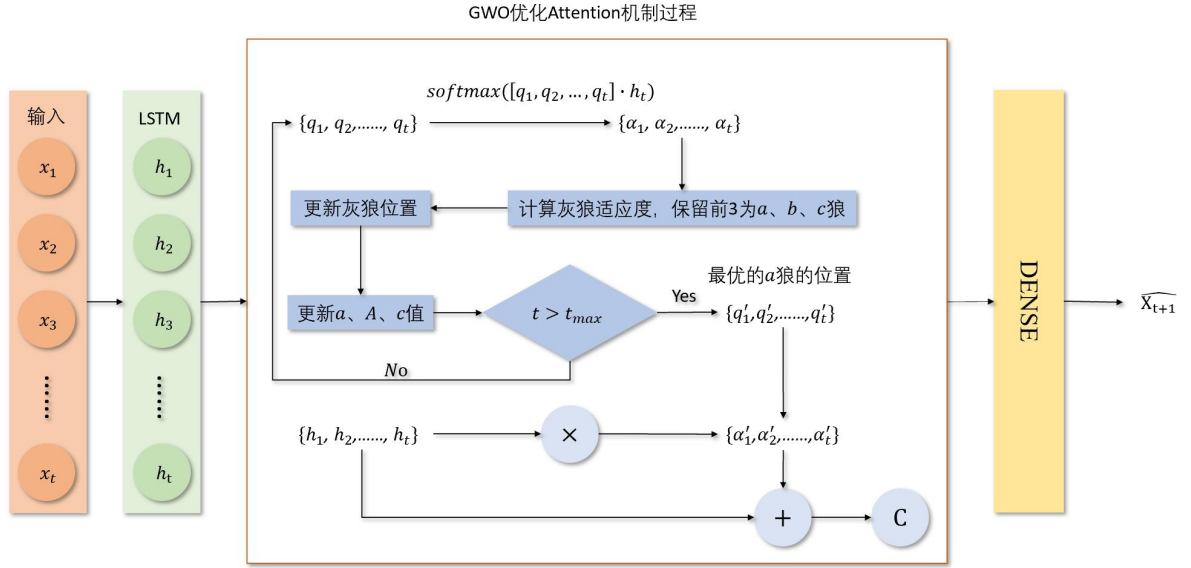


图 8 GWO-Attention-LSTM 流程图

五、基于 EAC-Attention-LSTM 和 GWO-Attention-LSTM 的短时交通流量预测模型

在通过最大似然估计法对时间序列训练神经网络时，通常假设各个时间步长上的误差是不相关的，然而由于数据的时间性质，很多情况下误差实际上是自相关的。为此，本文采用自相关补偿法，依次结合 Attention-LSTM 和 GWO-Attention-LSTM 模型，校正自相关误差影响，探究模型预测短时交流流量的性能。

（一）自相关误差补偿方法

本文考虑一阶自相关情况下的自相关补偿方法，令 $e_t = \rho e_{t-1} + \epsilon_t$ ，则一阶自相关误差下一个时序预测模型 $\widehat{X}_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) + \rho e_{t-1} + \epsilon_t$ 。其中 ϵ_t 可用于模型参数 θ 的训练， ρ 的确定成为提升模型预测精度关键。为此，通过对序列中相邻两项的线性操作来获得自相关系数 ρ 与序列预测值的相关关系。

$$\begin{cases} X_t = \widehat{X}_t + e_t & \text{①} \\ X_{t-1} = \widehat{X}_{t-1} + e_{t-1} & \text{②} \end{cases} \quad (28)$$

① - ρ ② 可得：

$$X_t - \rho X_{t-1} = \widehat{X}_t + \rho e_{t-1} + \epsilon_t - \widehat{X}_{t-1} + \rho e_{t-1} \quad (29)$$

$$= f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) - \rho f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) + \epsilon_t$$

其中相邻两项中的预测误差在一阶迟滞后通过自相关系数的引入相互抵消，即 $e_t - e_{t-1} \approx 0$ 。

模型参数通过最小化公式(29)的 MSE，得到一次优化后的模型参数 θ' 。一般真实的 ρ 值未知，因此用估计值 $\hat{\rho}$ 代替：

$$\begin{aligned} MSE &= \sum_t ||X_t - \hat{\rho} X_{t-1}||_2^2 \\ &= \sum_t f(X_{t-1}, X_{t-2}, \dots, X_{t-1}; \theta) - \hat{\rho} f(X_{t-2}, \dots, X_{t-k-1}; \theta) + \epsilon_t \end{aligned} \quad (30)$$

利用公式(30)中 θ' 计算模型的实际预测误差 e_t ：

$$e_t = X_t - \widehat{X}_t = X_t - f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) \quad (31)$$

将误差 e_t 对 e_{t-1} 进行线性回归来更新模型的自相关误差补偿系数 $\hat{\rho}$

$$\hat{\rho} = \frac{\sum_{t=2}^T e_t e_{t-1}}{\sum_{t=1}^{T-1} e_t^2} \quad (32)$$

至此模型实现自相关系数 ρ 和模型参数 θ 的一轮更新。随着模型不断迭代， θ 完成收敛。

通过对模型函数的更新，本文算法实现了 ρ 和 θ 的联合更新。但由于存在目标 $X_t - \hat{\rho}X_{t-1}$ 并非直接与模型输出相关，而是与两个模型的差值相关，这使模型优化变得复杂。为此，本文需要使用同一个输入集得单个模型来近似方程：

$$\begin{aligned} f(X_{t-1}, X_{t-2}, \dots, X_{t-k}; \theta) - \hat{\rho}f(X_{t-2}, X_{t-3}, \dots, X_{t-k-1}; \theta) \\ \approx f(X_{t-1}, X_{t-2}, \dots, X_{t-k-1}; \hat{\rho}, \theta) \end{aligned} \quad (33)$$

即 $\hat{\rho}$ 的确定方法和 MSE 依次被优化为公式(34)：

$$\begin{aligned} X_t - \hat{\rho}X_{t-1} &= f(X_{t-1}, X_{t-2}, \dots, X_{t-k-1}; \hat{\rho}, \theta) \\ MES &= \sum_t f(X_{t-1}, X_{t-2}, \dots, X_{t-k-1}; \hat{\rho}, \theta) \end{aligned} \quad (34)$$

虽然模型的目标已与输出相关，且两个模型的输出已各为一，但模型的输入序列 X_t 与目标序列 $X_t - \hat{\rho}X_{t-1}$ 并不相同，而通常邀请这两者序列形式相同。因此进一步优化 $\hat{\rho}$ 的确定方法为 X 式，使两序列都采取 $X_t - \hat{\rho}X_{t-1}$ 的形式。

$$\begin{aligned} X_t - \hat{\rho}X_{t-1} &= f(X_{t-1} - \hat{\rho}X_{t-2}, \dots, X_{t-k} - \hat{\rho}X_{t-k-1}; \theta) \\ &= f(X_{t-1}, X_{t-2}, \dots, X_{t-k-1}; \theta, \hat{\rho}) \end{aligned} \quad (35)$$

(二) EAC-Attention-LSTM 模型

将 EAC 融入 Attention-LSTM 模型中，具体流程如下：

1. 给定时间步长 W ，假设一个步长内数据为 $\{X_{t-w}, X_{t-w+1}, \dots, X_{t-1}\}$;
2. 定义差分变量 $\widetilde{X}_{t-k} = X_{t-k} - \rho X_{t-k-1}, k = 1, 2, \dots, W$ 。训练时每个样本输入为序列 $[\widetilde{X}_{t-w}, \widetilde{X}_{t-w+1}, \dots, \widetilde{X}_{t-1}]$ ，目标标签为 $\widetilde{X}_t = X_t - \rho X_{t-1}$ ，使用自相关误差补偿方位确定 LSTM 模型参数 θ_{LSTM} ;
3. 对每个时刻 i ，考虑上一时刻的隐藏层数据 h_{i-1} 和记忆状态 c_{i-} ，利用公式(36)得出当前时刻的 h_i 和 c_i ：

$$h_i, c_i = LSTM(\widetilde{X}_i, h_{i-1}, c_{i-1}; \theta_{LSTM}) \quad (36)$$

4. 为计算 h_i 注意力权重，利用公式(37)-(38)得到上下文向量 C

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^W \exp(e_j)} \quad (37)$$

$$C_t = \sum_{i=1}^w \alpha_i h_i \quad (38)$$

5. 将上下文向量 C 输入输出层（Dense 层），运用 $\hat{X}_t = \text{ReLU}(W_0[C_t; h_{t-1}] + b_0)$ 计算得到最终的预测结果。

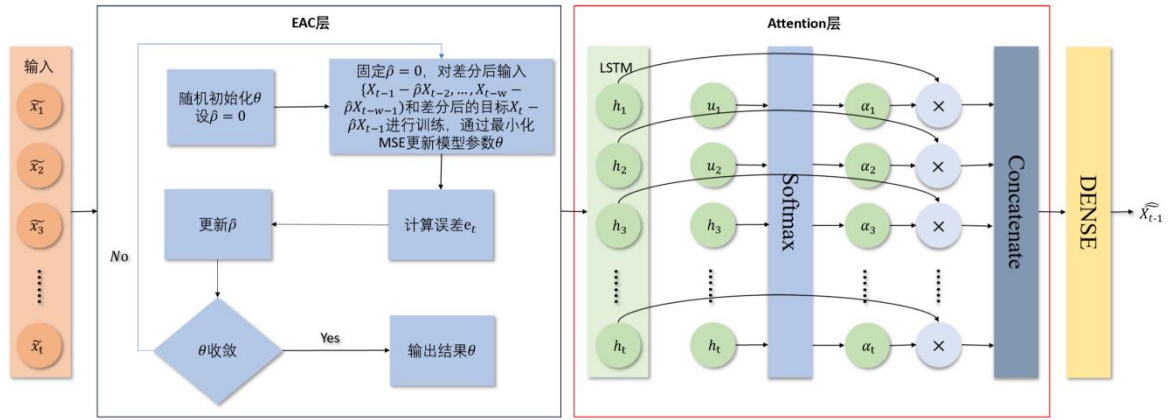


图 9 EAC-Attention-LSTM 流程图

（三）EAC-GWO-Attention-LSTM 模型

本模型是将 EAC 融入 GWO-Attention-LSTM 模型，先使用 GWO 对 Attention 机制进行优化，并且输出 α_i ，再按照前文步骤，将上下文向量输入 Dense 层即可。具体流程图如下：

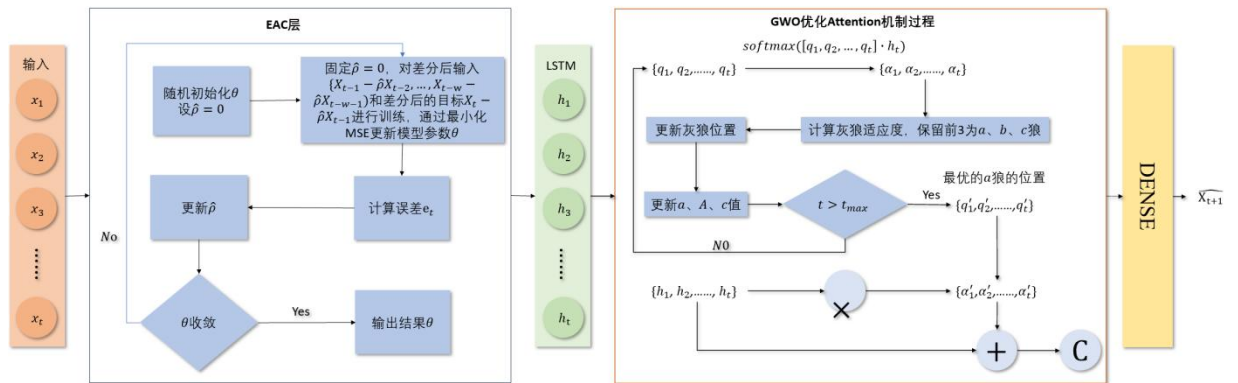


图 10 EAC-GWO-Attention-LSTM 流程图

六、模型模拟与分析

将处理好的嘉兴市交通流量数据，借助 python 软件，依次用上述模型进行求解，并计算各预测模型的性能，对比得出最优的短时交通流量预测模型。

（一）预测模型评估指标

由于真实值 y_i 中存在较多为 0 的数据，用 MAPE 直接会使其值趋于无穷大，而用 SMAPE 则会导致分母不稳定、指标波动剧烈，并且 SMAPE 值往往会被拉高，所以使用 SMAPE 对本研究的数据而言会过于敏感，无法真实反映模型在大部分时段的预测精度。因此，本文选取MSE、RMSE作为本次预测模型的评估指标，其中误差越大，指标值均越大。

1. 平均绝对误差MAE

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (39)$$

2. 均方根误差RMSE

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|^2} \quad (40)$$

(39) - (41) 式中 m 表示样本数量； $y_i = \{y_1, y_2, \dots, y_n\}$ 表示时刻 i 是实际交通流量值； $\hat{y}_i = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ 表示时刻 i 是交通流量的预测值。

（二）预测模型参数设置

基于 python 语言环境，对比模型设定为：ARIMA、RNN、LSTM、Attention-LSTM、GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 模型。

重要超参数的选择：

在实验中，所有神经网络的激活函数均设为 ReLU 函数，训练集：验证集：测试集比例均为 7：1.5：1.5，输入层和输出层的神经元数量分别为 12 和 1。模型的优化器采用自适应矩估计 ADAM。其中除有关 EAC 的损失函数采用(34)式优化后的均方误差MSE外，其他神经网络均采用MSE作为损失函数；除 GWO-Attention-LSTM 学习率为 0.01 外，其他模型的学习率均为 0.001。Attention 神经元数量等于 LSTM 的神

元数量。其余各部分融合模型的超参数设置如下：

- LSTM、RNN 的神经元数量为 64，batch_size 为 64，层数为 2，时间步长为 12，epochs 值分别为 50、75，早停次数分别为 5、10。
- Attention-LSTM 的神经单元数量为 128，LSTM 层数为 1，batch_size 为 128，epochs 为 75，早停次数为 10，时间步长为 12。
- GWO 中狼群数量 l 设为 30，权重参数值范围为 $[-1, 1]$ ，搜索空间维度 d 等于时间步长，其 epochs 为 5，早停次数设为 30。若连续 30 个序列的均方误差不再下降，则模型停止训练并保存最优模型。其中，GWO-Attention-LSTM 和 EAC-GWO-Attention-LSTM 的 GWO 最大迭代数分别设为 25 和 10，batch_size 分别为 64、32，时间步长分别为 12、8。
- EAC 的 epochs 为 300， θ 学习率为 0.003， $\hat{p}$ 学习率为 0.01，其中 EAC-Attention-LSTM 的 batch_size 为 64，时间步长为 12，LSTM 层数为 1、神经元数量为 64。
- ARIMA 模型各朝向数据参数设置见表 3。

表 3 ARIMA 各朝向数据的参数设置

ARIMA	(p, d, q)
crossing_1_east	(2, 1, 1)
crossing_1_north	(3, 0, 1)
crossing_1_south	(0, 0, 0)
crossing_1_west	(0, 1, 0)
crossing_2_east	(0, 0, 0)
crossing_2_north	(0, 0, 3)
crossing_2_south	(0, 0, 0)
crossing_2_west	(1, 0, 2)
crossing_3_east	(2, 1, 3)
crossing_3_north	(0, 0, 2)
crossing_3_south	(0, 0, 3)
crossing_3_west	(2, 0, 2)

注：各对比模型具体超参数设置见附录 1。

(三) 预测结果分析

本文借助 python 软件，分别对 ARIMA、RNN、LSTM、Attention-LSTM、GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 模型进行训练。所有模型的预测误差比较见图 11，具体数值见附录 2。

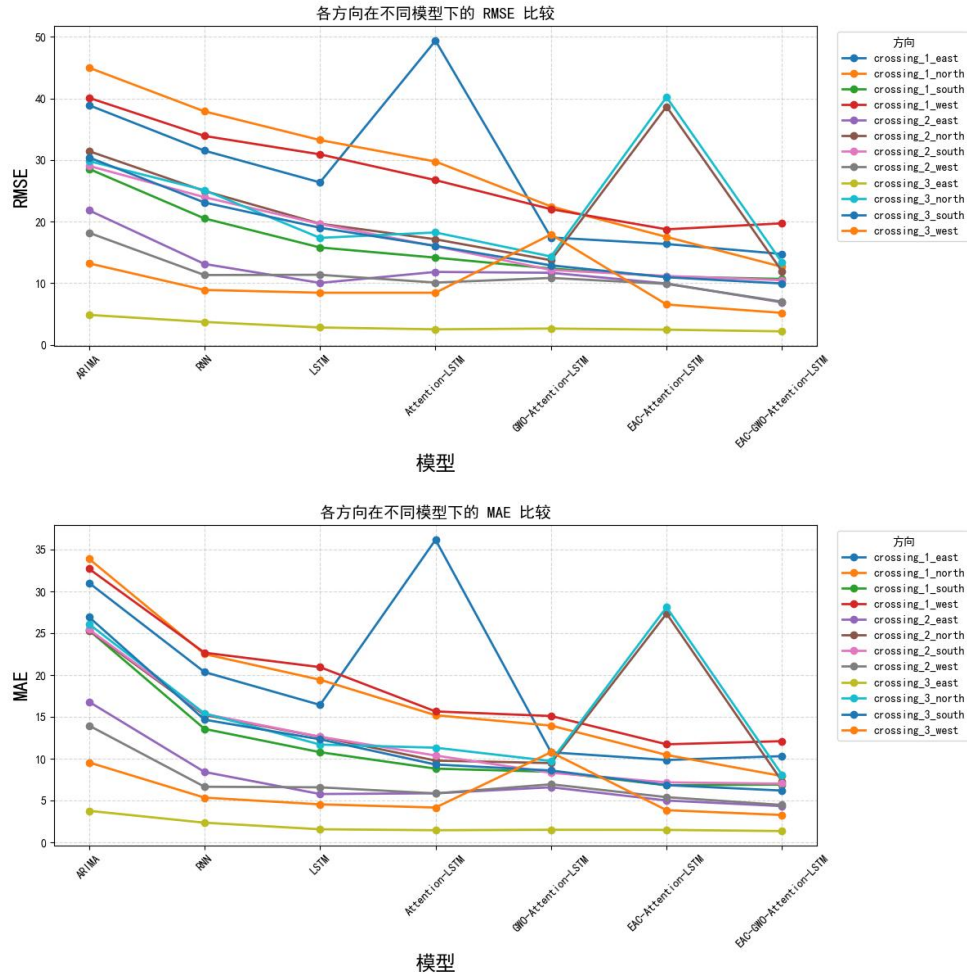


图 11 所有模型预测误差比较(MAE、RMSE)

由图 11 可得，EAC-GWO-Attention-LSTM 模型的预测结果优于其他模型，得到以下结果：

在 12 个朝向的路线中，与其他模型相比，EAC-GWO-Attention-LSTM 模型的MAE值(辆)分别平均降低了 15.95、6.93、4.67、4.56、2.58、3.73；RMSE值(辆)分别平均降低了 17.21、11.10、7.49、7.96 、3.81 、5.73。具体降低值见附录 3。

可见 EAC-GWO-Attention-LSTM 模型在精度上相较于传统预测模型有了显著的

提升，为嘉兴市短时交通流量预测的最优模型。

将 EAC-GWO-Attention-LSTM 模型的预测结果与实际值进行对比，由图 12-14 可得该模型能够准确地预测实际交通流量时间序列的趋势，模型的预测值折线图与实际值的折线图更为一致。

此外，预测点与几个实际点的峰值高度接近，即证明 EAC-GWO-Attention-LSTM 模型具有良好的非线性拟合能力。

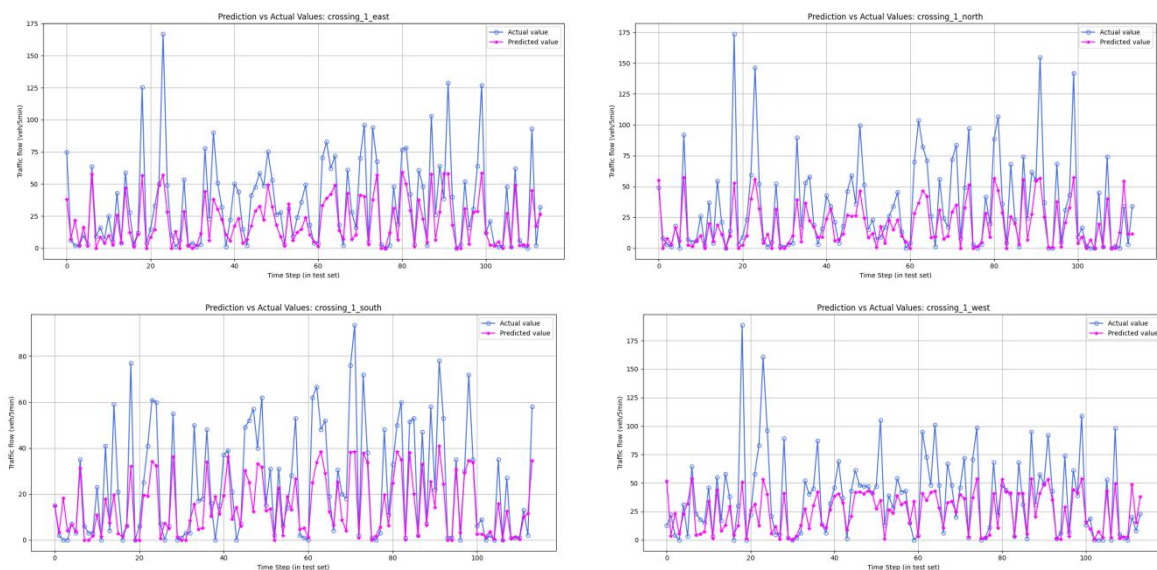


图 12 预测结果与实际值对比(依次为 crossing_1_east/north/south/west)

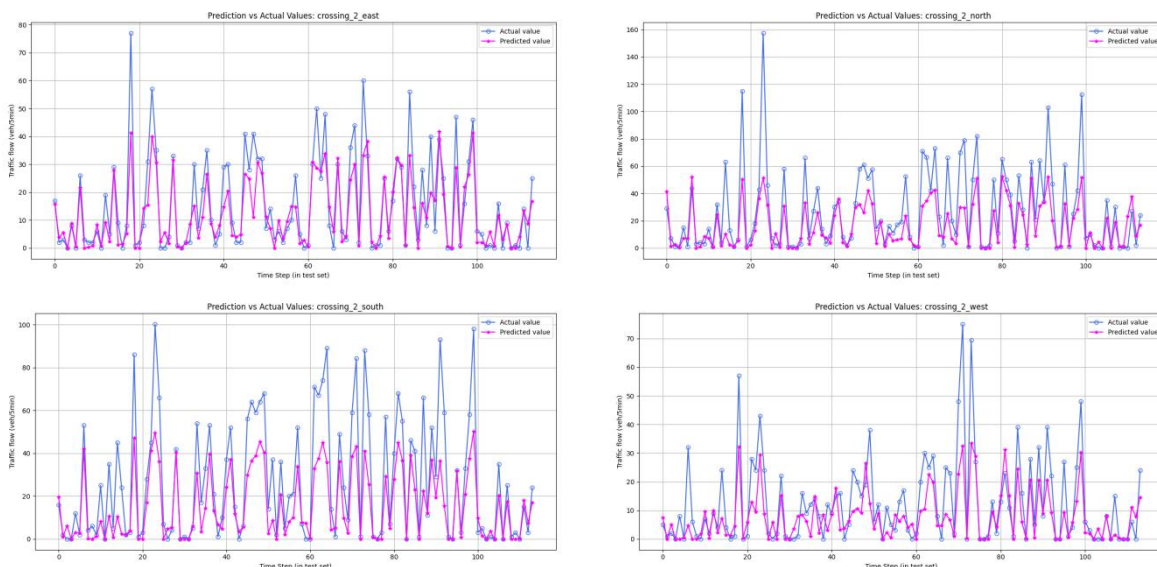


图 13 预测结果与实际值对比(依次为 crossing_2_east/north/south/west)

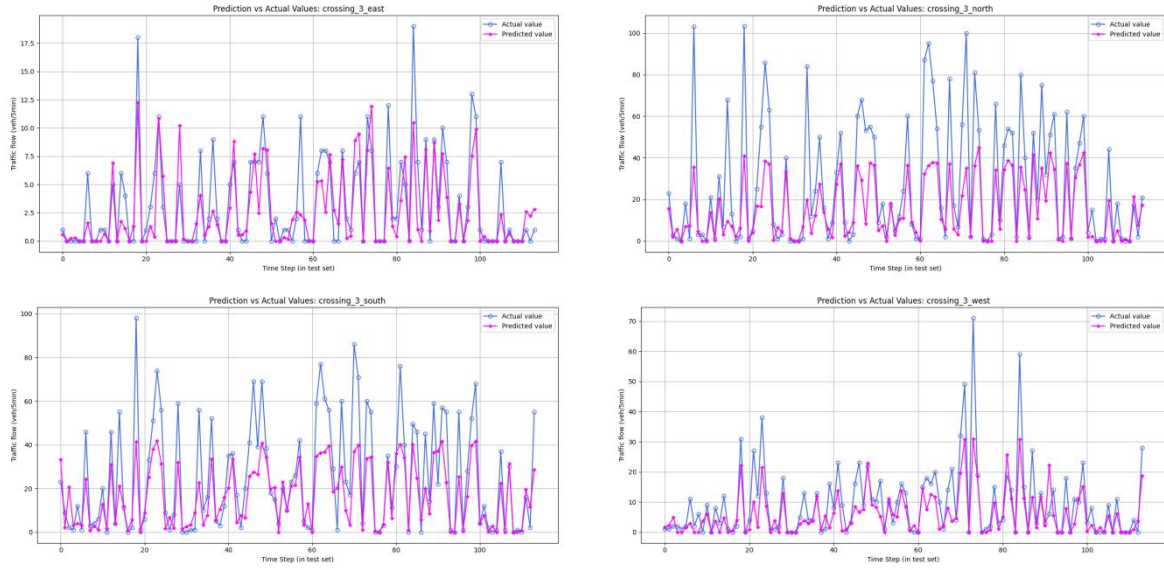


图 14 预测结果与实际值对比(依次为 crossing_3_east/north/south/west)

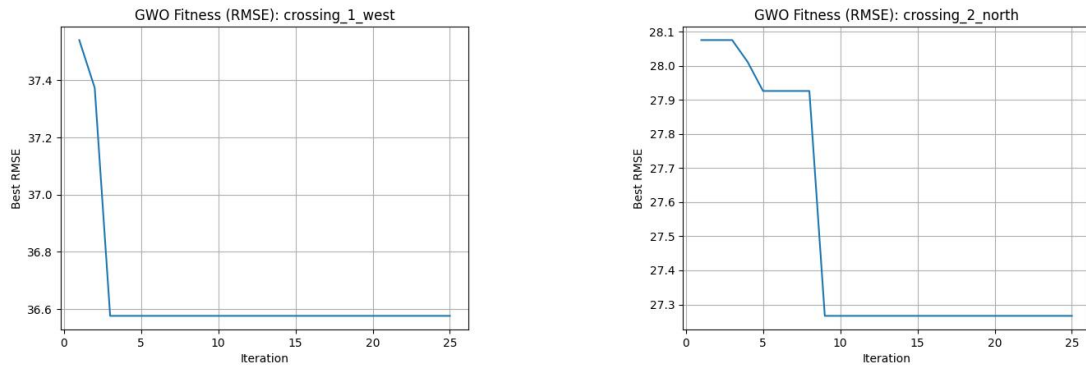
(四) 优化分析

由于正文的限制，本文在该部分只展示部分朝向的优化分析过程。具体各个朝向优化分析详见附录 4-6。

1. GWO 优化分析

GWO 算法通过迭代过程中不断更新狼的捕猎位置，从而逐步接近并包围目标猎物。狼 α 的迭代次数与适应度值(RMSE)变化过程详见附录 4。

其中经过 3 次迭代后，crossing_1_west 的 RMSE 稳定在 36.587；经过 9 次迭代后，crossing_2_north 的 RMSE 稳定在 27.245；经过 13 次迭代后，crossing_3_north 的 RMSE 稳定在 225.169。结果表明，灰狼优化算法具有收敛速度快、适用性强的优点。对于所有路段都能快速获得最佳适应的灰狼，从而获得注意力机制的最优舒适权重参数。



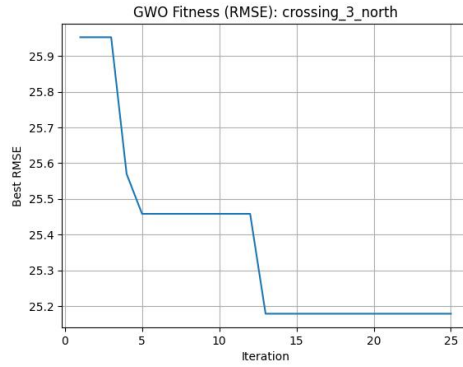


图 15 迭代次数与 RMSE 变化过程(依次为 crossing_1_west gwo_rmse、crossing_2_north_gwo_rmse、crossing_3_north_gwo_rmse)

2. EAC 优化分析

EAC 算法通过在训练过程中不断更新误差相关系数 ρ ，使模型逐步补偿残差中的一阶自相关结构。 ρ 随迭代轮次的变化过程图详见附录 5。表 4 给出 3 个路口各个朝向的 ρ 值随训练轮次的变化及其收敛特征：

表 4 各朝向 ρ 收敛迭代次数及稳定值

朝向	收敛迭代次数	ρ 收敛稳定值
crossing_1_east	75	0.08
crossing_1_north	250	0.35
crossing_1_south	120	0.11
crossing_1_west	110	0.49
crossing_2_east	105	0.15
crossing_2_north	80	0.10
crossing_2_south	100	0.12
crossing_2_west	80	0.42
crossing_3_east	100	0.03
crossing_3_north	100	0.07
crossing_3_south	120	0.15
crossing_3_west	98	0.12

由表 4 得，绝大多数朝向在百轮以内即可达到稳定，且对无自相关的序列(如各路

口西向)几乎瞬时收敛。EAC 能够有效针对不同路口、不同朝向的残差自相关强度，快速学习并得到最优的 ρ 值，其中高 ρ 值方向的残差自相关最强，EAC 修正潜力最大，低 ρ 值方向自相关弱，但也拥有一定 EAC 修正能力。

因此，EAC 能够在训练中高效识别并补偿误差自相关，为后续 Attention 机制分配权重提供基础，显著提升模型的稳定性和精度。

3. Epochs 和损失函数曲线对比

本文采取批量(batch)训练，一个 epoch 通常包含多次迭代，每次迭代处理一个 batch。在每个 epoch 结束后，在验证集上计算一次损失，从而得到“epoch-loss”曲线。在训练过程中设置了早停机制，这种方法能够有效避免过拟合。

为对比检验模型配置的优劣，本文以轮次多少为判断指标。如果两个设置最终损失相近，达到同样低损失所需 epoch 最少的更优。即模型越早跳出训练，其性能就越好。为此，本文对 RNN、LSTM、Attention-LSTM，GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM 模型的 Epochs 和损失函数曲线进行比较。由于正文的限制，本文以 crossing_3 各朝向的 Epochs 和损失函数曲线关系图为例进行展示，具体各模型变化图见附录 6。

对比分析图 16-21，可得 EAC-GWO-Attention-LSTM 模型迭代次数最少，损失值 MSE 的下降速度最快，达到同样低损失所需 epoch 最少。因此该模型相较于其他模型具有更好的模型性能。

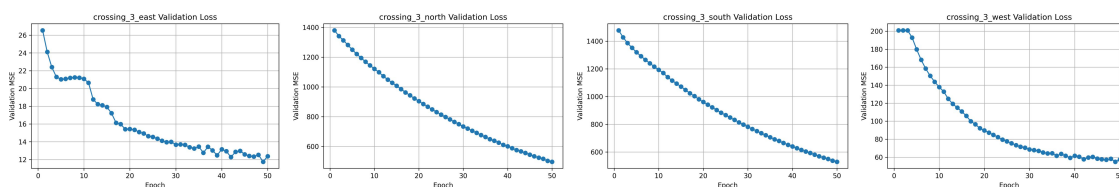


图 16 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

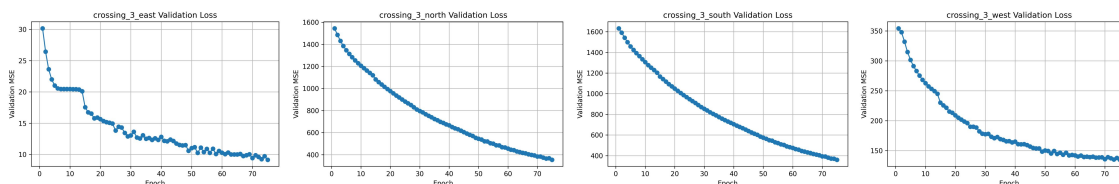


图 17 LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

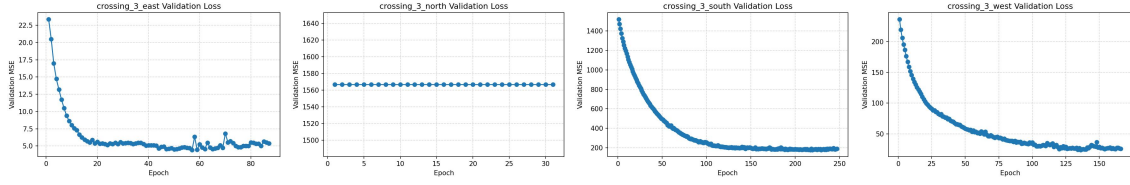


图 18 Attention-LSTM 的 MSE 随 Epochs 变化曲线
(依次为 crossing_3_east/north/south/west)

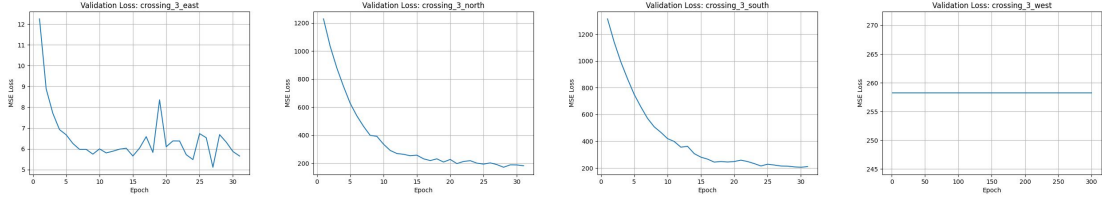


图 19 GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线
(依次为 crossing_3_east/north/south/west)

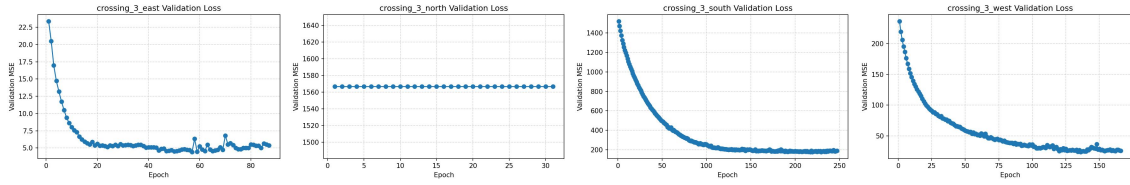


图 20 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_3_east/north/south/west)

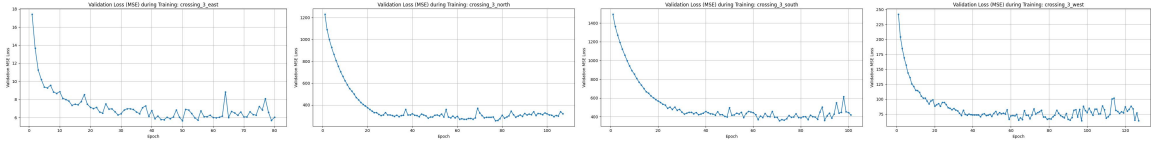


图 21 EAC-GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_1_east/north/south/west)

七、研究结论与反思

（一）研究结论

为使短时交通流量预测模型更准确且易于训练，建立 EAC-GWO-Attention-LSTM 模型。考虑时间序列特征信息的重要性和自相关误差的影响，在传统的 LSTM 模型中，利用自相关误差补偿法度自相关误差进行校正，结合 GWO 算法对注意力机制的初始权重参数进行优化。

本文以嘉兴市的交通量数据为研究对象，基于多模型对比分析视角，将处理好的数据，借助 python 软件，通过实验数据的模拟验证了 EAC-GWO-Attention-LSTM 的 MAE、RMSE 的预测误差值明显低于另外六个模型。同时，其能够准确地预测实际交通流量时间序列的趋势，模型的预测值折线图与实际值的折线图基本一致，预测点与几个实际点的峰值高度接近，即证明该模型具有良好的非线性拟合能力预测准确性。此外，通过模型优化分析证明，EAC-GWO-Attention-LSTM 更易于训练。

因此，将 EAC-GWO-Attention-LSTM 应用于当地的短时交通预测，能够在一定程度上提高该城市交通管理的智能化水平，为出行者提供实施可靠的交通信息，从而有效缓解交通拥堵、优化出行路径以及减少汽车尾气的排放。

（二）研究反思

本研究以嘉兴市实际交通流量数据为基础进行研究，引入 GWO 优化注意力机制初始权重参数和考虑自相关误差补偿方法的 Attention-LSTM 模型，能有效提高模型预测精度，增强模型对交通流量数据内在结构的捕捉能力，具有一定的创新意义。

同时，本研究从传统模型到前沿模型的完整预测体系，能够全面评估不同类型模型的适用性及优劣，对比得出的 EAC-GWO-Attention-LSTM 模型具有高实际价值，能够直接为该城市交通管理决策提供有力的数据支持和理论依据。

但是随着模型复杂性逐步提高，尤其是融合了 GWO 和自相关误差补偿方法的模型，参数量和计算复杂度显著增加，可能影响模型实际部署的便捷性和实时预测能力。

未来本研究将从以下方面进行改进：

1. 由于数据量不够大，并且考虑的相关因素较少，导致本研究的模型在长期交通流量预测方面效果欠佳。未来可以收集更广范围和更长时间周期的数据，进一步提

高模型的泛化能力和长期预测准确性。

2. 本研究假设误差项仅存在一阶自相关，但在真实的交通流量数据中，由于数据本身的复杂性和非线性特性，误差项可能不仅存在一阶自相关，还可能存在二阶甚至更高阶的自相关结构。后续研究中可以进一步考虑更高阶的误差自相关结构，进行对数据内部特征的充分挖掘，从而进一步提高预测精度。
3. 本研究由于计算机的算力限制，一些超参数的设置主要依赖经验设定，后续研究可进一步通过网格搜索或随机搜索更系统地确定最优参数，避免人为主观因素影响。

本研究的模型方法不仅局限于交通流量预测，还可推广应用于物流预测、城市能源需求预测等其他短期时间序列预测领域，具备较好的方法迁移性。

八、参考文献

- [1] 交通运输部关于印发《数字交通“十四五”发展规划》的通知-政府信息公开-交通运输部
- [2] 杨丽,吴雨茜,王俊丽,等.循环神经网络研究综述[J].计算机应用,2018,38(S2):1-6+26.
- [3] 李梅,宁德军,郭佳程.基于注意力机制的 CNN-LSTM 模型及其应用[J].计算机工程与应用,2019,55(13):20-27.
- [4] 王晓全,邵春福,尹超英,等.基于 ARIMA-GARCH-M 模型的短时交通流预测方法[J].北京交通大学学报,2018,42(04):79-84.
- [5] 李洁,林永峰.基于多时间尺度 RNN 的时序数据预测[J].计算机应用与软件,2018,35(07):33-37+62.
- [6] 徐皓,陈国凡,王双英.基于 RNN 模型的短时交通流量预测[J].今日制造与升级,2023,(09):50-52.
- [7] 晏臻,于重重,韩璐,等.基于 CNN+LSTM 的短时交通流量预测方法[J].计算机工程与设计,2019,40(09):2620-2624+2659.DOI:10.16208/j.issn1000-7024.2019.09.038.
- [8] 彭文,王金睿,尹山青.电力市场中基于 Attention-LSTM 的短期负荷预测模型[J].电网技术,2019,43(05):1745-1751.DOI:10.13335/j.1000-3673.pst.2018.1554.
- [9] 蒋云展.面向物联网时序数据预测的多视角稀疏化 Transformer 技术研究[D].西安电子科技大学,2023.DOI:10.27389/d.cnki.gxadu.2023.001101.
- [10] Zhang, H., Dai, G. The strategy of traffic congestion management based on case-based reasoning. Int J Syst Assur Eng Manag 10, 142 – 147 (2019).
- [11] Yao, B.; Ma, A.; Feng, R.; Shen, X.; Zhang, M.; Yao, Y. A deep learning framework about traffic flow forecasting for urban traffic emission monitoring system. Front. Public Health 2022, 9, 804298
- [12] Xiaolei Ma, Zhimin Tao, Yinhai Wang, Haiyang Yu, Yunpeng Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Transportation Research Part C: Emerging Technologies, Volume 54,

2015, Pages 187-197,

- [13] Cao, B.; Gao, M.T. Research on short-term traffic flow prediction based on LSTM. *Mod. Comput. (Prof. Ed.)* 2018, 25, 3 – 7.
- [14] Zhong, C.H.; Wang, M.; Zhao, W.; Zhang, Y. Short term traffic flow prediction of intersection based on LSTM. *Highw. Traffic Technol.* 2019, 35, 6.
- [15] Lan, T.; Zhang, X.; Qu, D.; Yang, Y.; Chen, Y. Short-Term Traffic Flow Prediction Based on the Optimization Study of Initial Weights of the Attention Mechanism. *Sustainability* 2023, 15, 1374.

附录 1 各模型具有参数设置表

表 5 RNN、LSTM、Attention-LSTM 模型参数设置

	RNN	LSTM	Attention-LSTM
时间步长	12	12	12
神经元数量	64	64	128
层数	2	2	1
batch_size	64	64	128
学习率	0.001	0.001	0.001
epochs	50	75	75
早停次数	5	10	10
训练集：验证集：测试集比例	7： 1.5： 1.5	7： 1.5： 1.5	7： 1.5： 1.5
优化器	Adam	Adam	Adam
激活函数	ReLU	ReLU	ReLU
输入层神经元数量	12	12	12
输出层神经元数量	1	1	1
损失函数	MSE	MSE	MSE
Attention 神经元数量	/	/	等于 LSTM 神经元数量
狼群数量	/	/	/
GWO 的最大迭代次数	/	/	/
搜索空间的维度	/	/	/
GWO 的 epochs	/	/	/
EAC 的 epochs	/	/	/
θ 学习率	/	/	/
$\hat{\rho}$ 学习率	/	/	/
EAC 的损失函数	/	/	/

表 6 GWO-Attention-LSTM、EAC-Attention-LSTM、EAC-GWO-Attention-LSTM
模型参数设置

	GWO-Attention- LSTM	EAC-Attention- LSTM	EAC-GWO-Atten- tion-LSTM
时间步长	12	12	8
LSTM 神经元数量	64	64	128
LSTM 层数	1	1	1
batch_size	64	64	32
学习率	0.01	0.001	0.001
LSTM 的 epochs	300	300	300
早停次数	30	30	30
训练集：验证集：测试集比例	7：1.5：1.5	7：1.5：1.5	7：1.5：1.5
优化器	Adam	Adam	Adam
LSTM 的激活函数	ReLU	ReLU	ReLU
输入层神经元数量	12	12	12
输出层神经元数量	1	1	1
损失函数	MSE	MSE	MSE
Attention 神经元数量	等于 LSTM 神经 元数量	等于 LSTM 神经 元数量	等于 LSTM 神经 元数量
狼群数量	30	/	20
GWO 的最大迭代次数	25	/	10
搜索空间的维度	等于时间步长	/	等于时间步长
GWO 的 epochs	5	/	5
EAC 的 epochs	/	300	300
θ 学习率	/	0.003	0.003
$\hat{\rho}$ 学习率	/	0.01	0.01
EAC 的损失函数	/	优化后的 MSE	优化后的 MSE

附录 2 各模型不同方向的 MAE、RMSE 误差值

表 7 各模型不同方向的 MAE 误差（辆）

方向	ARIMA	RNN	LSTM	Attention- LSTM	GWO-Attent ion-LSTM	EAC-Attent ion-LSTM	EAC-GWO-At tention-LSTM
crossing_1_east	31.02	20.37	16.45	36.17	10.79	9.87	10.33
crossing_1_north	33.89	22.52	19.46	15.22	13.98	10.47	7.95
crossing_1_south	25.36	13.58	10.80	8.84	8.46	6.85	6.94
crossing_1_west	32.70	22.68	20.96	15.67	15.12	11.75	12.12
crossing_2_east	16.82	8.44	5.81	5.89	6.61	5.03	4.37
crossing_2_north	25.32	15.24	12.63	9.80	9.51	27.35	7.32
crossing_2_south	25.46	15.38	12.65	10.41	8.35	7.20	7.04
crossing_2_west	13.98	6.67	6.61	5.88	6.97	5.43	4.50
crossing_3_east	3.80	2.38	1.60	1.49	1.54	1.52	1.38
crossing_3_north	26.11	15.43	11.71	11.34	9.75	28.11	8.10
crossing_3_south	26.94	14.69	12.34	9.33	8.62	6.87	6.22
crossing_3_west	9.58	5.38	4.58	4.19	10.80	3.88	3.30

表 8 各模型不同方向的 RMSE 误差（辆）

方向	ARIMA	RNN	LSTM	Attention- LSTM	GWO-Attent ion-LSTM	EAC-Attent ion-LSTM	EAC-GWO-At tention-LSTM
crossing_1_east	38.90	31.52	26.37	49.37	17.43	16.37	14.75
crossing_1_north	44.98	37.89	33.24	29.75	22.46	17.51	12.66
crossing_1_south	28.56	20.52	15.81	14.16	12.39	11.08	10.70
crossing_1_west	40.09	33.91	30.92	26.74	22.02	18.75	19.72
crossing_2_east	21.84	13.13	10.05	11.84	11.70	9.96	6.89
crossing_2_north	31.44	24.98	19.67	17.13	13.73	38.67	11.94
crossing_2_south	29.04	23.97	19.63	16.02	12.15	11.18	10.50
crossing_2_west	18.20	11.33	11.37	10.11	10.86	9.90	7.01
crossing_3_east	4.86	3.72	2.82	2.52	2.65	2.47	2.19
crossing_3_north	29.83	25.11	17.36	18.24	14.34	40.20	13.35
crossing_3_south	30.42	23.11	19.01	16.06	12.90	10.97	9.96
crossing_3_west	13.23	8.91	8.46	8.46	17.93	6.55	5.19

附录 3 EAC-GWO-Attention-LSTM 与其他各模型预测误差的对比

表 9 EAC-GWO-Attention-LSTM 与其他模型不同方向的 MAE 误差（辆）减少值

方向	ARIMA	RNN	LSTM	Attention- LSTM	GWO-Atten tion-LSTM	EAC-Atten tion-LSTM
crossing_1_east	20.69	10.04	6.12	25.84	0.46	-0.46
crossing_1_north	25.94	14.57	11.51	7.27	6.03	2.52
crossing_1_south	18.42	6.64	3.86	1.9	1.52	-0.09
crossing_1_west	20.58	10.56	8.84	3.55	3	-0.37
crossing_2_east	12.45	4.07	1.44	1.52	2.24	0.66
crossing_2_north	18	7.92	5.31	2.48	2.19	20.03
crossing_2_south	18.42	8.34	5.61	3.37	1.31	0.16
crossing_2_west	9.48	2.17	2.11	1.38	2.47	0.93
crossing_3_east	2.42	1	0.22	0.11	0.16	0.14
crossing_3_north	18.01	7.33	3.61	3.24	1.65	20.01
crossing_3_south	20.72	8.47	6.12	3.11	2.4	0.65
crossing_3_west	6.28	2.08	1.28	0.89	7.5	0.58
平均差值	15.95	6.93	4.67	4.56	2.58	3.73

表 10 EAC-GWO-Attention-LSTM 与其他模型不同方向 RMSE 误差（辆）减少值

方向	ARIMA	RNN	LSTM	Attention- LSTM	GWO-Atten tion-LSTM	EAC-Atten tion-LSTM
crossing_1_east	24.15	16.77	11.62	34.62	2.68	1.62
crossing_1_north	32.32	25.23	20.58	17.09	9.8	4.85
crossing_1_south	17.86	9.82	5.11	3.46	1.69	0.38
crossing_1_west	20.37	14.19	11.2	7.02	2.3	-0.97
crossing_2_east	14.95	6.24	3.16	4.95	4.81	3.07
crossing_2_north	19.5	13.04	7.73	5.19	1.79	26.73
crossing_2_south	18.54	13.47	9.13	5.52	1.65	0.68
crossing_2_west	11.19	4.32	4.36	3.1	3.85	2.89
crossing_3_east	2.67	1.53	0.63	0.33	0.46	0.28
crossing_3_north	16.48	11.76	4.01	4.89	0.99	26.85
crossing_3_south	20.46	13.15	9.05	6.1	2.94	1.01
crossing_3_west	8.04	3.72	3.27	3.27	12.74	1.36
平均差值	17.21	11.10	7.49	7.96	3.81	5.73

附录 4 GWO 优化分析

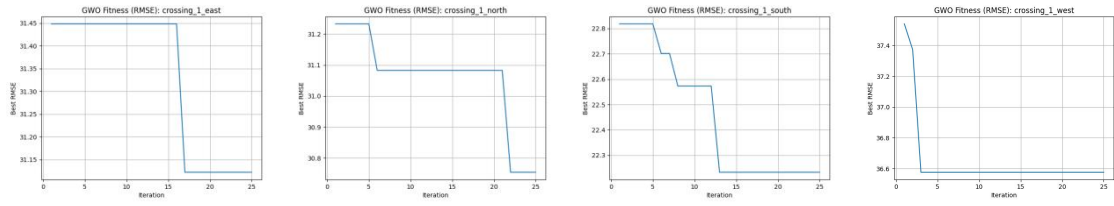


图 22 迭代次数与 RMSE 变化过程(依次为 crossing_1_east/north/south/west_gwo_rmse)

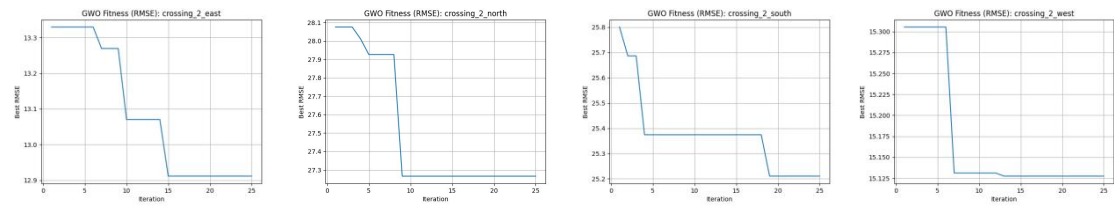


图 23 迭代次数与 RMSE 变化过程(依次为 crossing_2_east/north/south/west_gwo_rmse)

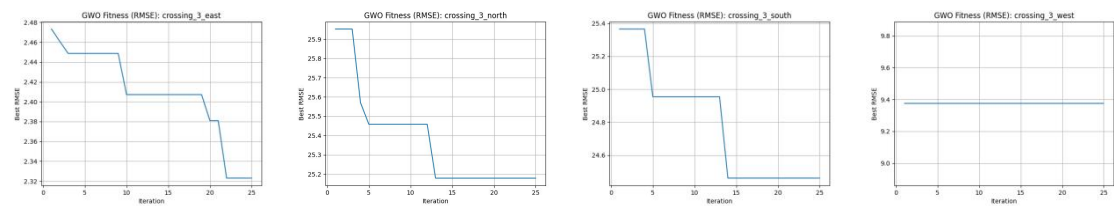


图 24 迭代次数与 RMSE 变化过程(依次为 crossing_3_east/north/south/west_gwo_rmse)

附录 5 EAC 优化分析

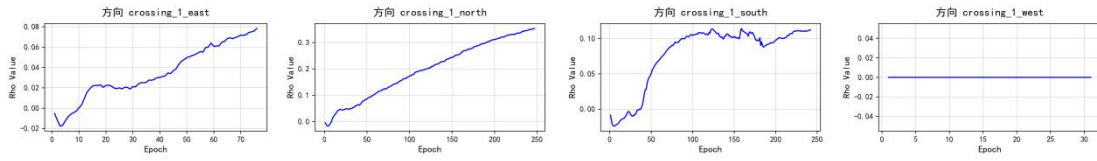


图 25 EAC 随 Epoch 收敛变化曲线依次为(crossing_1_east/north/south/west)

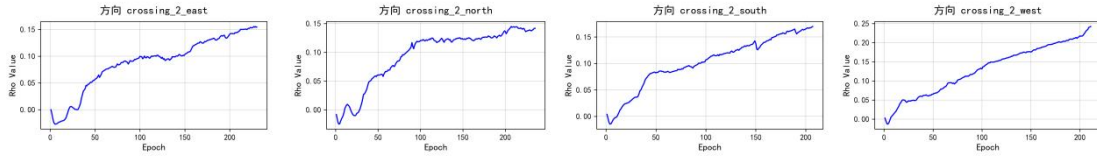


图 26 EAC 随 Epoch 收敛变化曲线(依次为 crossing_2_east/north/south/west)

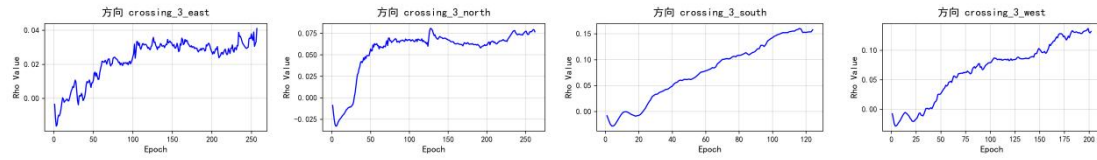


图 27 EAC 随 Epoch 收敛变化曲线(依次为 crossing_3_east/north/south/west)

附录 6 Epochs 和损失函数曲线对比

RNN 的损失函数分析

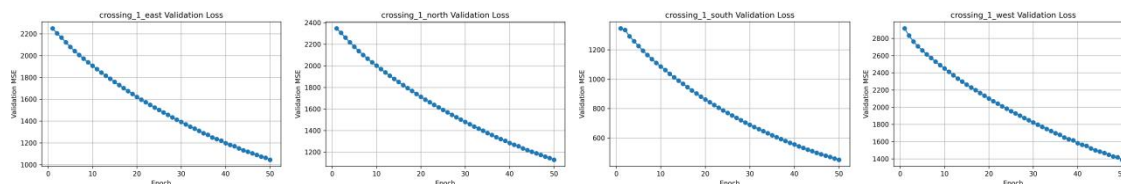


图 28 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_1_east/north/south/west)

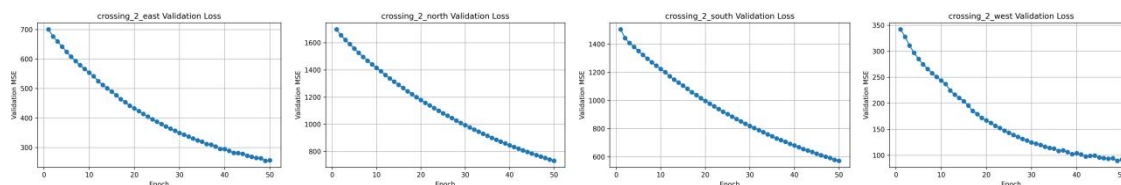


图 29 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_2_east/north/south/west)

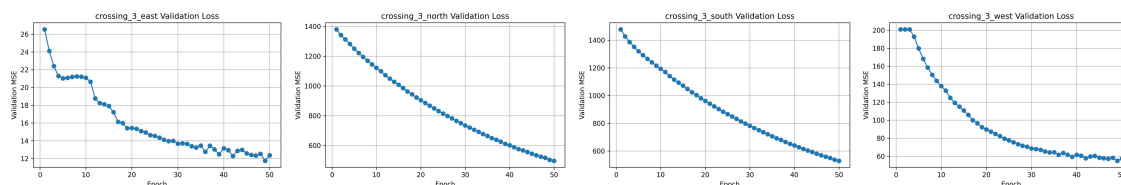


图 30 RNN 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

LSTM 的损失函数分析

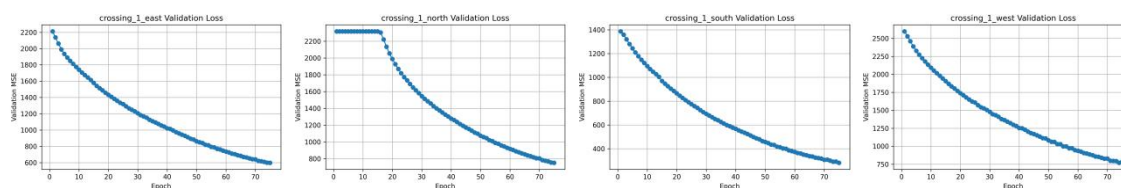


图 31 LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_1_east/north/south/west)

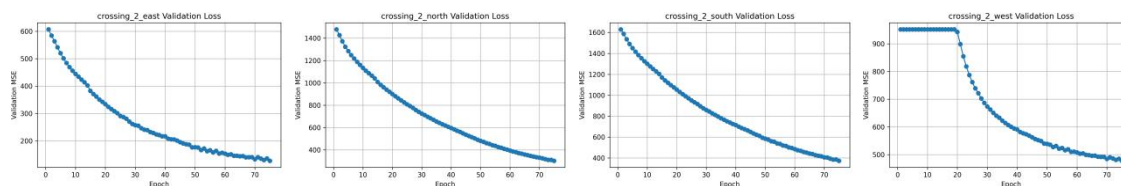


图 32 LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_2_east/north/south/west)

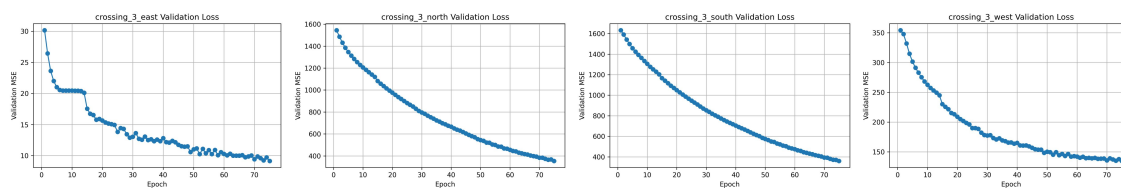


图 33 LSTM 的 MSE 随 Epochs 变化曲线依次为(crossing_3_east/north/south/west)

Attention-LSTM 的损失函数分析

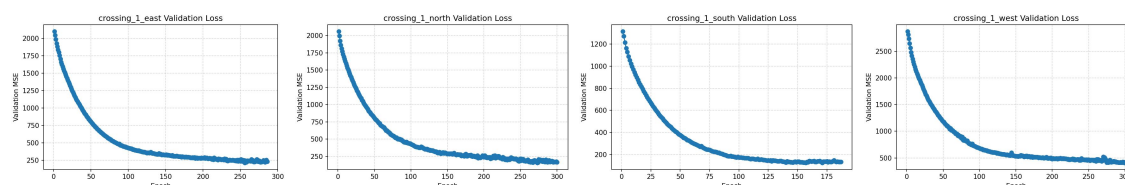


图 34 Attention-LSTM 的 MSE 随 Epochs 变化曲线

依次为(crossing_1_east/north/south/west)

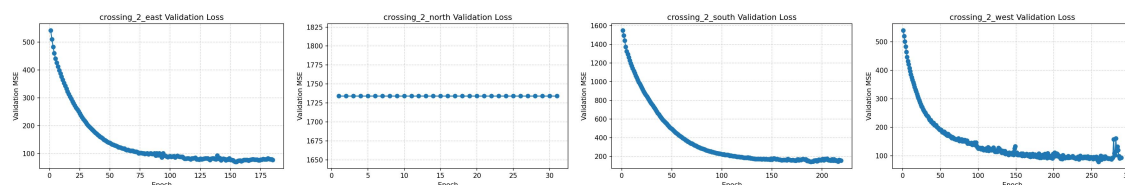


图 35 Attention-LSTM 的 MSE 随 Epochs 变化曲线

依次为(crossing_2_east/north/south/west)

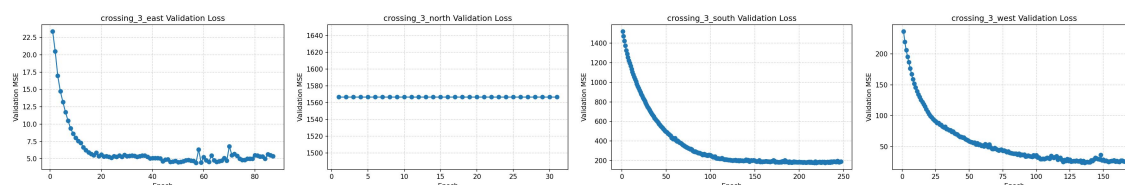


图 36 Attention-LSTM 的 MSE 随 Epochs 变化曲线

依次为(crossing_3_east/north/south/west)

GWO-Attention-LSTM 的损失函数图

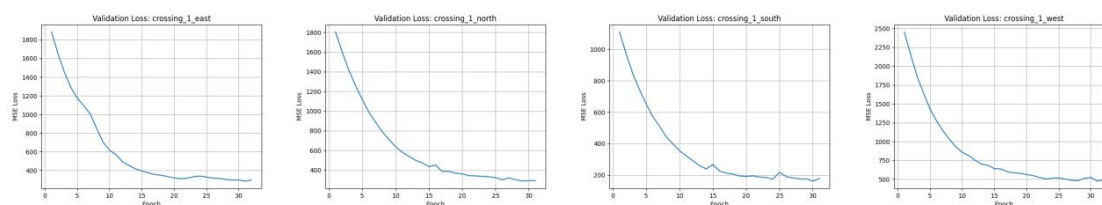


图 37 GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线

依次为(crossing_1_east/north/south/west)

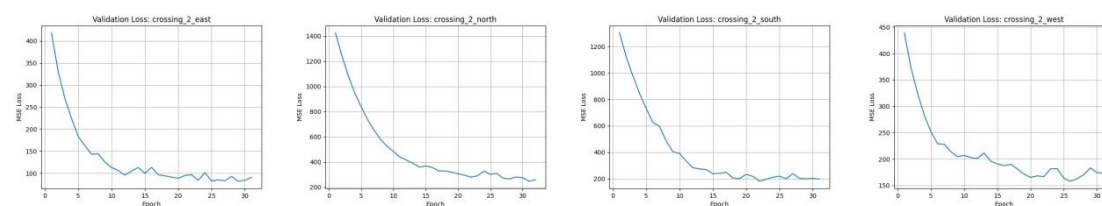


图 38 GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线

(依次为 crossing_2_east/north/south/west)

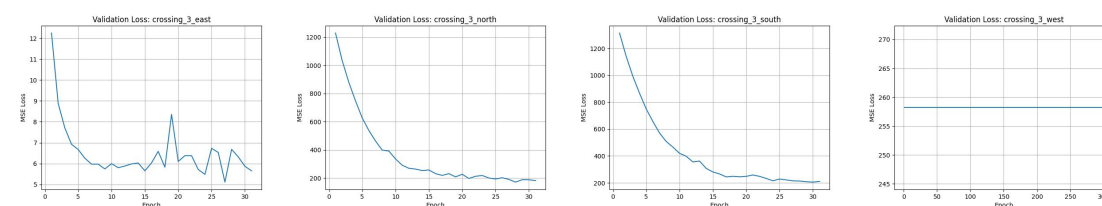


图 39 GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线

(依次为 crossing_3_east/north/south/west)

EAC-Attention-LSTM 损失函数图

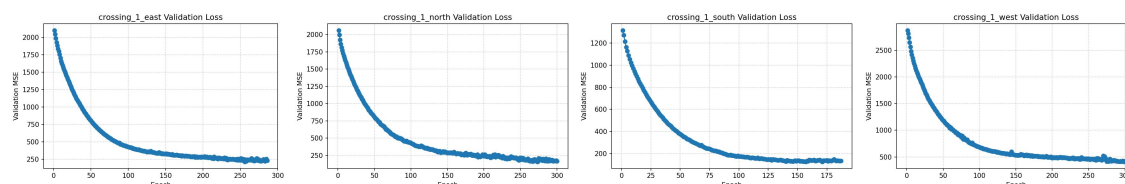


图 40 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线

依次为(crossing_1_east/north/south/west)

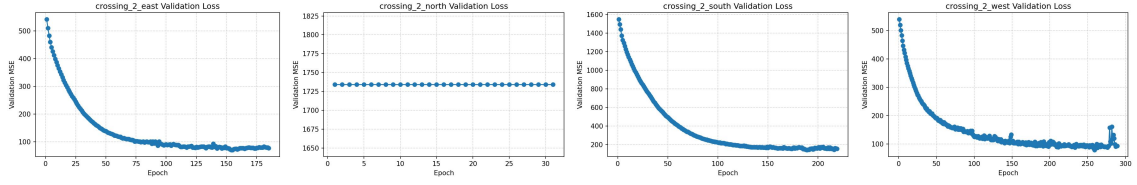


图 41 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_2_east/north/south/west)

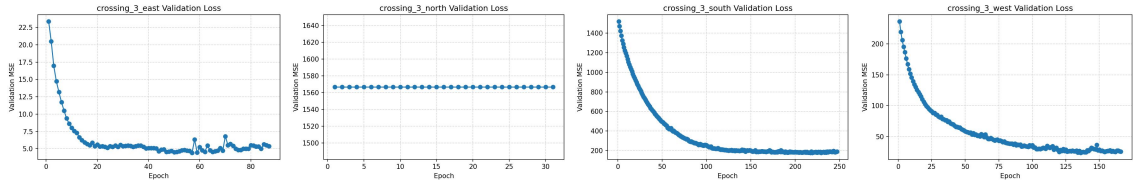


图 42 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_3_east/north/south/west)

EAC-GWO-Attention-LSTM 的损失函数图

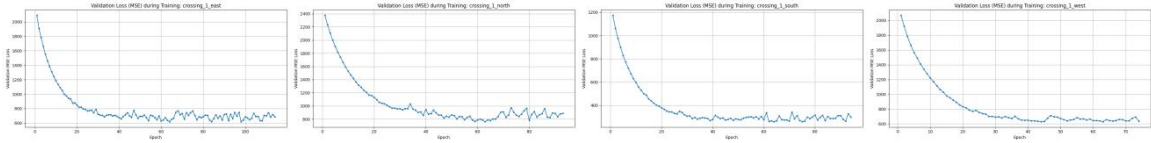


图 43 EAC-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_1_east/north/south/west)

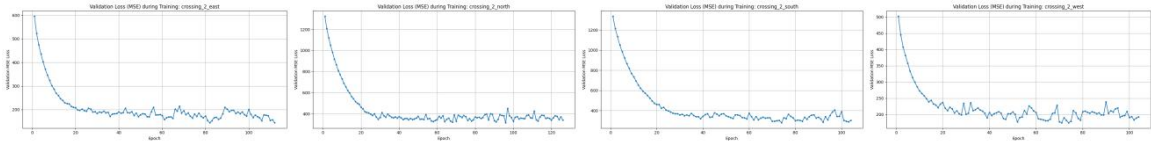


图 44 EAC-GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_1_east/north/south/west)

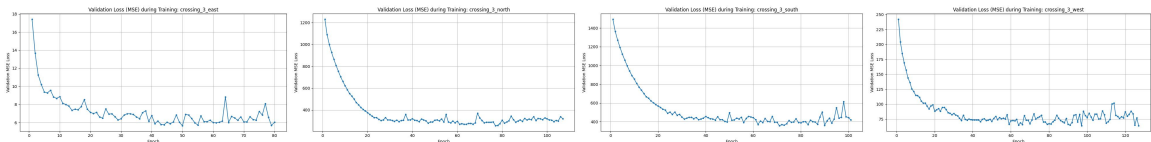


图 45 EAC-GWO-Attention-LSTM 的 MSE 随 Epochs 变化曲线
依次为(crossing_1_east/north/south/west)