



计算机科学

COMPUTER SCIENCE

基于注意力机制和ConvLSTM的船舶交通流量预测算法

李刚, 宋文, 陈致远

引用本文

李刚, 宋文, 陈致远. 基于注意力机制和ConvLSTM的船舶交通流量预测算法[J]. 计算机科学, 2023, 50(11A): 230800067-7.

LI Gang, SONG Wen, CHEN Zhiyuan. Ship Traffic Flow Prediction Algorithm Based on Attention Mechanism and ConvLSTM [J]. Computer Science, 2023, 50(11A): 230800067-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[一种融合CNN和Swin Transformer的医学显微图像分割模型](#)

Medical Microscopic Image Segmentation Model Based on CNN Structure and Swin Transformer
计算机科学, 2023, 50(11A): 230200119-8. <https://doi.org/10.11896/jsjcx.230200119>

[基于GRU与自注意力网络的声源到达方向估计](#)

Sound Source Arrival Direction Estimation Based on GRU and Self-attentive Network
计算机科学, 2023, 50(11A): 220900135-7. <https://doi.org/10.11896/jsjcx.220900135>

[基于TCN-BiLSTM的入侵检测算法研究](#)

Study on Intrusion Detection Algorithm Based on TCN-BiLSTM
计算机科学, 2023, 50(11A): 230300142-8. <https://doi.org/10.11896/jsjcx.230300142>

[基于替代模型的批量零阶梯度符号算法](#)

Batch Zeroth Order Gradient Symbol Method Based on Substitution Model
计算机科学, 2023, 50(11A): 230100036-6. <https://doi.org/10.11896/jsjcx.230100036>

[面向边缘计算的轻量级网络硬件加速设计](#)

Lightweight Network Hardware Acceleration Design for Edge Computing
计算机科学, 2023, 50(11A): 220800045-7. <https://doi.org/10.11896/jsjcx.220800045>

基于注意力机制和 ConvLSTM 的船舶交通流量预测算法

李 刚¹ 宋 文² 陈致远¹

1 国能黄骅港务有限责任公司 河北 沧州 061113

2 山东大学海洋研究院 山东 青岛 266237

摘 要 船舶交通流量预测是港口智能交通系统的关键技术之一,对港口运输的效率和安全性起到至关重要的作用。针对现有预测方法难以有效提取船舶交通流量数据中的时空相关性特征信息的问题,提出了一种基于注意力机制和 ConvLSTM 的预测方法(ACLN)。ACLN 首先通过深度的 ConvLSTM 构建编码网络,有效提取船舶交通流量数据中的时空相关性特征信息;其次通过注意力机制对提取的时空相关特征信息进行重要性关注,使模型在预测的过程中自动关注更重要的特征信息;最后通过多层的 ConvLSTM 和 CNN 构建预测网络,对提取的时空特征信息进行解析并输出预测结果。在真实的港口船舶交通流量数据上验证了所提方法的有效性,实验结果表明,所提方法的预测性能明显优于目前公开的预测方法,能够对一定区域进行有效的长短时预测,具有一定的实用价值。

关键词: 智能交通;智能港口;船舶流量预测;时空特征提取;卷积神经网络;长短期记忆网络;卷积长短期记忆网络;注意力机制
中图分类号 TP393

Ship Traffic Flow Prediction Algorithm Based on Attention Mechanism and ConvLSTM

LI Gang¹, SONG Wen² and CHEN Zhiyuan¹

1 CHN ENERGY Huanghua Port Affairs Co., LTD, Cangzhou, Hebei 061113, China

2 Insitute of Marine Science and Technology, Shandong University, Qingdao, Shandong 266237, China

Abstract Ship traffic flow prediction is one of the key technologies of port intelligent transportation system, which plays a vital role in the efficiency and safety of port transportation. Aiming at the problem that the existing prediction methods are difficult to effectively extract the spatio-temporal feature information from the ship traffic flow data, a prediction method based on attention mechanism and ConvLSTM(ACLN) is proposed. ACLN first constructs a encoding network through the deep ConvLSTM to effectively extract the spatio-temporal feature information from the ship traffic flow data. Secondly, the attention mechanism is used to pay attention to the importance of the extracted spatio-temporal feature information, so that the model can automatically pay attention to the more important feature information in the process of prediction. Finally, the prediction network is constructed by multiple layers of ConvLSTM and CNN to parse the extracted spatiotemporal feature information and output the prediction result. The effectiveness of the proposed method is verified on the real port ship traffic flow data. Experimental results show that the prediction performance of the proposed method is significantly better than that of the state-of-art prediction methods, and it can perform the long and short time prediction effectively in a certain area, and has a certain practical value.

Keywords Intelligent transportation, Intelligent port, Ship flow prediction, Spatio-temporal feature extraction, Convolutional neural network, Long short-term memory network, Convolutional long short-term memory network, Attention mechanism

1 引言

近年来,随着经济全球化快速发展,海运规模不断扩大,港口船舶流量大幅上升,给港口日常运营和管理带来很大压力^[1]。港口船舶流量预测对港口高效、安全运营起着至关重要的作用^[2]。精准的流量预测模型可为港口生产组织、航道管理等业务和人员提供准确的数据支持,以最大限度优化配置港口生产资源、发挥通航能力,从而有效提升整体运营效率和服务水平。船舶交通流量预测是港口智慧化运营管理的关键技术之一,高精度的港口船舶流量预测可以为管理维护人员提供有价值的信息,继而及时调整当前生产和维护

调度计划,提高港口的运营效率^[3]。

目前,船舶交通流量预测的方法主要分为基于传统统计学模型的方法和数据驱动的方法^[4]。基于统计学的方法通常采用定性和定量预测,即通过专家判断,经验分析以及推理总结来进行预测,主要的方法有自回归移动平均(Autoregressive Integrated Moving Average, ARIMA)^[5]、卡尔曼滤波模型^[6]、贝叶斯模型^[7]等。Yang 等^[8]提出了采用基于 ARIMA 模型的船舶交通流量预测方法,通过与非凸低秩稀疏分解模型相结合,提高了预测精度。Li 等^[5]采用 SARIMA 模型建立了船舶交通流量预测模型,实现了对船舶交通流量的长期预测,预测效果相比 ARIMA 预测模型有所提升。Guo 等^[6]

基金项目:国家自然科学基金(62102228);山东省自然科学基金(ZR2021QF063);交通运输部交通强国建设试点任务项目

This work was supported by the National Natural Science Foundation of China(62102228), Shandong Provincial Natural Science Foundation(ZR2021QF063) and Pilot Task Project of the Ministry of Transport for the Construction of a Strong Transportation Country.

通信作者:李刚(gang.li_gg@chnenergy.com.cn)

采用了随机的自适应卡尔曼滤波模型进行短期交通流量预测,与传统卡尔曼滤波相比,该方法具有更好的鲁棒性。由于船舶交通流量具有非线性、周期性和随机性等特点,基于统计学模型的预测方法是在假设环境以及运动限制的情况下进行预测,因此该方法的预测精度较低且无法满足现阶段船舶流量预测的需求。

基于数据驱动的船舶交通流量预测方法主要分为基于传统机器学习方法和深度学习的方法^[9-10]。基于传统机器学习的预测方法主要有 BP 神经网络^[11]、支持向量机(Support Vector Machine, SVM)^[12]以及随机森林^[13]等,已被广泛应用于船舶交通流量预测。但是,基于传统机器学习的预测方法需要依赖于复杂的特征工程才能保证其可靠性,无法实现端到端的动态学习预测,当环境状况发生变化时,基于该方法的预测模型精度无法得到保证。为了能够自适应地提取特征,具有端到端的学习能力,基于深度学习的船舶交通流量预测方法成为目前研究的热点。

当前主要有基于卷积神经网络(Convolutional Neural Networks, CNN)^[14]和循环神经网络(Recurrent Neural Network, RNN)^[15]等深度学习架构的船舶交通流量预测方法。Wang 等^[16]提出了一种基于 CNN 预测方法,利用不同分辨率的六边形进行数据划分并作为 CNN 的输入来预测极端天气下的船舶流量。基于 CNN 网络架构的预测方法由于只能通过不断增大卷积核的尺寸和数量才能提取较大范围的时空特征信息,因此限制了模型在船舶交通流量预测中的性能。Liu 等^[17]将 RNN 模型应用在船舶流量预测上,通过 RNN 模型来捕捉船舶交通流量变化的时间相关性。RNN 在处理较长的时间序列时,由于其循环网络的结构不可避免地会出现梯度消失的问题,因此无法有效提取船舶流量的长期有效信息。长短期记忆网络(Long Short-term Memory Network, LSTM)^[18]是 RNN 网络的一种变体,通过引入 3 种门控单元,有效解决了上述问题。Hao 等^[19]提出将多层双向的 LSTM 网络用于短期的铁客运流量预测中,进一步提高了 LSTM 的预测性能。然而, LSTM 只能提取不同历史时间节点之间的相关性特征信息,无法获取当前区域的空间相关性特征信息。

港口区域内的船舶交通流量预测主要受到时间和空间两方面因素的影响。(1)时间相关性:未来时刻船舶流量的变化会受到当前时刻和历史时刻船舶流量变化的影响;(2)空间相关性:船舶按照航线行驶,且多为合并以及相互交叉等复杂航线,即当前海域的船舶流量变化会受到相邻海域船舶流量的影响,表现出空间相关性。然而,目前的研究只考虑时间相关性或者空间相关性,没有同时融合两方面的特征,限制了船舶交通流量预测模型的精度。同时,对于具有船舶交通流量中更重要特征信息的时间节点,应该给予更多的关注,赋予其更大的重要性权重,而目前的研究中尚没有考虑这一点。

针对上述研究中存在的不足,本文提出了一种新的基于注意力机制(Attention)和卷积长短期记忆网络(Convolutional LSTM Network, ConvLSTM)的船舶交通流量预测模型(Attention and Convolutional LSTM Network, ACLN)。ConvLSTM^[20]对于处理时空预测问题具有一定的优势,能够有效提取多元时间序列数据中的时空特征信息。注意力机制^[21]可以为不同的特征信息分配不同的权重,提高模型的特征提取能力。ACLN 主要包含编码网络和预测网络两大

结构。编码网络中首先通过深度的 ConvLSTM 对预处理后的船舶流量数据进行时空特征提取,其次注意力机制对提取的时空特征信息进行重要性关注,得到具有时空特征关注信息的编码特征。预测网络对编码网络提取的时空特征信息进行解析,最后通过卷积网络输出目标区域的预测流量值。ConvLSTM 充分融合了 CNN 空间特征提取和 LSTM 时间特征提取的优点,能够同时有效提取时空特征信息,且注意力机制可以实现对提取时空特征进行重要性关注,使模型更关注那些更重要的时间节点和区域特征,提高模型的预测性能。通过采集的真实港口的流量数据验证了所提方法的有效性,实验结果表明,本文提出的预测方法与目前公开的预测方法相比,预测性能有明显提升。

2 模型架构

2.1 卷积长短期记忆网络

船舶交通流量预测是一个典型的时空序列预测问题。常见的基于 RNN 架构的预测模型(RNN, LSTM, GRU 等)只能从数据中提取时间相关性特征信息,无法获得空间相关性特征信息,大大影响了模型的预测性能。因此,为了解决传统 LSTM 网络无法有效同时提取船舶交通流量数据时间和空间特征信息的问题,本文提出在编码网络中采用多层的 ConvLSTM^[20]提取船舶交通流量数据中的时空特征信息。ConvLSTM 是在 LSTM 网络基础上的改进,通过将 LSTM 网络输入和隐藏状态与各个门结构之间的全连接方式替换为卷积层,在捕捉时空序列数据时间相关性特征信息的同时也会考虑到空间相关性,图 1 和图 2 为 LSTM 和 ConvLSTM 的结构对比,ConvLSTM 的计算过程如下:

$$\begin{cases} f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ C_{t-1} + b_f) \\ i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ C_{t-1} + b_i) \\ \tilde{C}_t = \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \\ C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \\ o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ C_{t-1} + b_o) \\ h_t = o_t \circ \tanh(C_t) \end{cases} \quad (1)$$

其中, W_{xf}, W_{hf}, W_{cf} 是遗忘门的权重; W_{xi}, W_{hi}, W_{ci} 是输入门的权重; W_{xi}, W_{hi}, W_{ci} 是输入门的权重; W_{xc}, W_{hc} 是单元状态的权重; W_{xo}, W_{ho}, W_{co} 是输出门的权重; “*” 表示卷积计算。可以看出 ConvLSTM 相比 LSTM 可以通过在多维数据中进行卷积操作来捕捉空间特征,其同时具备 LSTM 网络的特点,可以有效地获取不同时间步的时间特征,并有效提取船舶交通流量数据中的时空特征信息。

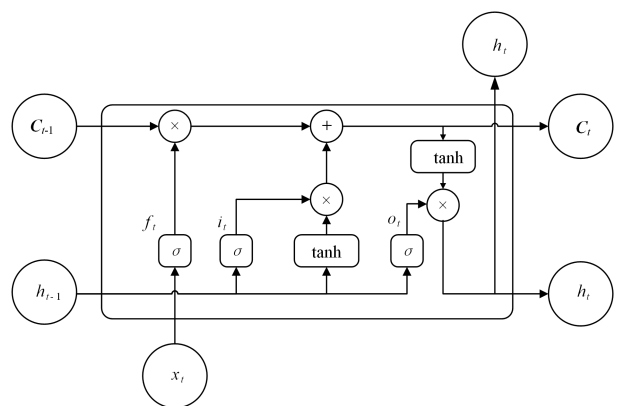


图 1 LSTM 模型结构

Fig. 1 Structure of LSTM

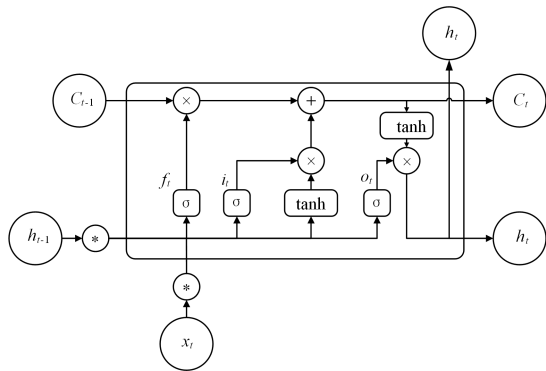


图2 ConvLSTM 模型结构

Fig. 2 Structure of ConvLSTM

2.2 注意力机制

注意力机制^[21]可以实现对时间序列中不同时间步的重要性进行关注,使模型自动关注含有更多重要特征信息的单元,提高模型的预测性能。Attention 本质是通过计算特征向量的加权和实现对不同时间步或空间区域的重要性关注,其计算主要分为以下 3 个步骤。首先,将输入映射为查询 Query 和

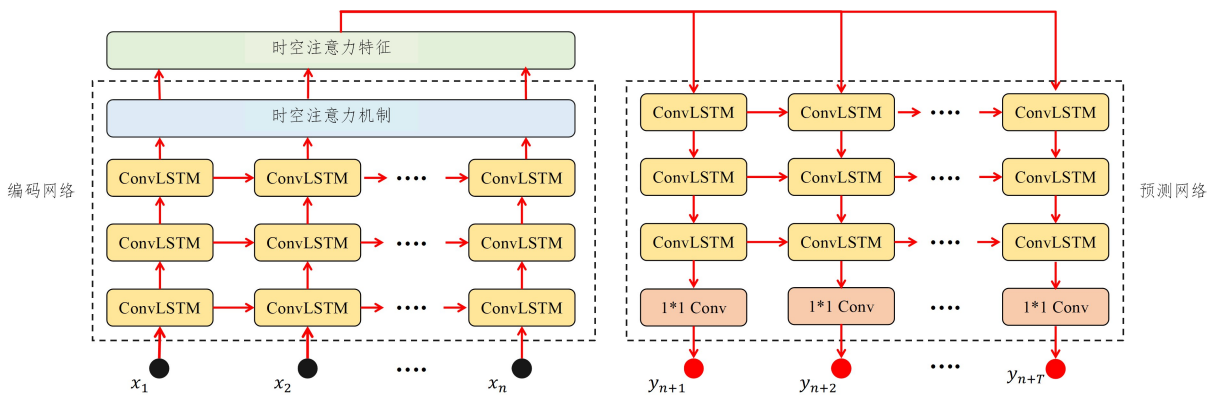


图3 基于 ACLN 的船舶交通流量预测模型架构

Fig. 3 Model architecture of ship traffic flow prediction based on ACLN

2.3.1 编码网络

如图 3 所示,编码网络由 3 层 ConvLSTM 单元组成。给定长度为 n 的船舶交通流量序列 $\mathbf{X} = \{x_t | t=1, 2, \dots, n\}$ 作为编码网络的输入,每一层 ConvLSTM 单元提取的时空特征信息会作为下一层的 ConvLSTM 单元的输入。当输入船舶交通流量序列中的 x_t 时,ConvLSTM 单元会通过卷积操作来确定未来单元状态 C_t^l 以及是否激活 3 个门,在经过多层的 ConvLSTM 单元后可以得到提取的时空特征信息 h_t^l ,计算过程如下:

$$\begin{cases} f_t^l = \sigma(\mathbf{W}_{xf}^l * x_t + \mathbf{W}_{hf}^l * h_{t-1}^l + b_f^l) \\ i_t^l = \sigma(\mathbf{W}_{xi}^l * x_t + \mathbf{W}_{hi}^l * h_{t-1}^l + b_i^l) \\ C_t^l = f_t^l \circ C_{t-1}^l + i_t^l \circ \tanh(\mathbf{W}_{xc}^l * x_t + \mathbf{W}_{hc}^l * h_{t-1}^l + b_c^l) \\ o_t^l = \sigma(\mathbf{W}_{xo}^l * x_t + \mathbf{W}_{ho}^l * h_{t-1}^l + b_o^l) \\ h_t^l = o_t^l \circ \tanh(C_t^l) \end{cases} \quad (3)$$

其中, i_t^l, f_t^l, o_t^l 分别为输入门、遗忘门、输出门; h_t^l 为第 l 层 ConvLSTM 单元的隐藏状态; C_t^l 为第 l 层 ConvLSTM 单元中记忆单元的状态; \mathbf{W}_{xi}^l 和 \mathbf{W}_{hi}^l 为第 l 层 ConvLSTM 单元输入门的权重; \mathbf{W}_{xf}^l 和 \mathbf{W}_{hf}^l 为第 l 层 ConvLSTM 单元遗忘门的权重; \mathbf{W}_{xo}^l 和 \mathbf{W}_{ho}^l 为第 l 层 ConvLSTM 单元输出门的权重; \mathbf{W}_{xc}^l 和 \mathbf{W}_{hc}^l 为第 l 层 ConvLSTM 单元状态的权重; $b_f^l, b_i^l, b_o^l, b_c^l$ 分别为第 l

键值对(Key, Value),将 Query 和每个 Key 进行点积,通过对结果进行缩放和 Softmax 函数得到 Attention 权重;然后将 Attention 权重和对应的 Value 求取加权和获得 Attention 的输出,计算公式如下:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (2)$$

其中, $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 分别表示 Query, Key, Value 矩阵, d_k 表示输入向量的映射维度。

2.3 基于注意力机制和 ConvLSTM 的船舶交通流量预测模型

本文所提出的基于 ACLN 的船舶交通流量预测模型如图 3 所示,主要包括编码网络、时空注意力机制、预测网络 3 个部分,模型的输入是采集的历史船舶交通流量数据。首先编码网络采用多层的 ConvLSTM 从原始输入数据中提取时空相关性特征,采用注意力机制对提取的时空特征信息重要性加权,使模型在训练的过程中关注对预测结果更重要的时间节点和空间区域信息;然后预测网络对编码网络提取的时空相关性特征信息进行解码;最后通过 1×1 的卷积层输出预测值。下面将详细介绍 ACLN 的计算过程。

层 ConvLSTM 单元输入门、遗忘门、输出门、记忆单元的偏置。

2.3.2 注意力机制计算过程

经过深度 ConvLSTM 提取的时空特征 $\{h_t^l | t=1, 2, 3, \dots, n; l=3\}$ 将作为时空注意力机制的输入,进行重要性特征信息关注。通过对提取的时空特征信息进行线性映射获得 Query, Key, Value 矩阵:

$$\mathbf{Q}_h = \mathbf{W}^Q h_t^3, \mathbf{K}_h = \mathbf{W}^K h_t^3, \mathbf{V}_h = \mathbf{W}^V h_t^3 \quad (4)$$

其中, $\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V$ 为可学习的参数。然后通过计算 \mathbf{Q}_h 和 \mathbf{K}_h 的点积运算,并沿着时间维度应用 Softmax 函数即可获得注意力权重:

$$\alpha_t^k = \text{softmax}\left(\frac{\mathbf{Q}_h \mathbf{K}_h^T}{\sqrt{d_k}}\right) \quad (5)$$

其中, α_t^k 表示编码网络中第 k 个隐藏状态 h_t^k 对船舶流量预测的重要性, d_k 代表线性映射的输入维度。通过计算编码网络中提取的时空特征信息加权和,得到注意力机制的输出:

$$a_t = \alpha_t^k \mathbf{V}_h \quad (6)$$

通过上述计算,即可获得带有注意力加权的时空特征信息 a_t 。

2.3.3 预测网络

预测网络由 3 层 ConvLSTM 和 1 层 CNN 组成,编码

网络中提取的时空特征信息经过时空注意力机制进行重要性关注后作为预测网络的输入。预测模型的任务是通过前 n 个时间步信息来预测未来 T 个时间步的船舶交通流量 $\{\hat{y}_t | t = n+1, n+2, \dots, n+T\}$ 。使用上一时刻的预测结果 \hat{y}_{t-1} 作为 t 时刻下的网络输入值,进行滚动预测, \hat{y}_{t-1} 由编码网络和注意力机制共同计算得到:

$$\hat{y}_{t-1} = \begin{cases} W[h_n^3; a_{t-1}] + b, & t = n+1 \\ W[\hat{y}_{t-1}; a_{t-1}] + b, & n+2 \leq t \leq T \end{cases} \quad (7)$$

其中, T 表示预测模型的预测序列长度,在 $t = n+1$ 时,预测网络的输入为编码网络 T 时刻下的最终隐藏状态 h_n^3 经过注意力机制加权后的结果;在 $n+2 < t < T$ 时,预测网络的原输入都为上一时刻得到的预测结果 \hat{y}_{t-1} ; W 和 b 将 $[h_n^3; a_{t-1}]$ 和 $[\hat{y}_{t-1}; a_{t-1}]$ 映射到预测网络输入的大小。预测网络的 ConvLSTM 层计算过程与编码网络一致,目的是为了解析带有注意力机制加权的时空特征信息,将预测网络最后一层 ConvLSTM 的隐藏状态 d_t^l 作为 1×1 卷积层的输入,即可获得船舶流量的最终预测值。

3 实验与分析

3.1 实验数据集

本文采用 AIS^[2,22] 的港口真实船舶流量数据来验证所提出方法的有效性。AIS 数据被广泛应用在船舶交通中,以解决船舶交通流量预测、船舶到达港口时间预测以及确定最佳航线等问题,为海事专业人员加强海上交通安全和提高船舶运行效率提供有效的信息。表 1 列出了 AIS 数据记录的信息。本文通过研究分析 AIS 数据中的海上移动业务识别码 (Maritime Mobile Service Identify, MMSI) 经度、纬度、定位时间来对某区域内的船舶交通流量进行预测。其中, MMSI 是船舶的唯一标识码,通过 MMSI 能够区分船舶;经纬度是为了筛选出预测区域内的 AIS 数据;定位时间是为了统计船舶流量。

表 1 AIS 数据信息

Table 1 AIS data information

船舶静态数据	船舶动态数据	航次数据
船舶识别码(MMSI)	纬度(LAT)	航线
国际海事组织编号	经度(LON)	货物种类
船名	对地航向(COG)	目的港
船舶呼号	对地航速(SOG)	到达时间
船舶类型	转向速率(ROT)	船舶吃水
GPS 定位天线位置	定位时间(UTC)	

3.2 预测区域划分

根据经纬度坐标筛选出预测区域的 AIS 数据,从而将预测区域划分为一个矩形区域,预测区域中经纬度的最高点和最低点分别为 $p_h = (lon_h, lat_h)$ 和 $p_l = (lon_l, lat_l)$, 其中 lon_h 和 lat_h 分别表示最高点的经度和纬度; lon_l 和 lat_l 分别表示最低点的经度和纬度。本文将预测区域划分为 $a \times b$ 个网格,每个网格都代表一个小的预测区域,用 G 表示预测区域中网格的集合:

$$G = \{g_{ij} | i \in [1, b], j \in [1, a]\} \quad (8)$$

其中, g_{ij} 表示预测区域中第 i 行、第 j 列的网格。预测区域中存在的船舶流量包括船舶流入量和船舶流出量,如图 4 所示,通过对预测区域进行网格划分,可以得到一段时间间隔内

预测区域中 $a \times b$ 个网格的船舶流入量和船舶流出量。

图 5 为通过 AIS 获取的国内某港口 2013 年 10 月 1 日某一区域一天的船舶流入量,采样时间间隔为 5 分钟,本文选取此海域 2013 年 10 月 1 日至 2013 年 10 月 31 日及其附近海域的 AIS 数据,共得到 8525 个样本数据,其中每个样本数据都包含预测区域内 7×7 个网格的船舶流入量和船舶流出量。选取 2013 年 10 月 1 日至 2013 年 10 月 25 日的数据作为训练集用于训练模型,共 6305 个样本数据;选取 2013 年 10 月 26 日至 2013 年 10 月 31 日的数据作为测试集用于评估模型性能,共 2220 个样本数据。

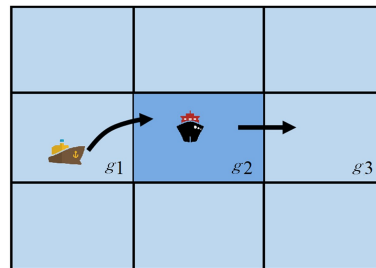


图 4 港口船舶流入/流出示意图

Fig. 4 Diagram of ship inflow/outflow in port

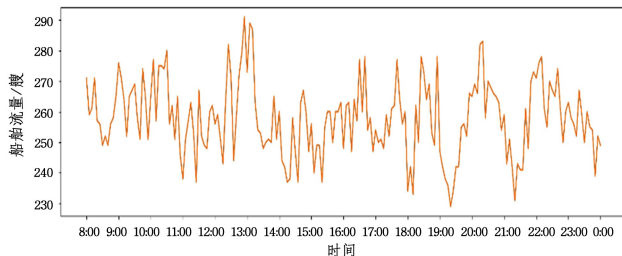


图 5 2013 年 10 月 1 日的船舶流入量

Fig. 5 Ship inflows on 1 October 2013

3.3 实验环境及参数设定

本文中的所有实验均在 windows 10 平台上进行,并配备了 NVIDIA RTX 2080 SUPER 显卡和 Intel 9900K CPU。本文使用 Python 语言和基于 Pytorch 的深度学习框架实现所提出的船舶交通流量预测模型以及用于比较其他预测模型。训练过程中的实验参数如表 2 所列,损失函数采用平均绝对误差损失函数 (Huber 损失函数),计算公式如式(11)所示,并采用 Dropout 技术来防止模型过拟合。

$$L_{\delta}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2, & |y - \hat{y}| \leq \delta \\ \delta|y - \hat{y}| - \frac{1}{2}\delta^2, & |y - \hat{y}| > \delta \end{cases} \quad (9)$$

其中, δ 为 Huber 损失函数的参数,当误差在 $[-\delta, +\delta]$ 时,等价于 MSE 损失函数,其在最优解附近,梯度更新会随着误差的减小而减小,更有利于收敛;当误差在 $[-\infty, \delta]$ 或 $[\delta, +\infty]$ 时,等价于 MAE 损失函数,因此具有较好的鲁棒性。

表 2 训练过程实验参数

Table 2 Experimental parameters used in training process

参数名称	参数设置	参数名称	参数设置
训练次数	100	优化器	Adam
学习率	1×10^{-3}	Dropout rate	0.5
批处理大小	24	损失函数	Huber

3.4 预测模型结构参数

为了确定所提预测模型的结构参数,我们进行网格搜索

以找到最佳的模型配置。本文提出的基于注意力机制和 ConvLSTM 的船舶交通流量预测模型结构参数细节如表 3 所列,其中编码网络和预测网络均由三层的 ConvLSTM 单元组成,每层的隐藏状态单元维度分别为 128,64,2,预测网络中最后通过一层二维 1×1 卷积进行输出。

表 3 预测模型参数设置

Table 3 Parameter settings of prediction model

网络层	单元层	参数设置
编码网络	ConvLSTM 单元 (三层)	卷积核大小: 5×5 ; 步长:1;填充:2; 隐舍单元: {128,64,2}
	ConvLSTM 单元 (三层)	卷积核大小: 5×5 ; 步长:1;填充:2; 隐舍单元: {128,64,2}
预测网络	二维卷积层	卷积核大小: 1×1 ; 步长:1;

3.5 预测模型评价指标

为了有效地评价所提出的船舶交通流量预测模型的性能,本文采用 4 种常见的预测模型评价指标,分别为平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Squared Error, RMSE)、标准化均方根误差 (Normalized Root Mean Square Error, NRMSE)、标准化平均绝对误差 (Normalized Mean Absolute Error, NMAE),其计算方式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

$$NRMSE = \frac{RMSE}{\bar{y}} \quad (12)$$

$$NMAE = \frac{MAE}{\bar{y}} \quad (13)$$

3.6 实验结果评估

为了有效地验证所提出的船舶交通流量预测模型的性能,本节分别对影响模型的多个因素进行了实验,并与当前公开的先进的预测模型进行了对比,为了保证实验结果的准确性,各项实验均重复 5 次取均值作为最终结果。

3.6.1 不同时间窗口大小对模型的影响

所提方法中编码网络输入窗口的大小对模型的预测性能有一定的影响,为了选取合适的窗口大小参数,本文进行了多组窗口实验结果,分别将窗口大小设置为 5,10,15,20,25 进行实验验证, RMSE 和 MAE 实验结果分别如图 6 和表 4 所示。可以看出,尽管较大的时间窗可以包含更多的历史信息,但是当时间窗口过大时也会包含更多噪声信息,影响模型的学习效果,当窗口大小为 10 时预测模型性能最佳,因此在后续实验中我们设置窗口大小为 10。

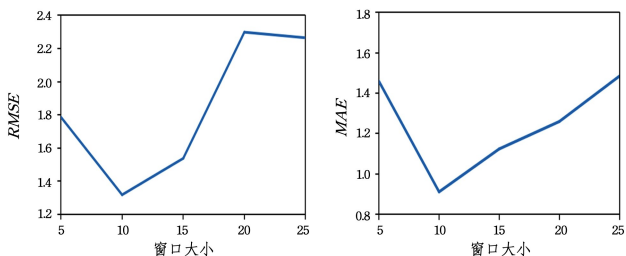


图 6 不同滑动时间窗口的实验结果

Fig. 6 Results with different sliding time windows

表 4 不同时间窗口的预测性能指标

Table 4 Prediction results of different time windows

窗口大小	RMSE	MAE
5	1.7856	1.4582
10	1.3169	0.9112
15	1.5369	1.1236
20	2.2985	1.2593
25	2.2647	1.4841

3.6.2 不同时间窗口大小对模型的影响

为验证所提方法的有效性,本文与近期具有代表性的预测方法进行船舶交通流量预测对比实验,包括基于支持向量回归 (SVR)^[22]、长短期记忆网络 (LSTM)^[23]、门控循环单元 (GRU)^[2]、双向 LSTM 和卷积神经网络 (BDLSTM-CNN)^[24] 方法的流量预测模型。为了验证所提出的注意力机制模型对预测性能提升的有效性,本文进行了模型的消融实验,与不含有注意力机制的预测模型进行了实验对比,预测方法的实验结果如表 5 所列,其中 ACLN-w/o 代表本文提出的不含有注意力机制的预测模型。可以看出,本文所提出的基于 ACLN 的预测方法在 4 个指标上都有明显的提升,相比目前公开的性能最好的预测方法在 MAE, RMSE, NMAE, NRMSE 4 个指标上预测性能分别提升 31.36%, 30.55%, 44.40%, 43.37%。

表 5 不同预测方法性能对比

Table 5 Performance comparison of different methods

预测方法	MAE	RMSE	NMAE	NRMSE
SVR ^[22]	2.13	2.74	0.50	0.51
LSTM ^[23]	1.29	1.89	0.24	0.22
GRU ^[2]	1.43	2.27	0.27	0.26
BDLSTM-CNN ^[24]	1.34	2.04	0.25	0.24
ACLN-w/o Attention	0.94	1.54	0.18	0.18
ACLN	0.91	1.31	0.13	0.12
性能提升/%	31	30	44	43

从表 5 中可以看出,基于传统机器学习方法 SVR 的预测性能与其他基于深度学习方法的预测性能相比明显较差,这是因为传统浅层机器学习的方法必须依赖特征工程才能保证其性能,因此限制了它的应用场景。基于 LSTM 和 GRU 的预测方法相比基于 SVR 的方法预测性能有所提升,但其只考虑了船舶交通流量预测中的时间相关性,忽略了空间相关性,因此降低了该方法的预测性能。基于 BDLSTM-CNN 的方法只是将两种方法简单组合来提取船舶交通流量数据的时空关系,没有充分考虑到两种特征之间的非线性关系,因此预测性能明显低于本文所提出的预测方法。同时可以看出,本文所提出的 ACLN 的预测方法相比没有注意力机制的方法预测性能明显提升,这是因为注意力机制可以有效地使模型自动关注更重要的时空特征信息,提高模型的预测能力。

3.6.3 预测模型结果分析

为了直观地分析所提出的基于 ACLN 的船舶交通流量预测模型的预测性能,本文给出了不同预测模型的预测流量值和真实流量值之间的对比,我们按照船舶交通流量预测中广泛采用的滚动向前一步预测的方式对测试集的样本进行预测。如图 7—图 10 所示,我们分别对比了基于 LSTM, GRU, BDLSTM-CNN 以及所提出的 ACLN 的预测效果。通过船舶流量预测值和真实值的对比可以看出,本文所提出的基于 ACLN 的预测方法性能明显优于其他方法,不管是在预测的

前期还是中后期,预测值都能很好地拟合船舶流量的真实值,证明了所提方法的有效性。相比于传统基于 LSTM 或 CNN 网络架构的预测方法,所提出的基于 ACLN 的预测方法可以有效地同时提取船舶流量中的时间和空间相关性特征信息,提高了模型对时空信息的处理能力。同时注意力机制可以使模型在预测的过程中关注对预测结果更重要的时空特征信息,进一步提高了预测性能。实验分析表明,ACLN 有效提高了船舶流量的预测精度并具备良好的长短期预测性能,具有较好的实用性。

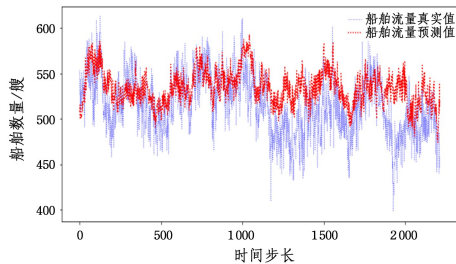


图 7 基于深度 LSTM 网络的预测结果

Fig. 7 Prediction results based on deep LSTM network

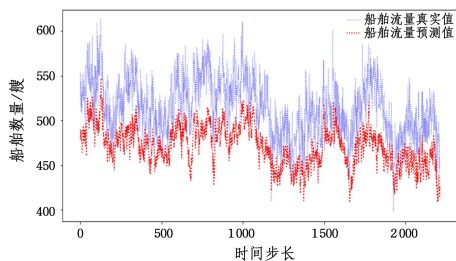


图 8 基于深度 GRU 网络的预测结果

Fig. 8 Prediction results based on deep GRU network

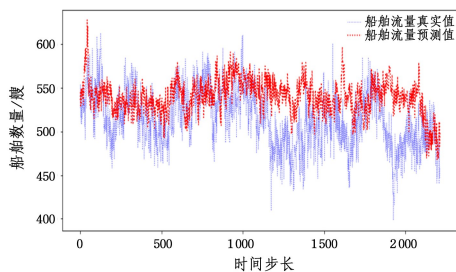


图 9 基于 BDLSTM-CNN 的预测结果

Fig. 9 Prediction results based on BDLSTM-CNN

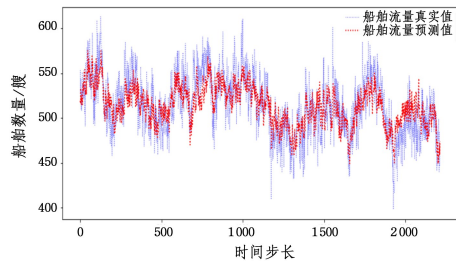


图 10 基于 ACLN 的预测结果

Fig. 10 Prediction results based on ACLN

结束语 针对现有预测方法难以有效同时提取船舶交通流量中的时间相关性特征和空间相关性特征的问题,提出了一种基于注意力机制和 ConvLSTM 的船舶交通流量预测方法。该方法首先通过多层的 ConvLSTM 构建编码网络,来

提取船舶交通流量数据中的时空相关性特征;其次注意力机制对提取的时空相关性特征进行重要性关注,使模型在预测的过程中关注更重要的特征信息;最后通过多层的 ConvLSTM 和 CNN 构建预测网络,对提取的带有注意力关注的时空特征进行解析并输出预测结果。通过 AIS 的真实港口流量数据验证了所提方法的有效性,实验结果表明,与目前公开的船舶流量预测方法相比,本文所提出的基于 ACLN 的船舶流量预测方法明显具有更好的预测性能,能够实现一定区域内的船舶流量进行有效的长短时预测。

参考文献

- [1] YAN X P, LIU C G. Review and prospect for intelligent waterway transportation system [J]. CAAI Transactions on Intelligent Systems, 2016, 11(6): 807-817.
- [2] SUO Y F, CHEN W K, YANG S H, et al. Prediction of ship traffic flow based on deep neural network [J]. Journal of Jimei University (Natural Science), 2020, 25(6): 430-436.
- [3] HE W, ZHONG C, SOTELO M A, et al. Short-term vessel traffic flow forecasting by using an improved Kalman model [J]. Cluster Computing, 2019, 22(S4): 7907-7916.
- [4] WANG K, LIU W, CHEN J W, et al. Study on method of ship traffic flow prediction based on VMD and LSTM [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2022, 46(1): 177-182.
- [5] LI X L, XIAO J L, LIU M J. Vessel traffic flow prediction based on the SARIMA model [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2017, 41(2): 329-332, 337.
- [6] GUO J, HUANG W, WILLIAMS B M. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification [J]. Transportation Research Part C: Emerging Technologies, 2014, 43: 50-64.
- [7] WANG R F, CHEN Z Q, LIU Z X. Link prediction in complex networks with syncretic naive Bayes methods [J]. CAAI Transactions on Intelligent Systems, 2019, 14(1): 99-107.
- [8] YANG S S, WU C S, LIU Z, et al. Vessel traffic flow prediction using non-convex low-rank and sparsity constraints [J]. Application Research of Computers, 2018, 35(1): 43-47.
- [9] YAN Z, YU C C, HAN L, et al. Short-term traffic flow forecasting method based on CNN+LSTM [J]. Computer Engineering and Design, 2019, 40(9): 2620-2624, 2659.
- [10] HUANG W, SONG G, HONG H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2191-2201.
- [11] LAI J H, LIANG S. Short-time traffic flow forecasting method based on BP neural network optimized by ACS [J]. Computer Engineering and Applications, 2014, 50(10): 244-248.
- [12] LI M, HAN D, WANG W. Vessel traffic flow forecasting by RS-VR with chaotic cloud simulated annealing genetic algorithm and KPCA [J]. Neurocomputing, 2015, 157: 243-255.
- [13] XIONG T, QI Y, ZHANG W B, et al. Short term traffic flow forecasting method based on temporal-spatial correlation [J]. Computer Engineering and Design, 2019, 40(2): 501-507.
- [14] DENG S, JIA S, CHEN J. Exploring spatial-temporal relations via deep convolutional neural networks for traffic flow prediction

- with incomplete data [J]. *Applied Soft Computing*, 2019, 78: 712-721.
- [15] ZHANG W, WU P, PENG Y, et al. Roll motion prediction of unmanned surface vehicle based on coupled CNN and LSTM [J]. *Future Internet*, 2019, 11(11): 243.
- [16] WANG X, LI J, ZHANG T. A machine-learning model for zonal ship flow prediction using AIS Data: a case study in the south atlantic states region [J]. *Journal of Marine Science and Engineering*, 2019, 7(12): 463.
- [17] LIU J X, LIU Z D, ZHOU F. A marine traffic flow forecasting model based on generalized regression neural network [J]. *Navigation of China*, 2011, 34(2): 74-77.
- [18] LI Q Y, JIN J, WANG B. Automatic sleep staging model based on the bi-directional LSTM convolutional network and attention mechanism [J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(3): 523-530.
- [19] HAO S, LEE D, ZHAO D. Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system [J]. *Transportation Research Part C: Emerging Technologies*, 2019, 107: 287-300.
- [20] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: a machine learning approach for Precipitation nowcasting [J]. *Advances in Neural Information Processing Systems*, 2015, 28: 802-810.
- [21] MNINH V, HEES H, GRAVES A, et al. Recurrent models of visual attention [C]// *Advances in Neural Information Processing Systems*, 2014.
- [22] KALTEH A M. Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform [J]. *Computers & Geosciences*, 2013, 54: 1-8.
- [23] XIE Z, LIU Q. LSTM networks for vessel traffic flow prediction in inland waterway [C]// *IEEE International Conference on Big Data and Smart Computing (BigComp)*. 2018: 418-425.
- [24] ZHOU X, LIU Z, WANG F, et al. Using deep learning to forecast maritime vessel flows [J]. *Sensors*, 2020, 20(6): 1761.



LI Gang, born in 1976, postgraduate, senior engineer. His main research interests include smart port construction, information and intelligence.