Forecasting Hand-Foot-and-Mouth Disease in Qingdao City Based on CNN-BiLSTM-Attention Model

Feiyan Zhang [1, a], Yan Liu [2, *, b], Haiping Duan [2, c], Yiwan Lai [1, d], Jing Jia [2, e], and XueLing Xin [2, f]

[1] School of Computer Science and Technology Ocean University of China,QingDao, China;

[2] Centers for DiseasePrevention Qingdao Municipal Center for Disease Control and Prevention QingDao, China.

[a] Zhangfeiyan9752@stu.ouc.edu.cn,*Corresponding Author:Yan Liu ,b ly2505@163.com

Abstract. Purpose:A deep learning model CNN-BiLSTM-Attention was proposed, which realized the prediction of the outbreak rule and trend of HFMD in Qingdao, and could provide scientific basis for the relevant departments to formulate preventive measures..Methods:Based on the data of temperature, humidity and HFMD in Qingdao city, a multi-source heterogeneous data set was constructed and a long-term time series prediction model was proposed based on attention mechanism, which can predict the future incidence trend of HFMD. Then, ablation experiments were performed to verify the importance of each module, and comparison experiments with infectious disease dynamic model (SEIIeQR) and statistical model (ARIMA, SARIMA) proved the accuracy of the proposed model..Results:The ablation experiments demonstrated that the CNN module,Attention module,and BiLSTM module could effectively enhance the model's performance.The comparative experiment results revealed the superiority of the deep learning model in predicting HFMD outbreaks in Qingdao City.Conclusion:The multi-source heterogeneous dataset incorporating average temperature and relative humidity data significantly improved the model's accuracy.The CNN-BiLSTM-Attention model outperforms traditional methods in predicting HFMD incidence in Qingdao City,thereby assisting relevant departments in Qingdao to adopt scientific preventive measures against HFMD.

Keywords: HFMD; CNN-BiLSTM-Attention; Disease Prediction.

1. Introduction

Hand-Foot-and-Mouth Disease (HFMD), as a highly contagious disease caused by a specific enterovirus, is especially active in summer and autumn, and its main victim group is infants, and severe cases may even lead to death.[1]. HFMD is prone to causing large-scale outbreaks, posing a severe threat to public health. In view of its powerful infectious capacity and transmission speed, it is particularly important to accurately predict its incidence trend and formulate effective prevention and control strategies.In existing HFMD prediction research, two primary methods are widely applied: the first is mathematical modeling based on case data[2], which is relatively straightforward and convenient to operate, but can only train univariate data, insufficiently handling complex nonlinear relationships and failing to fully consider other important factors related to HFMD incidence, thereby limiting further improvements in prediction accuracy. The second method utilizes deep learning techniques[3], such as Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN) models, which possess the capability to process nonlinear and multivariate data[4]. However, the currently employed model structures are relatively simple, making it difficult to extract significant features from complex and dynamic data and adapt to changes in the importance of data across different time periods.In response to these issues, this study proposes an innovative prediction framework. The main contributions are as follows:

(1) Construction of a Multi-source Heterogeneous Dataset: This framework integrates multi-source heterogeneous data, encompassing HFMD case data and meteorological factors[5], while considering the lagged effects of meteorological factors on HFMD. Experiments with different sliding windows (7 days, 14 days, 21 days) are conducted to determine the optimal value.

(2) Introduction of a Long Time-Series Prediction Model Based on Attention Mechanism: This model design enhances the identification and capture of key features in the data through the attention mechanism. It leverages its robust time-series analysis capabilities to adapt to and reflect the dynamic characteristics of data across different time periods, thereby improving the prediction accuracy and robustness of HFMD incidence trends.

(3) Ablation Experiments: To validate the importance of the Attention module, CNN module, and BiLSTM module, ablation experiments are performed. It is found that the chained structure of the BiLSTM has the greatest impact on performance, followed by the CNN module, and finally the Attention module.

(4) Comparative Experiments: Comparative experiments are conducted with infectious disease dynamics models (e.g., SEIIeQR) and statistical models (e.g., ARIMA, SARIMA)[6] to demonstrate the superiority of the CNN-BiLSTM-Attention model. The results provide scientific evidence for relevant authorities in Qingdao to prevent HFMD outbreaks.

2. Method

2.1 Data Preprocessing

This study selected cases from 2010 to 2019 in Qingdao (provided by Qingdao CDC, with data desensitization) and meteorological data (obtained from China Meteorological Center). The z-score standardization method was used to process the data to ensure the stability, robustness, and prediction accuracy of the model. The mathematical formula is shown as:

$$y = \frac{(x-\mu)}{\sigma} \tag{1}$$

2.2 Evaluation Metrics

This study selects three evaluation metrics to measure model performance: Mean Squared Error (MSE), Mean Absolute Error (MAE), and R^2. MSE is a commonly used evaluation metric that calculates the squared difference between the actual value and the predicted value for each sample. The mathematical formula is as $MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{c}_i - \bar{c}_i)^2$. n represents the number , $\hat{c}_i$ represents the actual value, and $\bar{c}_i$ represents the predicted value from the model. A smaller value MSE indicates that the predicted values are closer to the actual values. Conversely, a larger value indicates a greater error between the predicted values and the actual results.

MAE is used to measure the average absolute deviation between the predicted values and the actual values. The mathematical formula is as $MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{c}_i - \bar{c}_i|$.which have the same meaning as MAS. Compared with MSE, MAE is more robust to extreme values because it uses absolute values instead of squared differences. However, MAE does not distinguish the direction or weight of errors, and cannot discern the tendency of errors in detail.

The value of R^2 ranges from 0 to 1, representing the proportion of the variance in the dependent variable that can be explained by the independent variables. A result closer to 1 indicates more accurate predictions, while a lower value suggests that the model needs improvement. The mathematical formula is as:

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{c}_i - \hat{c}_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(\hat{c}_i - \bar{c}_i)^2} . \tag{2}$$

2.3 CNN-BiLSTM-Attention model structure

The CNN-BiLSTM model [7]combines the strengths of Convolutional Neural Networks (CNNs) in feature extraction with the ability of Bidirectional Long Short-Term Memory Networks (BiLSTMs) to handle long-sequence contextual information, making it suitable for processing sequence data with strong spatiotemporal dependencies. However, for long sequence data, the pure CNN-BiLSTM model may suffer from performance degradation due to its inability to adequately focus on key features at each time step. To address this, an attention mechanism is introduced to dynamically adjust the weights of inputs at different time steps, enabling the model to automatically focus on the most important parts of the data and ignore secondary information. By integrating CNN's feature extraction, BiLSTM's sequence modeling, and attention's focal enhancement, this model significantly improves its processing capability and prediction accuracy for complex sequence data. The model structure is shown in Figure 1.

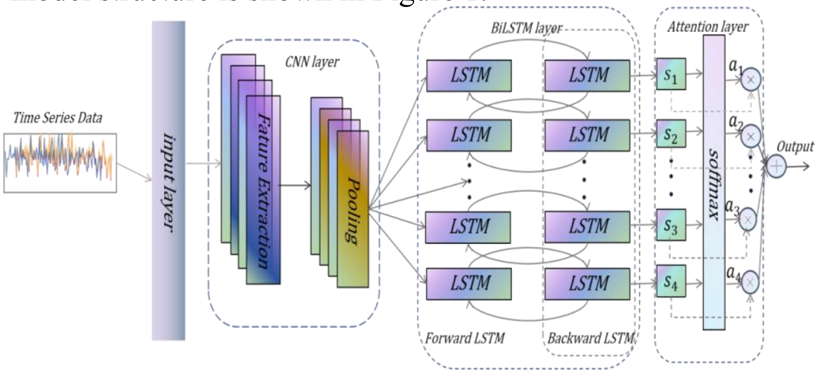

fig 1 CNN-BiLSTM-Attention model structure

The Attention layer[8], through precise weight calculations, accurately captures and enhances key factors, optimizing the model's predictive performance. Its computational mechanism is illustrated in Figure 2.

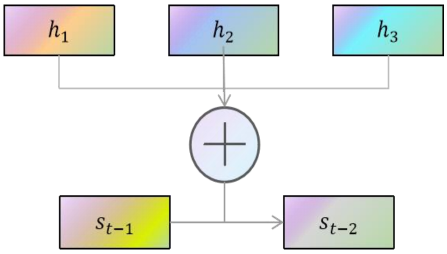

fig 2 The computational mechanism of Attention.

In the above figure, the hidden state values output by the BiLSTM layer are denoted as h_t, a_t represent the weight values of attention, which are used to calculate which factors and time periods have a stronger correlation with the predicted results. Which formula is shown in Formula 3:

$$a_t = \frac{\exp(m_t)}{\sum_{i=1}^{k} \exp(m_i)} \tag{3}$$

In the above formula, m_t is how h_t determine the size of the attention weight. The calculation formula is shown as $m_t = s_{t-1} * h_t$, s_{t-1} is the output value of the attention layer at t time step , the calculation formula is shown as $s_t = \sum_{j=1}^{t} \exp a_j * h_j$.

The activation function used in this layer is Sigmoid, and a fully connected layer is chosen as the output layer of this layer. The mathematical expression $data_n$ for the final output result obtained after calculation is shown in Formula 4:

$$data_n = sigmoid(w_{st} + b) \tag{4}$$

2.4 Experimental Procedure

Based on historical time series datasets data of Hand, Foot, and Mouth Disease (HFMD) and meteorological data, the CNN-BiLSTM-Attention model is used to predict HFMD data at future time points $data_n$. The training process is shown in Figure 3:

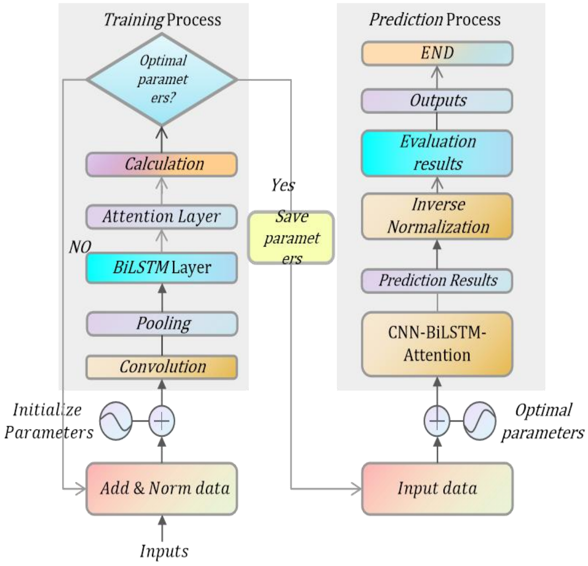

fig 3 CNN-BiLSTM-Attention Model Training Process

The training steps are as follows:

Preprocess data and initialize model: Standardize input time series data and initialize CNN-BiLSTM-Attention model parameters.

(2) Feature extraction and transformation: Use CNN for feature extraction, capture local features, reduce dimensions, reshape features, and convert to a sequence format suitable for BiLSTM.

(3) Sequence modeling and attention: Model transformed features with BiLSTM, capture contextual information, and introduce Attention to calculate contribution of different time steps and features, enhancing focus on important information.

(4) Training, optimization, and evaluation: Define a loss function, use optimization methods to adjust model parameters, minimize loss, and improve prediction accuracy.

2.4.1 Model Parameter Selection

Based on the model structure and training process, the research implemented code writing using Python and set the initial parameters for each layer. Through multiple experiments and tests on the training set[9], the optimal parameter combination was determined based on the evaluation of the loss function, as shown in Table 1:

Table 1 Parameter Settings for the CNN-BiLSTM-Attention Model

Hyperparameters	Search Range	Optimal Value
Hidden_size	[8,16,32,64,128,256,512]	128
Learning_rate	[0.001,0.005,0.01,0.05,0.1]	0.005
Dropout	[0.1,0.2,0.3,0.4,0.5]	0.2
Num. of Kernels	[4,8,16,32,64,128]	16
Pooling Depth	[2,3,4,5]	3
Batch_size	[60,80,100,120,160,200]	120
Attention	[RELU,Sigmoid,Tanh]	Sigmoid
Batch_size	[60,80,100,120,160,200]	120
Time_step	[7,14,21]	14

2.5 CNN-BiLSTM-Attention Model Prediction of HDMF

The data in this paper is divided into a training set (2010-2017), a testing set (2018), and a validation set (2019). After determining the parameters, the effectiveness is verified on the testing set, and the model is optimized by taking the average loss value through multiple training sessions on the validation set. In response to the characteristics of the HFMD epidemic in Qingdao, the data is further subdivided into the pre-outbreak period (May), the peak period (June), and the later period, taking into account the lag period and incubation period of meteorological factors[10]. By employing a sliding window mechanism, predictions are made for the spread of HFMD over the next 7 days, 14 days, and 21 days, respectively[11]. The experimental results are shown in Figure 4.

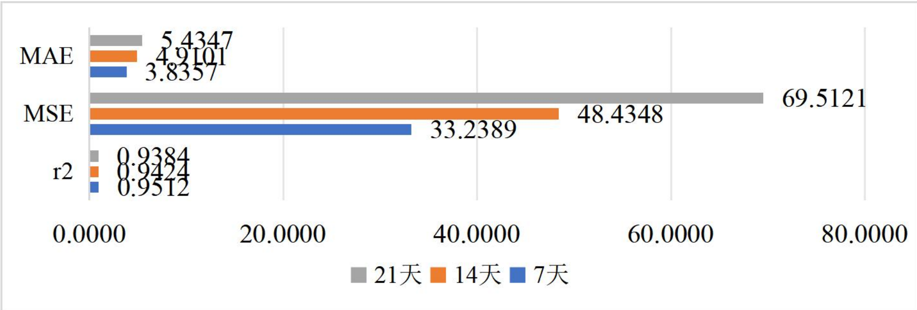

Fig 4 Experimental Results of CNN-BiLSTM-Attention Model in Different Time Lengths

The experimental results indicate that as the prediction time length increases, the effectiveness decreases. This may be attributed to two aspects: firstly, the highly dynamic nature of the real-world environment, where non-model learning factors such as policy interventions have a significant impact; secondly, prediction errors accumulate over time, leading to an enlargement of subsequent prediction deviations. In consideration of the lagged influence of meteorological factors on HFMD and in conjunction with the experimental results presented in Figure 4, this paper adopts a 7-day sliding window to predict the trend of hand, foot, and mouth disease in 2019. By applying moving average smoothing to both the actual and predicted values, the annual prediction situation is visually presented, as shown in Figure 5.

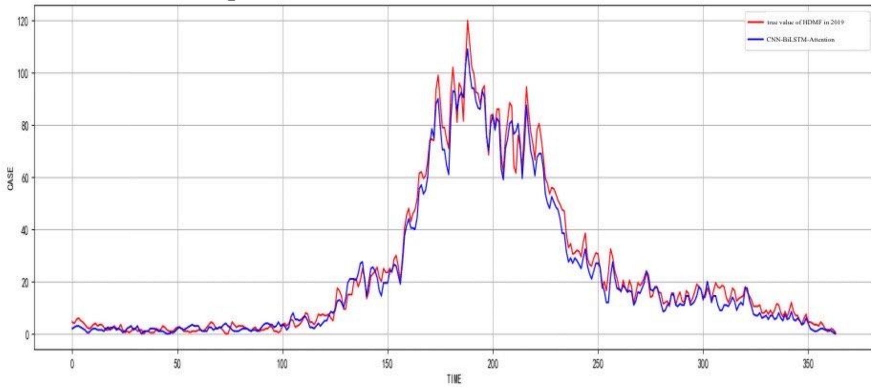

Fig 5 The prediction effectiveness of CNN-BiLSTM-Attention model on hand, foot, and mouth disease (HFMD) in Qingdao in 2019.

The results indicate that during the first 100 days, data sparsity due to scarce cases limited the prediction accuracy of the model. However, as data accumulated, the model significantly improved its predictive ability after 100 days, especially demonstrating high accuracy in predicting the trend and inflection point during the peak period of 2019, although there were still errors in predicting the specific number of peak cases (which may be related to various factors, including but not limited to the complexity of disease transmission mechanisms, the impact of seasonal variations on disease prevalence, and other potential influencing factors not included in the model, such as population migration and changes in vaccination coverage[12]). This highlights the importance of data scale for model performance. Therefore, future research should focus on further refining the model structure.

3. Comparative Experiment

To verify the superiority of the CNN-LSTM-Attention model in predicting hand, foot, and mouth disease (HFMD) in Qingdao, this study compares it with classical statistical models (ARIMA, SARIMA), the SEIIeQR infectious disease model, and the deep learning Informer model. Since Qingdao is part of Shandong province, the SEIIeQR model is representative, and the Informer model[13] has shown excellent performance in time series prediction, both were included in the comparison.

3.1 Experimental Analysis

In this study, the same historical case dataset was used to train and test the ARIMA, SARIMA, SEIIeQR, and CNN-BiLSTM-Attention models. Since statistical models and the SEIIeQR model do not support multi-variable input, meteorological data was not included. Figure 6 presents a comparison of the prediction results for the peak period of hand, foot, and mouth disease (HFMD) in Qingdao using various models:

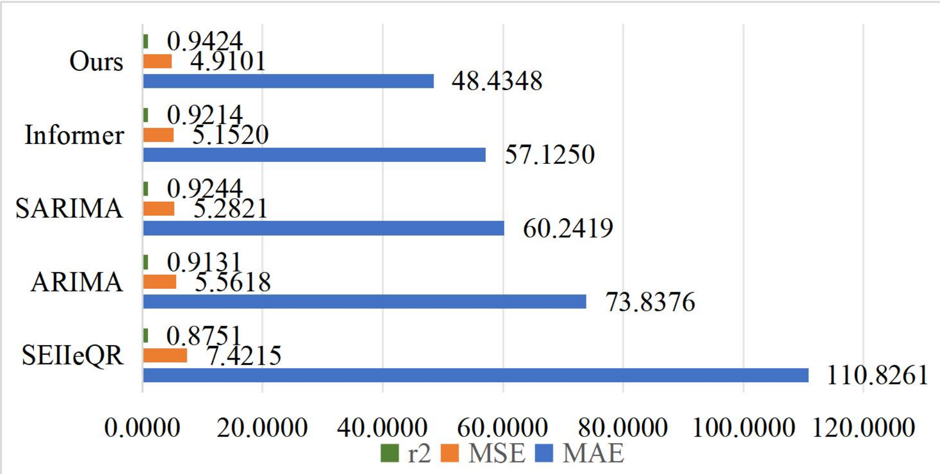

Fig 6 Experimental results of HFMD prediction by various models in the comparative experiment

3.1.1 Quantitative Analysis

From Figure 6, it can be observed that in terms of the MAE metric, the CNN-BiLSTM-Attention model showed a reduction of 17.94%, 24.38%, 52.45%, and 128.46% compared to the Informer, SARIMA, ARIMA models, and the SEIIeQR model, respectively. In terms of the MSE metric, the reductions were 11.03%, 7.58%, 4.93%, and 51.15%, respectively. For the specific metric not mentioned, the model used in this study showed improvements of 2.28%, 1.95%, 3.21%, and 7.69% compared to the others.

In summary, the deep learning model CNN-BiLSTM-Attention used in this study demonstrated better performance in accuracy compared to the above-mentioned prediction methods. following sections analysis the specific reasons for its superiority.

3.1.2 Qualitative Analysis

According to the classification of prediction methods, the infectious disease dynamic model SEIIeQR performed the worst, while deep learning models performed the best. The reasons are as follows:

(1) Infectious disease dynamic models are sensitive to parameters, some of which are difficult to obtain accurately due to recent changes like HFMD vaccination. Additionally, they do not consider meteorological factors, reducing prediction accuracy. However, the SEIIeQR model provides insights into HFMD dynamics and can guide targeted prevention measures.

(2) Compared to the ARIMA model, the SARIMA model performs better due to its incorporation of seasonal factors, aligning with Zhang et al.'s research[14].

(3) The deep learning model used in this study outperforms the Informer model, especially with limited data. Deep learning models excel because they combine meteorological factors and have strong adaptive capabilities, consistent with Chen et al.'s research[15] and highlighting the importance of incorporating meteorological factors for improved prediction accuracy.

Figure 7 is a comparison of the models in the experiment for predicting hand, foot, and mouth disease in Qingdao throughout 2019, which more intuitively demonstrates the aforementioned differences.

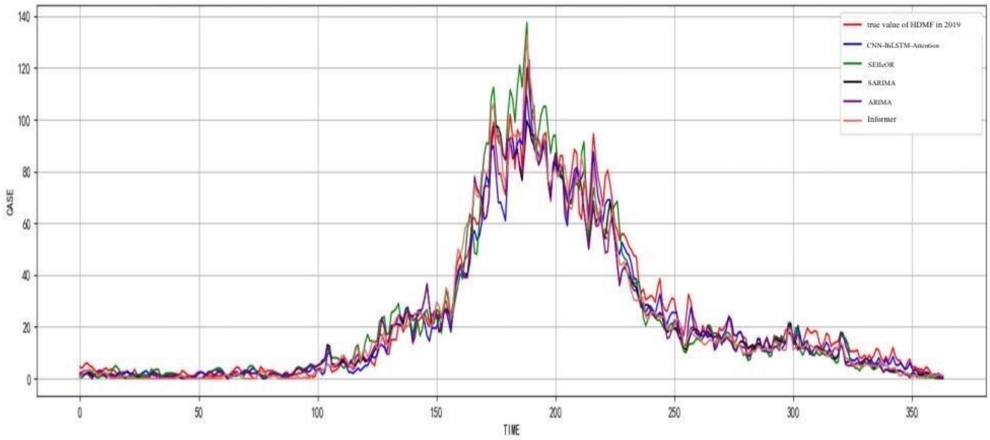

Fig 7 Comparison of the prediction effects of various models on hand, foot, and mouth disease (HFMD) in Qingdao in the experimental study.

4. Ablation Experiment

4.1 Experimental Analysis

The ablation experiment used the same data as the main model to predict the peak periods of hand, foot, and mouth disease (HFMD) for the pruned network model. The average loss was calculated over ten training sessions with the validation set, providing a visual demonstration of the impact of each module on the model's performance. The experimental results are shown in Figure 8.

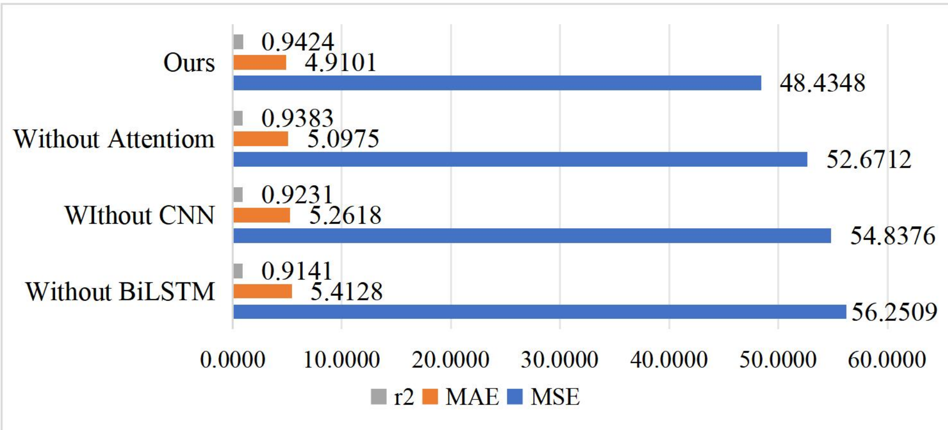

Fig 8 Comparison of the results of various models in predicting hand, foot, and mouth disease (HFMD) in the ablation experiment.

(1) Attention Module: A comparison was made between the CNN-BiLSTM model and the model in this study. When the Attention layer was removed, the Mean Absolute Error (MAE) and Mean Squared Error (MSE) decreased by 3.82% and 8.75% respectively, while another metric increased by 0.44%. This indicates that the lack of the attention mechanism module leads to a decline in prediction performance, verifying that the attention mechanism, which assigns weights to different factors at different time intervals, can enhance the model's prediction performance[16].

(2) CNN Module: A comparison was conducted between the model in this study and the BiLSTM-Attention model. When the CNN module was removed, the MAE and MSE decreased by 7.16% and 13.2% respectively, while R^2 increased by 2.09%. This suggests that the absence of the CNN module results in a decrease in prediction performance, confirming that the CNN module can extract local features from the data to some extent, thereby improving the model's prediction performance.

(3) BiLSTM Chain Structure: A comparison was made between the model in this study and the CNN-LSTM-Attention model. When the chain structure in the prediction module was removed, the MAE and MSE decreased by 10.24% and 16.13% respectively, while R^2 increased by 3.1%. This demonstrates that the BiLSTM module enhances model accuracy by effectively integrating contextual information, highlighting the importance of the chain structure.

Using the model parameters mentioned above to predict hand, foot, and mouth disease (HFMD) in Qingdao in 2019, the experimental results are shown in Figure 9. The predicted values of the model in this study are the closest to the real values, indicating that the incorporation of the Attention, CNN, and BiLSTM modules significantly improves the prediction performance of the original LSTM model. The integration of these modules can effectively enhance the prediction accuracy of the model, proving the effectiveness of the CNN-LSTM-Attention model.

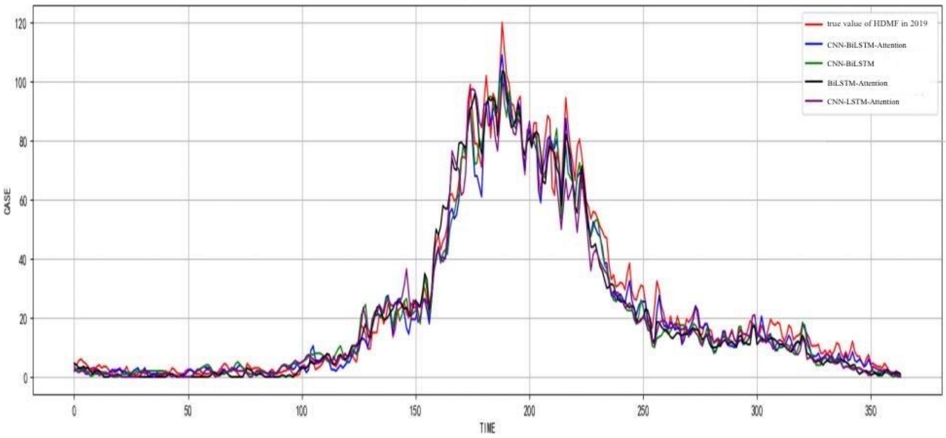

Fig 9 Prediction Effectiveness of Various Models on Hand, Foot, and Mouth Disease in the Ablation Experiment

5. Summary

This study focuses on predicting the hand, foot, and mouth disease (HFMD) in Qingdao using the CNN-BiLSTM-Attention model, with a particular emphasis on forecasting the peak periods of daily HFMD incidence in Qingdao from May to July 2019. The study describes the structure of the model and its training process, and evaluates the prediction performance using several evaluation metrics. To understand the impact of different components of the model's architecture on the prediction results, an ablation experiment was conducted for further validation. The experimental results demonstrate that the CNN module, Attention module, and bidirectional LSTM module used in the proposed model can effectively enhance the model's performance, proving the importance of each module. Finally, to showcase the superiority of the proposed method, comparisons were made with statistical methods (ARIMA, SARIMA), infectious disease dynamic models (SEIIeQR), and the latest deep learning model (Informer), demonstrating the superior predictive accuracy of the model presented in this paper. Additionally, through comparative analysis, the experiments indicate that incorporating temperature and humidity data from meteorological factors and their lag periods (7 days) can effectively improve the model's accuracy.

Referencing

[1] L.N. A ,M.M.A. J ,H. O .Characterization of enteroviruses from hand, foot and mouth disease cases in the Philippines, 2012 – 2017[J].International Journal of Infectious Diseases,2020,101(S1):501-501.

[2] Liu Jianhua. Study on the Prediction of Hand-Foot-and-Mouth Disease Epidemic Trend and Simulation of Transmission Dynamics [D]. Huazhong University of Science and Technology, 2017.

[3] Suraj V ,Abdur M R ,Usanut S , et al.Digital Diagnosis of Hand, Foot, and Mouth Disease Using Hybrid Deep Neural Networks[J].IEEE ACCESS,2021,9143481-143494.

[4] Yiran W ,Ping S ,Jiangchen L , et al.A hybrid model for hand-foot-mouth disease prediction based on ARIMA-EEMD-LSTM[J].BMC Infectious Diseases,2023,23(1):879-879.

[5] Y. Lai et al., "The early warning model of HFMD which is implemented by the multivariable deep learning neural network," 2023 IEEE 6th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2023, pp. 542-548, doi: 10.1109/ITNEC56291.2023.10082278.

[6] Luo Dongmei, Xu Jiabo, Feng Xing, et al. Prediction and Control Analysis of Hand-Foot-and-Mouth Disease Epidemic in Xinjiang Based on Infectious Disease Dynamics Model [J]. Mathematics in Practice and Theory, 2018, 48(22): 154-160.

[7] Mayur W ,Rao S C A ,Ajith A .CBMAFM: CNN-BiLSTM Multi-Attention Fusion Mechanism for sentiment classification[J].Multimedia Tools and Applications,2023,83(17):51755-51786.

[8] Tao Junwen, Zhang Tao, Zhuang Xuefei, et al. Comparison of Prediction Effects of Dynamic Bayesian Network Model and SARIMA Model on Hand-Foot-and-Mouth Disease [J]. Modern Preventive Medicine, 2020, 47(21): 3851-3854+4010.

[9] Gajowniczek K ,Liang Y ,Friedman T , et al.Semantic and Generalized Entropy Loss Functions for Semi-Supervised Deep Learning[J].Entropy,2020,22(3):334-334.

[10] Li Jun, Gu Minhua, Qian Cheng, et al. Exploring the Association Between the Incidence of Hand-Foot-and-Mouth Disease and Meteorological Factors in Jiangyin City Based on the Distributed Lag Non-linear Model (DLNM) [J]. Chinese Journal of Disease Control & Prevention, 2019, 23(04): 412-416. DOI: 10.16462/j.cnki.zhjbkz.2019.04.009.

[11] ZhiMin H ,HuHu W ,YanJuan W , et al.Spatiotemporal analysis of hand, foot and mouth disease data using time-lag geographically-weighted regression.[J].Geospatial health,2020,15(2):337-347.

[12] Railey F A ,Lembo T ,Palmer H G , et al.Spatial and temporal risk as drivers for adoption of foot and mouth disease vaccination[J].Vaccine,2018,36(33):5077-5083.

[13] Chang Wanjie, Liu Linlin, Cao Yu, et al. Research on Virus Transmission Prediction Based on Informer Algorithm [J]. Journal of Liaoning Shihua University, 2024, 44(01): 80-88.

[14] Y. Wang, C. Xu, S Zhang et al., "Development and evaluation of a deep learning approach for modeling seasonality and trends in hand-foot-mouth disease incidence in mainland China", Sci Rep, vol. 9, pp. 8046, 2019, [online] Available: https://doi.org/10.1038/s41598-019-44469-9.

[15] S. Liu, J. Chen, J Wang et al., "Predicting the outbreak of hand foot and mouth disease in Nanjing China: a time-series model based on weather variability", Int J Biometeorol, vol. 62, pp. 565-574, 2018, [online] Available: https://doi.org/10.1007/s00484-017-1465-3.

[16] Huangfu Xiaoying, Qian Huimin, Huang Min. A Review of Deep Neural Networks Incorporating Attention Mechanism [J]. Computer and Modernization, 2023, (02): 40-49+57.