

RESEARCH ARTICLE

Wavelet-Based ResNet: A Deep-Learning Model for Prediction of Significant Wave Height

XIANGJUN YU¹, YARONG LIU², ZHIMING SUN², AND PAN QIN^{ID}²¹Department of Military Oceanography and Hydrography, Dalian Naval Academy, Dalian 116018, China²School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China

Corresponding author: Pan Qin (qp112cn@dlut.edu.cn)

ABSTRACT Predicting significant wave height (SWH) is significant for coastal energy evaluation and utilization, port construction, and shipping planning. It has been reported that SWH is difficult to forecast for the complex marine conditions and chaos in nature. Current methods either require reliable prior information or reach the upper limit of prediction accuracy. To this end, this paper proposes a wavelet-based residual network to predict SWH with high accuracy. First, the time-series data of wave-related factors collected by the ocean buoy station is decomposed using the wavelet transformation. Then, the transformation results are used as the inputs to train the residual neural network. Finally, the data obtained from the NOAA's National Data Buoy Center is used to prove the outperformed prediction accuracy of the proposed method. The analysis results suggested that wavelet transformation can improve the prediction performance of the neural network, and the proposed model achieves better performance compared with several other deep neural network schemes.

INDEX TERMS Convolution network, data mining, ocean wave time series, ResNet, wavelet decomposition.

I. INTRODUCTION

Marine meteorological forecasting, especially wave elements forecasting, significantly impacts human marine activity. Accurate estimation of wave parameters not only serves as a major reference for coastal energy evaluation and utilization [1], [2] but also provides essential guidelines for port construction and shipping planning [3]. The significant parameters for wave characterization include significant wave height (SWH), wave period, and wave speed. Among them, SWH is one of the most relevant and necessary parameters to evaluate the wave energy source and ocean meteorological condition, and thus, the estimation of SWH is a core question of wave characterization [4].

The measurements of SWH usually consist of using remote sensing methods [5], [6] or using wave-buoy-type in-situ sensors [7]. The remote sensing methods often require additional data reprocessing and have a significant deviation, which still needs to be corrected using the buoy station data. Currently, the prediction of SWH mainly relies on the wave buoy observations data, which provide time series of wave

state information in fixed points [8]. Thus, it is necessary to use large and reliable buoy datasets to estimate SWH accurately. Currently, the largest marine meteorological database is the National Oceanic and Atmospheric Administration (NOAA), which consists of a great deal of real-time and long-term marine buoys observation data, and are the significant sets used for the estimation and evaluation of ocean wave SWH [9].

The SWH estimation problem has been considered by using various methods, including numerical wave models [10], [11], time series models [12], and artificial neural network (ANN) [13]. The numerical models, like WAVEWATCH III of wide ocean, Simulating Waves Nearshore (SWAN) of offshore, are based on the approximation and simulation of wave-spectrum. Through numerical calculation, SWH and other parameters can be derived from the wave action equation to describe future wave states [14]. This approach can achieve comparatively ideal results given sufficient external prior knowledge and has been the main approach for wave forecasting over a long period in the past. Rogers [15] used SWAN to predict the wave SWH trend near the Southern California Bridge and compared the results with measured data to investigate the effect of island

The associate editor coordinating the review of this manuscript and approving it for publication was Zhipeng Cai ^{ID}.

topographical resolution on prediction. Zheng [16] applied WAVE-WATCH III framework to simulate the seasonal and long-term trends of SWH in the South China Sea and further analyzed the statistical data distribution of SWH in different time and areas. The shortcoming of numerical models is that they require reliable prior information such as geography and meteorology of the predicted location. Thus, slight deviations of prior information will be gradually amplified and greatly impact the accuracy of the final prediction. Therefore, data-driven methods like time series models and machine learning algorithms have recently acquired great attention for SWH estimation to improve prediction performance by digging into the inherent characteristics of historical data and reducing the reliance on prior knowledge. Experimental results have revealed a better performance of machine learning methods over the statistical predictive model in SWH prediction [17].

In the classical time series and machine learning approaches for SWH, empirical mode decomposition (EMD) and wavelet transform (WT) have been used to extract features from wave time-series data [18], [19], [20] and the prediction is consequently conducted through models like autoregressive (AR) moving average, support vector regression (SVR), and symbiotic organism search (SOS) [21]. For example, Salcedo-Sanz [4] applied SVR to SWH prediction using the shadowing effect of radar images and achieved excellent performance. Meanwhile, Duan [22] performed EMD ahead of SVR prediction for feature extraction and reached a higher accuracy on NOAA buoys datasets. However, with the rapidly increase of marine data dimensions, the complex relationships between unstructured or semistructured data limits the upper bound of the prediction accuracy of traditional machine learning analysis approaches. These shortcomings have been gradually addressed by developing neural network based models. With the increasing computing speed of graphics computing devices, ANN has been widely applied for wave forecasting.

EMD and wavelet have also been combined with ANNs [23]. Shahabi [24] proposed a GMDH network to predict the SWH of north Atlantic coast based on the buoys data and achieved better results at the 6-step to 12-step period prediction than time-series and machine learning models. Pushpam [25] used the long-short-term memory (LSTM) network to reconstruct and predict the wave height of Bay of Bengal, which achieved better performance than the traditional forward ANNs and recursive neural networks (RNN). Kaloop [26] proposed a wavelet-particle-swarm optimization extreme learning machine (ELM) to estimate the ocean wave height, and the experiment results on buoys data of the US south-east coast outperformed SOS, LSTM, and SVR. Wang [27] applied deep neural network for the calibration of HY2B SWH by using input from parameters provided by the altimeter and greatly improved the performance of HY2B. The limitations of the existing ANN based models are that RNN and its variants are of high computation cost for its sequential calculation, and are thus lower in efficiency. On the other hand, ELM is limited by the single hidden

layer structure and thus runs into the problem of insufficient nonlinear fitting ability. Residual network (ResNet) is a powerful framework that can learn a wide range of complex relationships from data, its efficiency and robustness has been convinced in a variety of applications [28]. The effectiveness of this model suggests a potential application in marine element forecast.

In this research, a wavelet-based residual network is proposed to construct an effective deep-learning method for prediction SWH. First, the time-series data of ocean wave data is wavelet-transformed for feature extraction and noise elimination after min-max normalization. Then, the processed data is used as the input of ResNet for prediction. The output of ResNet is transformed into one-dimensional sequence and then flattened into a one-dimensional vector through a two-layer linear block as the final prediction of ocean SWH.

The rest of the paper is organized as follows: In the second part, the overall framework and detailed settings of the proposed model are illustrated. Performance validation and results analysis are given in the third part. The conclusions are given in the last part.

II. METHODOLOGY

Figure 1 illustrates the overall framework, in which the proposed model takes the buoys data as input and predicts the SWH of a certain time step period.

A. PREPROCESSING

The features of buoy data include SWH, gust wind speed (GS), average wind speed (WS), dominant wave period (DPD), average wave period (APD), and air temperature (AT). The features are preprocessed to keep the network weights and biases staying to avoid extremely large parameters of the trained networks. First, all the features are normalized into the interval $[0, 1]$ by

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (1)$$

where x denotes the feature measurement and x' is the corresponding normalized feature value. It is worth noting that the minimum and maximum values used for normalization are selected from the training set.

WT transformation is conducted to the normalized data to further denoise and identify the short-term change and long-term trend. Using the time and frequency characteristics of wavelet functions, WT can adaptively sense the frequency changes of signal in the time metrics. For this reason, WT is suitable for the frequency contents analysis of signals with no redundant components in the time domain [29]. By using WT, the data are decomposed as the following

$$x(t) = \sum_k c_0[k] \phi_{0,k}(t) + \sum_{j=0} \sum_k d_j[k] \psi_{j,k}(t), \quad (2)$$

where $\psi_{j,k}(t)$ are the wavelet signals, $\phi_{0,k}(t)$ are the scale signals of wavelet basis function. Figure 2 shows the preprocessing results.

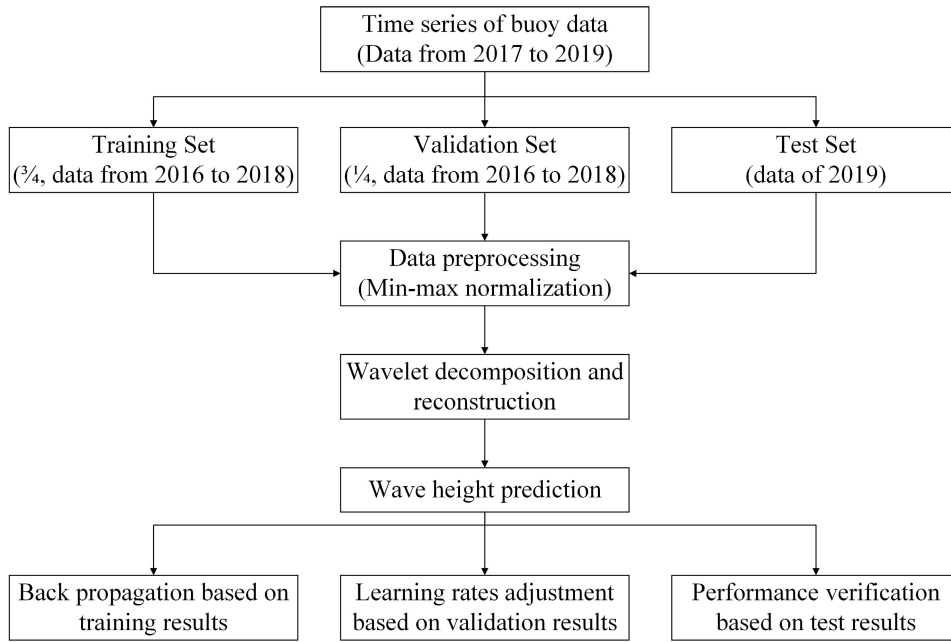


FIGURE 1. Framework of proposed method.

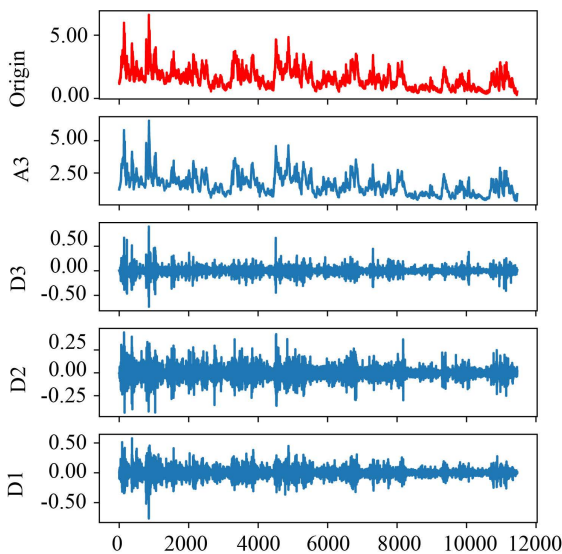


FIGURE 2. Origin and three-level wavelet decomposition of SWH sequence in June 2019.

B. ARCHITECTURE OF NETWORK

The major characteristics of a convolution network (CNN) are weight sharing, local connection, and down sampling. Compared with RNN and its variant LSTM, CNN can better utilize the multi-core parallel computing performance of computing devices and hereby has a faster computation speed [30]. Let ω be the convolution kernel, x be the input, b be the bias, then the convolution operation of CNN can be expressed as

$$h = f(w * x + b) \tag{3}$$

where $*$ denotes the convolution operator and $f(\cdot)$ denotes the activation function.

The utilization of CNN in time series sequence prediction usually takes two ways: one-dimensional convolution and two-dimensional convolution [31], [32]. In this article, the wavelet decomposition results of different features with the same order are subjected to different channels of the same input group. The one-dimensional sub-sequences of WT are split equally and reshaped into two-dimensional matrix as the input of certain dimension of input channels.

As depicted in Figure 3, the overall framework of the proposed model contains eight forward layers and one residual shortcut connection between the output of layer 1 and layer 5 to avoid accuracy degradation. The first six layers are convolutional and the remaining two are linear. The convolutional layers have 3×3 filters and follow two rules: (1) each layer is followed by a zero entries padding of size 1 to keep feature map size unchanged; and (ii) the stride of all filters is set as 1. For residual connection, a 1×1 convolution with 80 channels is used for dimension matching. Specifically, the first convolutional layer filters the $8 \times 9 \times 24$ input with 80 kernels, and the output is taken as input of the second layer which is further filtered by 160 kernels. The third convolutional layer has 320 kernels, the fourth has 160, and the fifth has 80. The output of the fifth convolutional layer is filtered with 24 kernels of size 3×3 , which keep the sizes of input and output layers. The output of convolutional block is flattened into a one-dimensional vector and is fed to a linear network with two layers (Figure 4), which produces a distribution of the SWH.

C. TRAINING STRATEGY

In the training stage, dropout is used in the last two linear layers to zero the outputs of hidden neurons with a probability of 0.3. The dropped neurons do not participate in forward

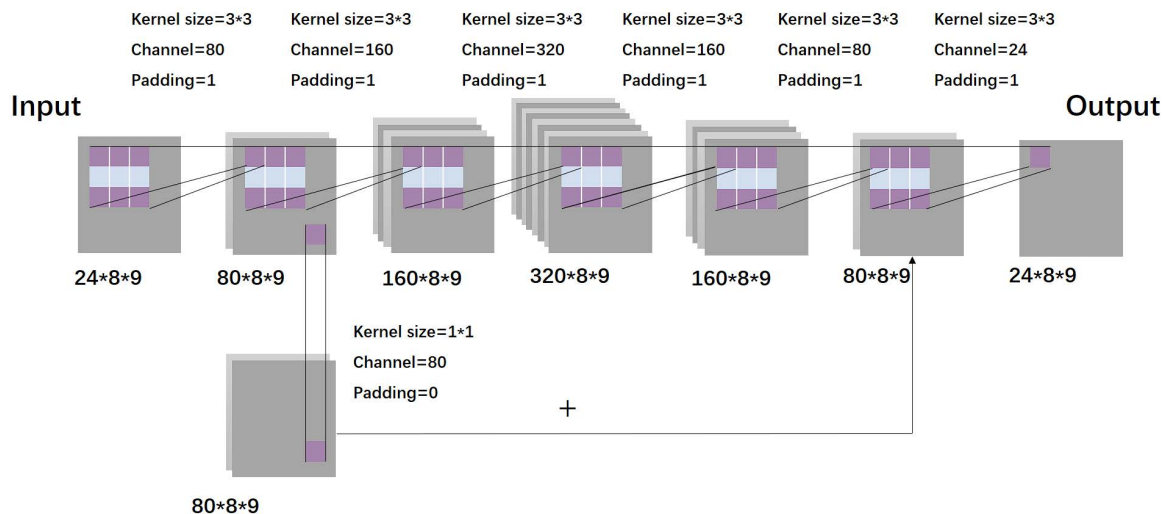


FIGURE 3. Structure of convolutional ResNet.

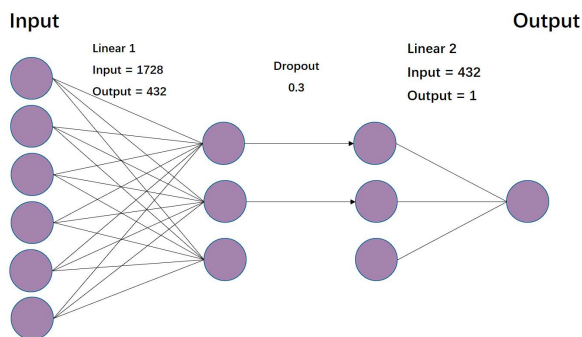


FIGURE 4. Structure of linear network.

pass and back-propagation in the training process, but in test time, all neurons are used but multiply weights by 0.3. The stochastic gradient descent (SGD) method was used to optimize the parameters. The descent of learning rate adopted the following strategy: (i) for each 5 rounds of unreduced training loss, reduce the learning rate by 50%, and (ii) if the loss value of the model on the validation set does not decrease after 17 training consecutive rounds, then stop the training process. The mean squared error (MSE) and the correlation coefficient R are used to evaluate the performance of the deep learning model.

III. EXPERIMENTAL RESULTS

A. DATASET

The proposed model is validated by using a buoy dataset from the NOAA database (<https://www.ndbc.noaa.gov>) [3]. The buoy station 46087 is located at 48.49° north latitude and 124.73° west longitude and the sea depth of this area is 260 meters. The raw buoy data obtain the features SWH, gust wind speed, average wind speed, dominant wave period, average wave period, and air temperature. The buoy data from January 2016 to December 2018 are used for training the

TABLE 1. Statistical properties of buoy data.

Variable	Mean	Variance	Maximum	Minimum
SWH	1.85	0.87	8.94	0.31
GS	6.70	3.97	28.50	1.60
WS	5.31	3.29	21.60	0.00
DPD	10.39	2.59	23.53	2.94
APD	7.22	1.39	14.27	3.27
AT	10.09	5.13	22.00	-0.20

network and the data from January 2019 to August 2019 are used for testing. The statistical properties of all features are presented in Table 1.

The buoy data of 72 last hours from time t are used as inputs. SWH values of 0.5 hours, 1.5 hours, 3 hours, 6 hours, and 12 hours ahead are used as the outputs.

B. PREDICTION PERFORMANCE

To evaluate the performance of proposed method on SWH prediction, the model is compared with several standalone models including CNN [3] and LSTM [33] with the same settings for learning rate, optimization algorithm, and batch size. Moreover, two hybrid models Wavelet-CNN and Wavelet-LSTM that combine the series decomposition technique were compared. Among them, CNN adopts the same convolutional and linear block with the proposed model. LSTM with and without wavelet impact are both set as the same two-layer bidirectional RNN, only differing in the parameter tuning process owing to different hidden and cell layers dimensions.

Experimental results of the proposed model and other methods on SWH prediction are given in Table 2. Comparison results suggest that wavelet decomposition of the origin time series sequence can better extract the inherent features of the ocean wave height variation and consequently improve the prediction performance of model in the training stage. The CNN and LSTM models with wavelet generally achieve better performance than those without wavelet impact.

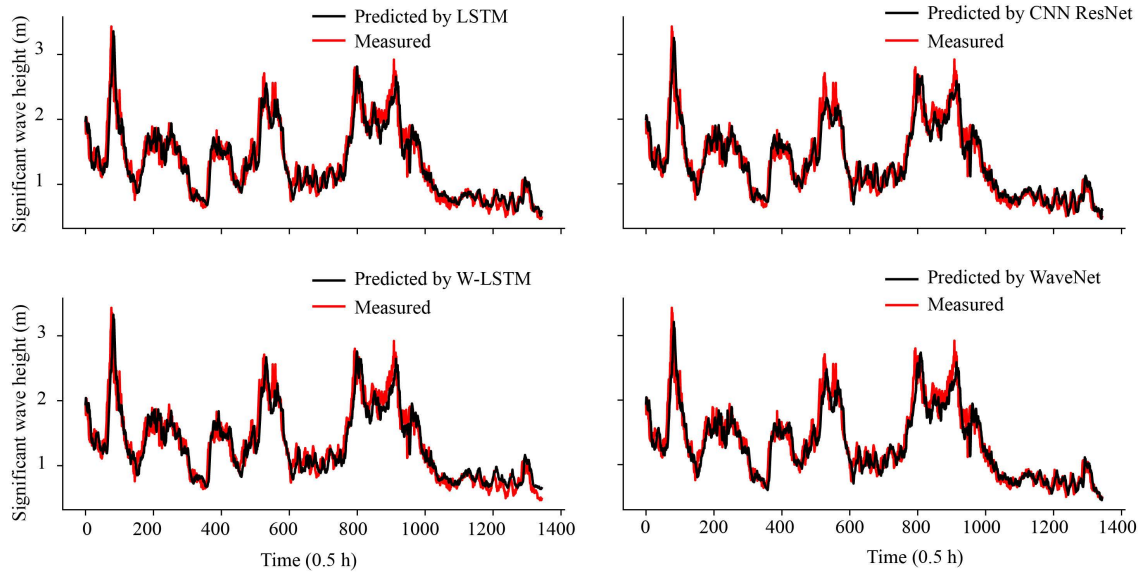


FIGURE 5. 6-step ahead prediction results of significant wave height of each model wave in June 2019.

TABLE 2. Prediction results of significant wave height of each model.

Model	1-step ahead		3-step ahead		6-step ahead		12-step ahead		24-step ahead	
	MSE	R	MSE	R	MSE	R	MSE	R	MSE	R
LSTM	0.022	0.98	0.04	0.97	0.07	0.93	0.15	0.86	0.24	0.76
Convolutional ResNet	0.020	0.98	0.04	0.97	0.07	0.94	0.13	0.88	0.27	0.74
Wavelet-LSTM	0.022	0.98	0.04	0.97	0.07	0.94	0.13	0.88	0.23	0.77
Wavelet-Convolutional ResNet	0.018	0.99	0.03	0.97	0.06	0.95	0.12	0.89	0.21	0.80

TABLE 3. Time consumption of CNN, LSTM, Wavelet-LSTM and Wave-CNN under different hours ahead SWH forecast.

Model	Decomposing time(s)	Training/Prediction time(s)				
		1 h	3 h	6 h	12 h	24 h
LSTM	/	1188.0/0.4	904.2/0.4	585.4/0.4	350.3/0.4	427.2/0.4
Convolutional ResNet	/	432.3/0.4	378.6/0.3	305.4/0.3	280.4/0.4	230.5/0.4
Wavelet-LSTM	0.8	1504.8/0.4	1106.5/0.4	1111.2/0.4	814.5/0.5	780.6/0.5
Wavelet-Convolutional ResNet	0.8	896.8/0.4	690.5/0.3	617.0/0.5	560.4/0.4	462.3/0.4

The results of multi-step-head predictions suggest that all four models can achieve good performance on short-term prediction, especially for 6 or fewer steps. As shown in Table 2, the values of MSE and R of the four models are all around 0.07 and 0.94. Meanwhile, Figure 5 also shows the good performance at 6-step-ahead prediction. Furthermore, Figure 5 also indicates that LSTM model with or without wavelet-based preprocess achieved better performance at early stage of prediction, while, the performance of both the CNN models are improved at later stage. However, for 12-step-ahead prediction, the wavelet decomposition improves the accuracy of CNN and LSTM, the values of MSE and R reach 0.12, 0.89 and 0.13, 0.88, respectively. Figure 6 shows that the performance of the CNN model without the wavelet-based preprocessing becomes worse at the late stage.

Figure 7 shows that the performance significantly degenerates for the 24-step-ahead prediction.

C. COMPUTING TIME

In this part, the training and prediction time of CNN, LSTM, Wavelet-LSTM and Wavelet-CNN are compared to evaluate the computational efforts of each model. In order to compare fairly, all experiments are conducted on a computer with 10GB RTX-3080 graphics card and 16GB memory.

The time consumption of Wavelet-LSTM and Wavelet-CNN consists of decomposing time of the original time series and training/prediction time of the models, while only training/prediction time for the single model CNN and LSTM. As shown in Table 3, the decomposing time of Wavelet-CNN and Wavelet-LSTM is both around 0.8 s, and

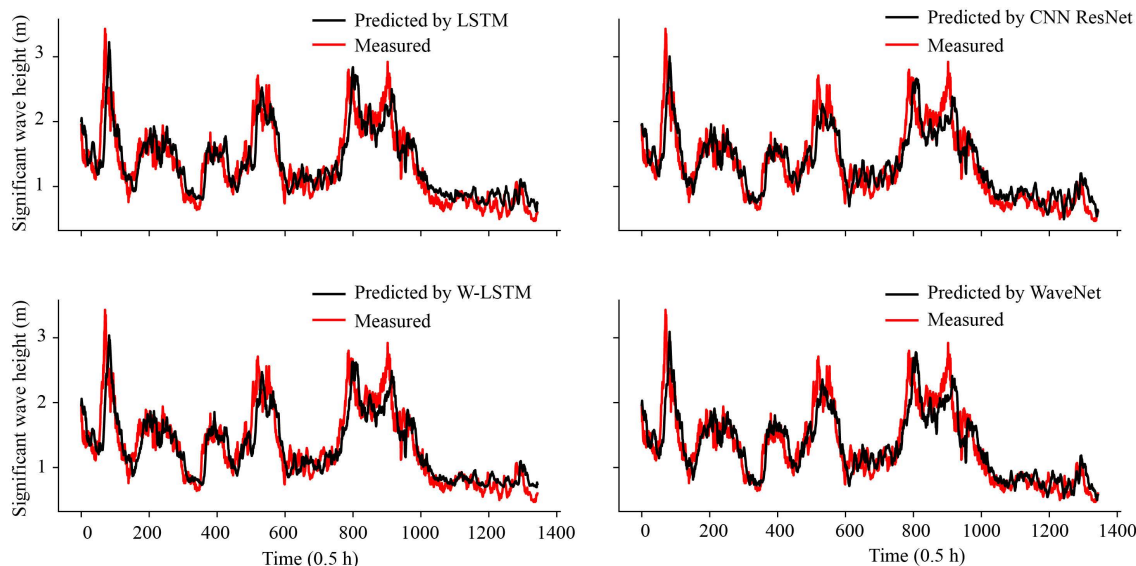


FIGURE 6. 12-step ahead prediction results of significant wave height of each model wave in June 2019.

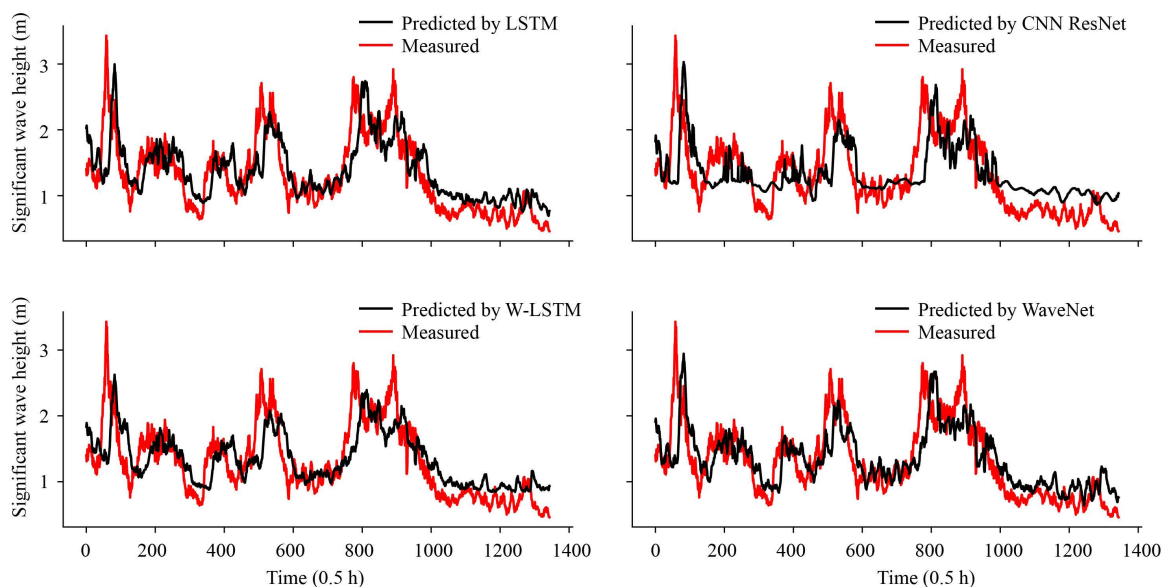


FIGURE 7. 24-step ahead prediction results of significant wave height of each model wave in June 2019.

the overall training time of hybrid models is generally longer than that of the single models. Specifically, Wavelet-LSTM has the longest training time, followed by Wavelet-CNN and LSTM, and the shortest of CNN. Moreover, the training time of four models all decreases with the increase of forecasting time. However, it can be seen that the prediction time of four models is basically the same, all about 0.4 s, which suggests that the efficiency of the four models is roughly the same in the actual prediction scenario.

IV. CONCLUSION

In this paper, a wavelet-based ResNet model is proposed for the prediction of SWH. This proposed model consists of

two parts, in which the wavelet decomposition extracts the learnable features from buoy data, and the ResNet provides SWH prediction in terms of the features decomposed by WT.

Overfitting is often an inevitable problem when training model for complex tasks like time series prediction due to the overly complicated model structure, especially when there is limited amount of training data. Several common techniques like data augmentation, regularization, dropout and early stopping can be adopted in order to avoid this situation [34], [35]. In this study, a combination of these training techniques is adopted to avoid model overfitting. Specifically, the datasets used for model training and testing are large enough, thereby ensuring a better coverage of the real data

distribution. In the training stage, dropout and adaptive learning rate techniques are adopted to reducing the complexity of model parameters. As a result, the proposed model achieves better performance than other approaches in the testing stage and has similar prediction accuracy as that in the training stages, indicating the model is not overfitting.

To validate the performance of the proposed model, comparative investigations among the LSTM, Wavelet-LSTM, Convolutional ResNet and Wavelet-ResNet models have been conducted. The experimental results indicate that the wavelet transformation can improve the prediction performance of the neural network. Through the wavelet basis function's time and frequency domain characteristics, the wavelet transformation can adaptively decompose the signals into trends with different scales and achieve effective features for the further deep-learning issue. Meanwhile, the wavelet-base ResNet can achieve better performance on long-term prediction. In the experiments of 24-step ahead estimation, the proposed model showed relatively slighter time delay and higher accuracy on the prediction of high wave heights, indicating that better performance was achieved. Nevertheless, the model still shows some limitations in long-term prediction, especially in the prediction of extreme SWH values, there is still a certain error gap between the prediction results of model and the measured data. Further work is needed in order to address this problem.

In summary, this paper verifies the practicability of wavelet transform and residual convolution neural network in SWH prediction. This not only provide a reference for the selection of signal decomposition methods in time series prediction like SWH forecasting, but also contribute a new method to SWH prediction. For the follow-up study, other similar time series decomposition or transformation methods and other neural network structures can be considered on the basis of this study, and their optimal combination can be explored to further improve the accuracy of model prediction results.

ACKNOWLEDGMENT

(Xiangjun Yu and Yarong Liu contributed equally to this work.)

REFERENCES

- [1] I. Lopez, J. Andreu, S. Ceballos, I. M. D. Alegria, and I. Kortabarria, "Review of wave energy technologies and the necessary power-equipment," *Renew. Sustain. Energy Rev.*, vol. 27, pp. 413–434, Nov. 2013.
- [2] F. Comola, T. L. Andersen, L. Martinelli, H. Burcharth, and P. Ruol, "Damage pattern and damage progression on breakwater roundheads under multidirectional waves," *Coastal Eng.*, vol. 83, pp. 24–35, Jan. 2014.
- [3] S. Yang, Z. Deng, X. Li, C. Zheng, L. Xi, J. Zhuang, Z. Zhang, and Z. Zhang, "A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast," *Renew. Energy*, vol. 173, pp. 531–543, Aug. 2021.
- [4] S. S. Sanz, J. C. N. Borge, L. C. Calvo, L. Cuadra, K. Hessner, and E. Alexandre, "Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface," *Ocean Eng.*, vol. 101, pp. 244–253, Jun. 2015.
- [5] P. Izquierdo, J. Borge, C. Soares, R. Gonzalez, and G. Rodriguez, "Comparison of wave spectra and scalar from nautical radar images buoy data," *J. Waterway Port Coastal Ocean Eng.-Asce*, vol. 131, no. 3, pp. 123–131, 2005.
- [6] J. Seemann, F. Ziemer, L.-C. Wu, M. Cysewski, and S. Flampouri, "The analysis of sea surface dynamics using a dopplerized X-band radar," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 2428–2430.
- [7] P. Abhigna, S. Jerritta, R. Srinivasan, and V. Rajendran, "Analysis of feed forward and recurrent neural networks in predicting the significant wave height at the moored buoys in bay of Bengal," in *Proc. Int. Conf. Commun. Signal Process. (ICCCSP)*, 2017, pp. 1856–1860.
- [8] L. Cornejo-Bueno, J. C. N. Borge, E. Alexandre, K. Hessner, and S. Salcedo-Sanz, "Accurate estimation of significant wave height with support vector regression algorithms and marine radar images," *Coastal Eng.*, vol. 114, pp. 233–243, Sep. 2016.
- [9] L. C. Bueno, J. C. N. Borge, P. G. Díaz, G. Rodríguez, and S. S. Sanz, "Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm—Extreme learning machine approach," *Renew. Energy*, vol. 97, pp. 380–389, Nov. 2016.
- [10] T. Yasuda, H. Ueyama, and H. Mase, "Empirical estimation of wave height over reef based on Boussinesq wave model," in *Proc. Int. Conf. Asian Pacific Coasts*, N. Viet, D. Xiping, and T. Tung, Eds. 2020, pp. 175–181.
- [11] H. Chun and K.-D. Suh, "Empirical formulas for estimating maximum wave height and period in numerical wave hindcasting model," *Ocean Eng.*, vol. 193, Dec. 2019, Art. no. 106608.
- [12] W.-Y. Duan, L.-M. Huang, Y. Han, and D.-T. Huang, "A hybrid EMD-AR model for nonlinear and non-stationary wave forecasting," *J. Zhejiang Univ.-Sci. A*, vol. 17, no. 2, pp. 115–129, 2016.
- [13] H. Kagemoto, "Forecasting a water-surface wave train with artificial intelligence—A case study," *Ocean Eng.*, vol. 207, Jul. 2020, Art. no. 107380.
- [14] H. Mase, K. Oki, T. S. Hedges, and H. J. Li, "Extended energy-balance-equation wave model for multidirectional random wave transformation," *Ocean Eng.*, vol. 32, no. 8, pp. 961–985, 2005.
- [15] W. E. Rogers, J. M. Kaihatu, L. Hsu, R. E. Jensen, J. D. Dykes, and K. T. Holland, "Forecasting and hindcasting waves with the swan model in the southern California bight," *Coastal Eng.*, vol. 54, no. 1, pp. 1–15, 2007.
- [16] K. Zheng, A. A. Osinowo, J. Sun, and W. Hu, "Long-term characterization of sea conditions in the east China sea using significant wave height and wind speed," *J. Ocean Univ. China*, vol. 17, no. 4, pp. 733–743, 2018.
- [17] A. Ali, A. Fathalla, A. Salah, M. Bekhit, and E. Eldesouky, "Marine data prediction: An evaluation of machine learning, deep learning, and statistical predictive models," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–13, Nov. 2021.
- [18] M. Rhif, A. B. Abbes, I. Farah, B. Martínez, and Y. Sang, "Wavelet transform application for/in non-stationary time-series analysis: A review," *Appl. Sci.*, vol. 9, no. 7, p. 1345, Mar. 2019.
- [19] J. Quilty and J. Adamowski, "Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework," *J. Hydrol.*, vol. 563, pp. 336–353, Aug. 2018.
- [20] R. Tür, "Maximum wave height hindcasting using ensemble linear-nonlinear models," *Theor. Appl. Climatol.*, vol. 141, nos. 3–4, pp. 1151–1163, Aug. 2020.
- [21] K. Zhao and J. Wang, "Significant wave height forecasting based on the hybrid EMD-SVM method," *Indian J. Geomarine Sci.*, vol. 48, no. 12, pp. 1957–1962, 2019.
- [22] W. Y. Duan, Y. Han, L. M. Huang, B. B. Zhao, and M. H. Wang, "A hybrid EMD-SVR model for the short-term prediction of significant wave height," *Ocean Eng.*, vol. 124, pp. 54–73, Sep. 2016.
- [23] P. C. Deka and R. Prahada, "Discrete wavelet neural network approach in significant wave height forecasting for multistep lead time," *Ocean Eng.*, vol. 43, pp. 32–42, Apr. 2012.
- [24] S. Shahabi, M.-J. Khanjani, and M.-R. Hessami, "Significant wave height forecasting using GMDH model," *Int. J. Comput. Appl.*, vol. 133, no. 16, pp. 13–16, Jan. 2016.
- [25] M. M. P. Pushpam and F. V. S. Enigo, "Forecasting significant wave height using RNN-LSTM models," in *Proc. Int. Conf. Intell. Comput. Control Syst.*, 2020, pp. 1141–1146.
- [26] M. R. Kaloop, D. Kumar, F. Zarzoura, B. Roy, and J. W. Hu, "A wavelet—Particle swarm optimization—Extreme learning machine hybrid modeling for significant wave height prediction," *Ocean Eng.*, vol. 213, Oct. 2020, Art. no. 107777.
- [27] J. Wang, L. Aouf, Y. Jia, and Y. Zhang, "Validation and calibration of significant wave height and wind speed retrievals from HY2B altimeter based on deep learning," *Remote Sens.*, vol. 12, no. 17, p. 2858, Sep. 2020.

- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.
- [29] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
- [30] M. Ibrahim, M. Torki, and M. ElNainay, "CNN based indoor localization using RSS time-series," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2018, p. 01044.
- [31] C. Lea, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks: A unified approach to action segmentation," in *Proc. Eur. Conf. Comput. Vis.*, vol. 9915, 2016, pp. 47–54.
- [32] M. U. Gudelek, S. A. Boluk, and A. M. Ozbayoglu, "A deep learning based stock trading model with 2-D CNN trend detection," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2017, pp. 1–8.
- [33] S. Fan, N. Xiao, and S. Dong, "A novel model to predict significant wave height based on long short-term memory network," *Ocean Eng.*, vol. 205, Jun. 2020, Art. no. 107298.
- [34] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *Comput. Sci.*, vol. 3, no. 4, pp. 212–223, 2012.
- [35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2012.



YARONG LIU received the M.S. degree from the Dalian University of Technology, Dalian, China, in 2021, where he is currently pursuing the Ph.D. degree with the Faculty of Electronic Information and Electrical Engineering.

His main research interests include deep learning, bioinformatics, and statistical modeling.



ZHIMING SUN received the B.S. and M.S. degrees from the Dalian University of Technology, Dalian, China, in 2019 and 2022, respectively.

His main research interests include deep learning and statistical modeling.



XIANGJUN YU received the Ph.D. degree in oceanography from Dalian Naval Academy, Dalian, China, in 2011.

He is currently a Lecturer in Military Oceanography with Dalian Naval Academy. His current research interest includes ocean numerical simulation and analysis.



PAN QIN received the Ph.D. degree in information science from Kyushu University, Fukuoka, Japan, in 2008.

He was a Research Fellow with the Faculty of Mathematics and the Institute of Mathematics for Industry, Kyushu University, from 2008 to 2013. He is currently an Associate Professor with the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China. His current research interests

include statistical modeling and stochastic control theory.

...