iOceanSee: A Novel Scheme for Ocean State Estimation Using 3D Mobile Convolutional Neural Network

HUAFENG WU¹, (Senior Member, IEEE), YUANYUAN ZHANG¹, JUN WANG², (Senior Member, IEEE), WEIJUN WANG¹, JIANGFENG XIAN¹, JING CHEN¹, XIANGYI ZOU³, AND PRASANT MOHAPATRA⁴, (Fellow, IEEE)

¹Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China
²Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL 32816, USA
³National Deep Sea Center, Shandong 266237, China
⁴Department of Computer Science, University of California at Davis, Davis, CA 95616, USA

Corresponding authors: Yuanyuan Zhang (2357612743@qq.com), Jun Wang (jun.wang@ucf.edu), and Prasant Mohapatra (pmohapatra@ucdavis.edu)

This work was supported in part by the National Natural Science Foundation of China under Grant 51579143, Grant 51709167, and Grant 61701299; in part by the Shanghai Committee of Science and Technology, China, under Grant 18040501700; in part by the Major Special Science and Technology Project of Hainan Province under Grant ZDKJZ2017012; in part by the top-notch innovative program for postgraduates of Shanghai Maritime University under Grant 2019YBR006; and in part by the Graduate Innovation Foundation of Shanghai Maritime University under Grant 2016ycx042 and Grant 2017ycx030.

ABSTRACT Ocean state estimation is a basic problem in the field of ocean engineering. Under the trend of data-driven, the development of intelligent ship decision-making, ocean energy system design and other aspects, are inseparable from the estimation of wave parameters in the ocean area. In recent years, researchers have developed remote sensing technology to monitor ocean waves. However, sensor-based methods all have a key limitation, which is high cost and fault worry. More importantly, one major limitation exists in current research: due to lack of change information and relying on a single feature of spatial data, the final predictive results are inaccurate. Adopting a 3D Convolutional Neural Network is a possible solution to improve the detection accuracy. Unfortunately, it cannot be deployed in the ocean environment due to lack of physical network connections. To resolve these issues, we develop a light-weight version of 3D Convolutional Neural Network, namely a low-cost, high-accuracy detection scheme to foresee ocean wave parameters using a 3D Mobile Convolutional Neural Network technique called iOceanSee in the marine environment. iOceanSee employs a mobile terminal composed of low-cost measuring equipment and non-interference (except light) device-an RGB camera to collect video data in real time. It extracts both space and time features through three-dimensional depthwise separable convolutions. More specifically, iOceanSee is able to capture the encoding motion information from multiple adjacent frames of the video, according to which period and height of the waves being evaluated. Our experimental results conclude that iOceanSee obtains comparable performance to 3D Convolutional Neural Network and outperforms other models in terms of measurement accuracy in the marine environments.

INDEX TERMS Ocean wave detection, 3D mobile convolutional neural network, deep learning, identification of wave parameters.

I. INTRODUCTION

Ocean state estimation plays an important role in coastal monitoring, port design, autonomous intelligent ship design, etc. Specific marine engineering operations require real-time prediction of the marine status at specific locations. However, the marine environment is relatively complex, especially wind, wave, current and other factors, often have a negative impact. Therefore, we must ensure that certain engineering tasks are completed accurately and safely, considering the marine parameters.

In the marine environment, wave parameters are more and more closely related to people’s marine operations, such as the design of marine energy system. How to detect wave...
parameters plays an important role in many marine engineering areas, such as the wave energy conversion [1] of marine energy management. For example, wave height, period, level, energy conversion and etc. are modern research hotspots.

As a renewable energy source, ocean energy converter has been studied by many scholars. Yin and Xiuxing [2] designed a novel kite-type ocean energy converter using current and wave energy, which can generate electricity under current and wave conditions at different ocean current velocities, wave periods and altitudes. Xia et al. [3] proposed a quasi-Habach magnetized field-modulated tubular linear generator (FMTLG) for wave energy conversion. The simulation results show that FMTLG is superior to TLPMG in anti-EMF and other aspects. Recent years have seen increasing research activity on wave height. Alexandre et al. [4] proposed a hybrid genetic algorithm based extreme-learning machine method to reconstruct effective waves. Considering the spatial correlation between the position values of adjacent buoys, they solved the problem of local reconstruction of failed buoys by utilizing the wave parameters of adjacent buoys. Salcedo-Sanz et al. [5] proposed to apply the Support Vector Regression (SVR) methodology to estimate the significant wave height of X-band marine radar image of sea surface by shadow effect, by which good prediction results in wave height can be obtained. Ti et al. [6] estimated the effective wave heights near shore by using the response surface method prediction equation. In the case of inadequate or difficult acquisition of near-shore measurements, they used the numerical wave model and semi-analytical fitting model derived from the conservation equation of wave energy flow to fit the response surface of near-shore wave height, establishing the prediction equation and predicting the effective wave height.

Unfortunately, most of the research methods such as x-band wave radar and meteorological remote sensing satellites have higher cost and accuracy limitations. Specifically, traditional ocean state estimation techniques usually include manual observations, wave buoys, x-band wave radars, and meteorological remote sensing satellites or weather forecasts [7]. There is a strong subjectivity in manual observation. Wave buoys have certain reliability, but the ability to resist wind and waves is weak. The use of x-band radar technology is not only expensive, but also requires regular calibration and maintenance. Meteorological remote sensing satellites are sensitive to clouds, and sensor-based methods bear a key limitation, which is high cost and fault worry.

In order to overcome the shortcomings of traditional estimation methods, researchers have carried out extensive exploration. They divided the estimation methods of ocean parameters into model-based and non-model-based methods. Model-based methods mainly use domain knowledge to build mathematical models. However, due to their excessive dependence on mathematical models and corresponding assumptions, these methods are prone to generate false recognition due to the dynamic impact of the marine environment. Non-model-based methods mainly use traditional machine learning or deep learning techniques to extract time and frequency features. The advantage of using this kind of method is that it does not depend on previous domain knowledge and is easier to learn. Therefore, comparing the two types of means, we choose a model-free method. Many scholars have studied the height of the waves, but there are few studies on the period of the waves.

In this paper, we studied these two kinds of ocean parameters. In order to improve the accuracy of the estimation, we must consider both time and space information. Existing methods are lacking in change information over time and exclusively rely on a single feature of spatial data, the final predictive results are less accurate. Adopting a 3D Convolutional Neural Network is a possible solution to improve the detection accuracy. The 3D convolution neural network (3D CNN) can extract the temporal and spatial features of the video by using the pictures of continuous frames in the video, so as to capture the object features better. Video data itself is a kind of high-dimensional data, which takes up more storage space for the storage of video data. Unfortunately, 3D CNN cannot be deployed in the ocean environment due to lack of physical network connections. In addition, 3D CNN contains many unneeded information which leads to non-negligible compute and storage overhead. To resolve these issues, we develop a light-weight version of 3D Convolutional Neural Network, namely a low-cost, high-accuracy detection scheme to foresee ocean wave parameters using a 3D Mobile Convolutional Neural Network technique called iOceanSee in the marine environment. iOceanSee employs a mobile terminal composed of low-cost measuring equipment and non-interference (except light) device-an RGB camera to collect video data in real time. It extracts both space and time features just for wave related data through three-dimensional depthwise separable convolutions. More specifically, iOceanSee is able to capture the encoding motion information from multiple adjacent frames of the video, according to which period and height of the waves being evaluated.

iOceanSee adopts measuring equipment and an RGB camera on vessels named YUMING and XIANGYANGHONG81 to collect real-time video data and measure the height and period of the waves through an underwater pressure wave meter simultaneously. In the training phase, we take collected video data as our training data. Using them as the input, we train our model using the online error back-propagation algorithm as described in [8] on NVIDIA TITAN XP GPUs. We can extract the information both in space and time, obtain the final characteristic representation by convolution and subsampling of adjacent video frames and evaluate the wave height and period according to the training results. The output data values are compared with the actual measured values and the results obtained by other methods to evaluate the performance of the model. The experimental results show that iOceanSee achieves near-optimal accuracy of the 3D CNN model which is superior to other baseline methods. In more detail, we train and optimize the 3D MCNN using multi-task loss function. In the inference phase, the input data of the network is the real-time video data acquired by the RGB
camera in the real-life marine environment. The value of the network output is the ocean wave estimation parameters predicted by our model. The schematic workflow diagram of the system is illustrated in Figure 1.

The rest of this paper is organized as follows: Next section outlines the related work of this paper. Section 3 describes the three-dimensional convolution operation used in the wave parameter detection scheme, the proposed model and implementation details. Section 4 presents details of the ocean wave dataset, experimental results and discussion, evaluation metrics and model choices. We make a summary in Section 5.

II. RELATED WORK

Although there are many research works in wave parameters prediction, their results are not accurate since most studies employ only single input data for prediction. In order to accurately predict or numerically calculate wave height and period data, there is a need for taking multiple eigenvalues into consideration. Recent years have seen several new methods are successfully exploited. For example, video recognition technology has been widely used in many fields, such as text recognition [9], face recognition [10], medical symptom recognition [11], human behavior re-identification [12], [13], crack identification [14] and so on. As the core machine of video recognition technology, deep learning model learns feature hierarchy by constructing high-level features from low-level features. It adopts supervised or unsupervised learning methods, adjusts the parameters of the system according to the output results, and automatically learns features according to the input data.

Convolution neural network [15], [16], as a deep learning model, alternately applies the training filter and local neighborhood pool operation to the original input image, which is mainly used for two-dimensional images. With the rapid development of artificial intelligence, two-dimensional convolution neural network has matured in the field of computer vision, and three-dimensional convolution neural network has made a major breakthrough. Two-dimensional convolution neural network can only be used for static analysis of two-dimensional images, so it can only extract spatial features. In contrast, three-dimensional convolution neural network can extract not only spatial features, but also temporal features. It can extract temporal and spatial features from images of continuous frames in video, so as to capture object features better. Currently, many scholars have studied 3D CNN. Ji et al. [17] proposed a new three-dimensional CNN action recognition model, which was successfully applied to the real environment of airport surveillance video to identify human behavior. Li et al. [18] employed Gabor filter to analyze the action in video data and adopt KTH dataset to validate the model, and obtain good results. Ren et al. [19] combined several three-dimensional convolution neural networks altogether to realize automatic segmentation of small tissues in head and neck CT images. Ceschin et al. [20] developed a computational framework for automatic classification of brain dysplasia from neonatal MRI, which used three-dimensional convolutional neural network to classify cerebellar dysplasia in full-term infants with congenital heart disease, and designed a program that can be extended to both internal and external brain imaging of neonates. Zhang et al. [21] proposed a new framework of deep three-dimensional convolution neural network (3D-cnns) feature...
network to learn the processing characteristics of mechanical parts CAD model and proved that the feature network can significantly improve the most advanced manufacturing feature detection technology by numerical experiments.

In recent years, many scholars explore how to employ the machine learning technique to study the wave parameters. Oh and Suh [22] developed an EOFWNN model, which takes into account the relationship between spatial distribution of meteorological variables and waves, and predicts wave heights at multiple locations. Compared with the previous wavelet neural network model, this model trains the observed wave data at each location, which is convenient and improves the prediction accuracy. Cornejo-Bueno et al. [23] proposed a wave parameter estimation method based on regression genetic fuzzy system (GFS, FRULER) to predict the wave height and energy. Jongchul and Dongseob [24] used CCD camera to monitor the wave parameters. They compared the wave information collected by the Hu-moments algorithm in Gyoam Beach, South Korea, with that obtained by the underwater pressure wave instrument, and concludes that the method can detect the wave period with slight bias. The main details of these algorithms are as follows.

### A. EOF-WAVELET AND NEURAL NETWORK HYBRID MODELS

Oh and Suh [22] combined empirical orthogonal function analysis and wavelet analysis with neural network, and formulated a hybrid model - EOFWNN. By considering the relationship between wave and meteorological variables, they forecasted the significant wave heights for various lead times of eight wave stations in the coastal waters around the East/Japan. In the study, they mainly introduced EOF analysis to explain the relationship between waves and meteorological data and enable the model to predict waves at multiple stations simultaneously. You can see that spatiotemporal data is decomposed by EOFs as follows:

\[
T(x,t) = \sum_n T_n(t) B_n(x) \tag{1}
\]

where, \( \{B_n(x)\} \) are called the eigenfunctions, and \( \{T_n(t)\} \) are the expansion of the coefficients that are functions of time. \( T(x,t) \) is called the EOF decomposition only if \( \{B_n(x)\} \) are mutually orthogonal to each other and \( \{T_n(t)\} \) are mutually uncorrelated, that is,

\[
B_n(x) \cdot B_m(x) = \frac{1}{N} \sum_{x=1}^{N} B_n(x) B_m(x) = \delta_{nm} \tag{2}
\]

\[
T_n(t) \cdot T_m(t) = \frac{1}{N} \sum_{t=1}^{N} T_n(t) T_m(t) = \lambda_n \delta_{nm} \tag{3}
\]

where \( \delta_{nm} \) is the Kronecker delta, and \( \lambda_n \) is the eigenvalues. Often, \( \{B_n(x)\} \) are called the loading vectors, \( \{T_n(t)\} \) represent the principal component (PC) time series.

### B. FUZZY RULE LEARNING THROUGH EVOLUTION FOR REGRESSION

Cornejo-Bueno et al. [23] employed genetic fuzzy system to estimate significant wave height and energy flux, which consists of three staged algorithm that combines an instance selection method for regression, a multigranularity fuzzy discretization of the input variables and an evolutionary algorithm to generate accurate and simple Takagi-Sugeno-Kang (TSK) fuzzy rules. The first two components of FRULER are part of a two-stage preprocessing step to improve the accuracy and simplicity of the fuzzy rules which are obtained by the evolutionary algorithm. The evolutionary algorithm searches for the best database configuration using the obtained fuzzy partitions and takes advantage of the Elastic Net method linearly, which combines the \( l_1 \) (Lasso regularization) and \( l_2 \) (Ridge regularization) penalties of the Lasso and Ridge method automatically learn the importance weight of each variable, minimizing the following equation:

\[
\hat{\beta} = \arg \min_{\beta} \| Y - X \cdot \beta \|_2^2 + \lambda \cdot \alpha \cdot \| \beta \|_2^2 + \lambda \cdot (1 - \alpha) \cdot \| \beta \|_1 \tag{4}
\]

where \( \beta \) is the coefficients vector, \( Y \) is the outputs vector, \( X \) is the inputs matrix, \( \lambda \) is the regularization parameter and \( \alpha \) represents the trade-off between \( l_1 \) and \( l_2 \) penalization.

Finally, the robust and very accurate prediction of the two target variables is obtained.

### C. HU-MOMENTS ALGORITHM

Jongchul and Dongseob [24] proposed a video-based beach monitoring system. Because the change of wave arriving in shallow water is very discontinuous in CCD image clipping, they use Hu-moments algorithm to detect the wave period with weak offset, thus clarifying the complex characteristics of wave distribution and its time variation. They first extracted the outline of each video patch from data captured at Gyoam Beach, and then using seven Hu-moments to estimate the similarity matrix of the shapes in video patches \( V_i \) and \( V_j \) which have different frames:

\[
I_{ij} = \sum_{g=1}^{7} \left| 1/C_{g}^{i} - 1/C_{g}^{j} \right| \tag{5}
\]

where \( V_i (i = 1, 2, \ldots, N) \) is a sample patch of N video frames, it can convert to an edge map \( F_i \) utilizing the Canny edge detector, \( C_{g}^{i} = \sin \left( h_{g}^{i} \right) \log_{10} \left[ h_{g}^{i} \right] \), the seven Hu-moments \( h_{g}^{i} \) \((g = 1, \ldots, 7)\) can be calculate by analyzing the first three central moments. After that, the radial symmetry kernel is used to represent the period and height of the wave.

### III. PROPOSED WORK

The detection of wave parameters has always been the key in marine engineering applications. In recent years, researchers have developed techniques for monitoring wave parameters using remote sensing technology, but sensor-based methods...
all have a key limitation, which is high cost and fault worry. In addition, most solutions rely on spatial data, i.e., a single feature to design their predictive mechanism, and thereby oversighting the change information over time. This leads to the inaccurate final results, which might dissatisfy the user level agreement. It may be noted that, in the actual marine environment, due to a strong dynamicity, change information in a continuous time window is of great importance. There is a need for taking both time and space information into consideration to accurately detect or predict the height and period of ocean waves. iOceanSee captures multi-dimensional features of real-time wave data and develops a wave parameter detection scheme based on 3D mobile convolution neural network. Compared with the sensor-based method, low-cost NVIDIA Jetson TX2 and non-interference (except light) devices—RGB camera are employed to collect video data, and fed together with the motion information of the video frames into the iOceanSee inference model. Given both time and space information, iOceanSee is able to precisely forecast the period and height of the waves. As a feasible working scheme, iOceanSee employs depth separable convolution to construct lightweight networks to reduce computational overhead. We develop a low-cost, high-accuracy detection scheme named iOceanSee to track and predict ocean wave parameters using a 3D Mobile Convolutional Neural Network. The overall framework of the proposed model is shown in Figure 2.

A. 3D CONVOLUTION NEURAL NETWORK

Regarding the convolution of two-dimensional convolution map, since the input is provided entirely by the images, the feature can only be calculated from the spatial dimension, and the time dimension can not be convoluted. When using video to analyze problems, it is necessary to capture motion information in multiple consecutive frames. For example, Ji et al. [17] proposed three-dimensional convolution in convolution layer by taking account into both spatial dimension and temporal dimensions. Multiple consecutive video frames are superimposed to form an input cube. After convoluting the cube with the three-dimensional core, and adding the corresponding bias, the result will become the input value of the next layer after passing through the activation function. Thus, the whole three-dimensional convolution process is completed. Each feature map is connected to several adjacent frames in the previous layer, and hence the motion information is captured.

Formally, the value at position \((x, y, z)\) on the \(j\)th feature map in the \(i\)th layer is denoted as \(v_{ij}^{xyz}\), which is expressed as

\[
v_{ij}^{xyz} = \tanh \left( b_{ij} + \sum_{m} \sum_{p} \sum_{q} \sum_{r} w_{ijm}^{pqr} (x+p)(y+q)(z+r) \right)
\]

where, \(\tanh(\cdot)\) is the hyperbolic tangent function, \(b_{ij}\) is the bias for this feature map, \(m\) indexes over the set of feature maps in the \((i-1)\)th layer connected to the current feature map, \(w_{ijm}^{pqr}\) is the value at the position \((p, q, r)\) of the kernel connected to the \(m\)th feature map in the previous layer, \(P_i\) and \(Q_i\) are the height and width of the kernel, \(R_i\) indicates the thickness of the convolution nucleus, in other words, it is the size of the 3D kernel along the temporal dimension.

B. OVERALL FRAMEWORK

The whole scheme realizes the real-time monitoring of wave height and period in the marine environment. This scheme uses the mobile terminal composed of NVIDIA Jetson TX2 and camera to monitor the waves in real time, collect the corresponding video data, and train the video data offline. The data acquisition device shown in Figure 1 is comprised of NVIDIA Jetson TX2, an RGB camera (1080 x 1920 pixels with 30 frame per second), a computer and NVIDIA TITAN XP GPUs. The RGB camera is connected to NVIDIA Jetson TX2 via USB interface. The network is implemented using NVIDIA Jetson TX2. NVIDIA TITAN XP GPUs are used to train our network. Actual parameters measured by an underwater pressure type wave meter.

For the actual marine environment, the changes of wave parameters in different regions are often affected by natural
factors such as strong winds. Hence, carrying out a real-time monitoring is challenging. It is even harder to conduct monitoring in real time. However, the detection of ocean wave parameters often makes a non-negligible impact on the development of other applications in the marine engineering field. In the marine environment, the video signal we collect is three-dimensional, and it represents an effective means of identification to employ the space-time information directly. Therefore, it is nature to apply a three-dimensional convolutional neural network to identify wave parameters.

We propose a three-dimensional mobile convolution neural network model for ocean wave parameter detection using an architectural design similar to the well-known 3D CNN [17] network, as shown in Figure 2 above. Inspired by literature [25], we employ depthwise separable convolution to construct lightweight networks to reduce computational overhead. The details are as follows.

As shown in Figure 2, seven frames of size $60 \times 40$ centered on the current frame and selected from video is chosen as inputs to the 3D MCNN model. A set of hardwired kernels is applied to generate multiple channels of information from the input frames, which ends up with 33 feature maps in the second layer expressed as H1 in five different channels, being denoted by gray, gradient-x, gradient-y, optflow-x, and optflow-y respectively. The gray channel contains the gray pixel values of the seven input frames. The feature maps in the gradient-x and gradient-y channels are acquired by calculating gradients along the horizontal and vertical directions on each of the seven input frames, and the optflow-x and optflow-y channels contain the optical flow fields along the horizontal and vertical directions, respectively, which is computed from adjacent input frames. Our prior knowledge concerning features is encoded by the hardwired layer, and this scheme usually contributes to better performance as it is compared to the random initialization.

Next, original 3D convolutions with a kernel size of $7 \times 7 \times 3(7 \times 7$ in the spatial dimension and 3 in the temporal dimension) are applied on each of the five channels, separately. We have the same time dimension as before, and only do separable convolution for space dimension.

Likewise, we apply 3D convolution in the next convolution layer Convolution4 (Abbreviated as C4 in Figure 2) with a kernel size of $7 \times 6 \times 3(7 \times 6$ in the spatial dimension and 3 in the temporal dimension). Consistent with what was mentioned earlier, we only do separable convolution for space dimension. Standard convolution layer of size $7 \times 6 \times 3 \times 30 \times 54$ is replaced by depthwise convolution layer with size $7 \times 6 \times 3 \times 30$ and pointwise convolution layer with size $1 \times 1 \times 3 \times 30 \times 54$. Standard convolution layer of size $7 \times 6 \times 3 \times 16 \times 24$ is replaced by depthwise convolution layer with size $7 \times 6 \times 3 \times 16$ and pointwise convolution layer with size $1 \times 1 \times 3 \times 16 \times 24$. They generated a total of 78 feature maps. You can see the details in Table 1. DeepwiseConv1 and Conv1 represent a set of corresponding deep convolution and point convolution, and DeepwiseConv2 and Conv2 are the same.

The next layer S5 is obtained by applying $3 \times 3$ subsampling on each feature map in the C4 layer, leading to the same number of feature maps with a reduced spatial resolution. At this stage, the size of the temporal dimension is already relatively small (3 for gray, gradient-x, gradient-y, and 2 for optflow-x and optflow-y), so we conduct convolution only in the spatial dimension at this layer. Because the size of the convolution kernel used is $7 \times 4$, the sizes of the output feature maps are reduced to $1 \times 1$. The C6 layer consists of 128 feature maps of size $1 \times 1$, each of which is connected to all 78 feature maps in the S5 layer.

The seven input frames have been converted into a 128D feature vector capturing the motion information in the input frames after multilayer convolution and subsampling. The output layer is made up of the same number of units as the number of wave features, and each unit is fully connected to each of the 128 units in the C6 layer. In this design, we essentially apply a linear classifier on the 128D feature vector for action classification.

Our 3D MCNN model also consists of two fully connected layers (FC1, FC2), which include 2048, 512 neurons, respectively. We mainly have two main tasks, one is wave period estimation and the other is wave height estimation. All trainable parameters in this model are initialized randomly and trained by the online error back-propagation algorithm as described in [6], which makes loss minimize.

It is worth noting that the lightweight implementation is only for two-dimensional spatial domains. For three-dimensional convolutional neural network, the characteristics of time domain should also be taken into consideration. During the splitting process, the time dimension remains unchanged, which transforms the original 3D convolution

### TABLE 1. Network configuration of the 3D MCNN model.

| Layer     | Configuration                                      | output size |
|-----------|----------------------------------------------------|-------------|
| Hardwire1 | —                                                  | $33 \times 60 \times 40$ |
| Convolution2 | DeepwiseConv1 7×7×3×21 54×34×21      | 54×34×30    |
|           | Conv1 1×1×3×21×30 54×34×30                        | 54×34×12    |
|           | DeepwiseConv2 7×7×3×12 54×34×12                   | 54×34×16    |
|           | Conv2 1×1×3×12×16 54×34×16                        | 54×34×16    |
| Subsampling3 | 2×2                                              | 27×17×46 |
| Convolution4 | DeepwiseConv1 7×6×3×30 21×12×30                     | 21×12×30    |
|           | Conv1 1×1×3×30×54 21×12×54                        | 21×12×54    |
|           | DeepwiseConv1 7×6×3×16 21×12×16                   | 21×12×16    |
|           | Conv2 1×1×3×16×24 21×12×24                        | 21×12×24    |
| Subsampling5 | 3×3                                              | 7×4×78 |
| Convolution6 | 7×4                                               | 1×1×128    |
| FC1       |                                                    | 2048 neurons|
| FC2       |                                                    | 512 neurons |
kernel into multiple lightweight 3D cubes, and the output results retain the original time domain information.

**C. IMPLEMENTATION DETAILS**

Table 1 details the Network configuration of the entire 3D MCNN model, including Layer, filters shape, output size and the number of fully connected layers neurons. Kernel size of DeepwiseConv is denoted as Height × Width × Time dimension × Input channels. Conv is Height × Width × Time dimension × Input channels × Output channels. The loss function for the two main tasks is defined as follows.

\[
L_p(\Theta) = \frac{1}{N} \sum_{i} \| F_p(X_i; \Theta) - P_i \|^2
\]

\[
L_H(\Theta) = \frac{1}{N} \sum_{i} \| F_H(X_i; \Theta) - H_i \|^2
\]

where, \( \Theta \) is the parameters set of the 3D MCNN model and \( N \) is the training sample numbers. \( L_p \) is the loss between the estimated period value \( F_p(X_i; \Theta) \) and the true value \( P_i \) measured by wave meter. And similarly, \( L_H \) is the loss between the estimated wave height \( F_H(X_i; \Theta) \) and the true value \( H_i \) measured by wave meter, respectively. Here, we adopt Relative Euclidean Distance in these two objective losses. We use back propagation to make the loss minimized.

\[
L(\Theta) = \alpha L_p + \beta L_H
\]

Then, we adopt a joint multi-task scheme to minimize the loss function. \( \alpha \) is used to control \( L_p, \alpha \in (0, 1) \), \( \beta \) is taken to regulate \( L_H, \beta \in (0, 1) \), and they satisfy formula \( \alpha + \beta = 1 \).

We use a linear multi-task iterative training process to train two tasks, train the input data in batches to minimize \( L_p \) and \( L_H \), and use \( \alpha \) and \( \beta \) to control the proportional weights of the two loss functions. Through multiple iterations of training, \( L(\Theta) \) is minimized.

**Algorithm 1** Training Process of Our Model

Input: Size-normalized patches with period and height values from the whole training data

Output: Parameters \( \Theta \) for 3D MCNN model

1. Initialize each parameter randomly;
2. \( t = 1 \)
3. while (\( t \leq T \)) do
4. Back propagation is used to learn \( \Theta \), until the joint loss function (Equation 9) minimize;
5. end while

We compare the results after training with those measured by the underwater pressure type wave meters and get the corresponding error analysis. Furthermore, the results are also compared with EOFWNN, FRULER, HM-CCD and 3D CNN to evaluate measurement performance.

**IV. EXPERIMENTS AND RESULTS**

We collect wave video image data on YUMING and XIANGYANGHONG81 on March 6, 10, 14, 18, 22, 26, 30 and April 3, 7, 11 and 15, 2019, respectively, and use opencv to preprocess wave video data to crop the video to the input size of our algorithm. Furthermore, we manually divide the wave video data set into training set and test set, and label the data using underwater pressure wave meter. The video data is used as the input of the 3D MCNN network, and the wave period and wave height are used as the output of the network. Then, we make use of the training set data to train our model, and use the test set data to verify the performance and generalization ability of the model. Finally, we compare it with other algorithms that include EOFWNN, FRULER, HM-CCD and 3D CNN. The experimental results show that the proposed 3D MCNN algorithm performs better than other algorithms except 3D CNN in predicting wave period and wave height. In terms of measurement accuracy, although our 3D MCNN is slightly inferior to 3D CNN, its calculation parameters are greatly reduced.

**A. OCEAN WAVE DATASET**

Wave video data acquisition was carried out at 8 stations on YUMING and XIANGYANGHONG81. These data were used to train and evaluate our algorithm. The locations of the data acquisition sites are shown in Figure 3.

We divide the data set collected on different sites into two parts: training set and test set. The video data used in training set and test set are from 8 sites. Training set consists of 184 video sequences with a time length of 1 minute, and the test set contains 11 video sequences with 30 minutes each. Figure 4 shows some example frames of the wave dataset on XIANGYAGNHONG81 and YUMING.

The proposed 3D MCNN algorithm can automatically extract motion information between frames of ocean wave video. Compared with other single-frame image processing algorithms, 3D MCNN and 3D CNN algorithm can not only accurately calculate the wave height, but also accurately estimate the wave period. We train different algorithms on the data sets collected from YUMING and XIANGYANGHONG81 ships, and then use model of each algorithm to calculate the period and height of the waves on the test set respectively, and compare the calculation accuracy of each algorithm.

Figures 5 and 6 show the actual measured average ocean wave period and height of wave video data collected at eight stations on the YUMING and XIANGYANGHONG81 and predicted values of various algorithm models, respectively. Both figures record video data of 11 days. As can be seen from the figures, blue curves represent actual measurements, and the estimated values from each model are represented by points of different colors. Each segment represents the actual value measured by the wave meter and the resulting value calculated by each algorithm model on the day of the horizontal coordinate date. For example, the period from March 6th to 10th indicates that the measurement value and the calculation value of each algorithm model are on March 6th. When the measured wave height and period are relatively high, there is a large deviation between the estimated value and the real value.
of each algorithm. Especially when the actual wave period is more than 9 seconds and wave height exceeds 0.6 meters, the deviation of the estimation of wave period increases, but the estimation errors of 3D MCNN and 3D CNN are both relatively smaller than other algorithms and to a lesser extent, our 3D MCNN is not as good as 3D CNN, while results of the two measurements are nearly close to each other.

**B. EVALUATION METRICS**

In practice, the indicators used to assess the accuracy of predictions vary. In this experiment, correlation coefficient square ($R^2$), index of agreement ($I$), root mean square error ($RMSE$) and its normalized ($NRMSE$), mean absolute error ($MAE$), are selected by us to evaluate the models.
The scatter plot is employed to show them. The calculation formulas are as follows:

\[
R^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \bar{x})^2} \tag{10}
\]

\[
I = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (|y_i - \bar{y}| + |x_i - \bar{x}|)^2} \tag{11}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_i (x_i - y_i)^2} \tag{12}
\]

\[
NRMSE = RMSE / (x_{\text{max}} - x_{\text{min}}) \tag{13}
\]

\[
MAE = \frac{1}{n} \sum_i |x_i - y_i| \tag{14}
\]

where, \(x_i\) and \(y_i\) are the measured wave height, predicted wave height, \(\bar{x}\) and \(\bar{y}\) are mean measured wave height and mean predicted wave height, \(x_{\text{max}}\) and \(x_{\text{min}}\) are the maximum and minimum measurements, and \(n\) is the number of measurements.

In the experiment, we compared our model with the following existing prediction methods including EOFWNN, FRULER, HM-CCD. In addition, we also use the most basic 3D CNN method for experiments, and compare the results of our proposed 3D MCNN with all other methods.

C. EXPERIMENTAL RESULTS AND DISCUSSION

1) HEIGHT COMPARISON OF OCEAN WAVE

We show a scatterplot representation of the prediction result of the index correlation coefficient square \((R^2)\). Based on the experimental results, we compared the five methods.

As shown in Figure 7, the fitting relationship between the predicted value, namely, the estimated value of the algorithm model and the actual measured value is clear. The determination coefficients of the five models are 0.9876, 0.9883, 0.9943, 0.9990, 0.9986, respectively. By observing and comparing the predicted and theoretical values of wave heights of the four models, we can find that 3D CNN gives a better fit. The error between the estimated and measured wave heights calculated by 3D CNN is smaller than others, with its higher correlation coefficient. The fitting degree of 3D MCNN is slightly lower than that of 3D CNN, and its error value is still smaller than that of other algorithms.

To ensure our choosing model is best, we use different performance indicators to measure the estimation accuracy of wave height, and compare 3D MCNN with four other different estimation methods, and the results are shown in Figure 8.

We use a bar chart to summarize the results of several evaluation indicators in Figure 8. As you can see from the picture, \(R^2\) and itself \(R\) of five model are very close to each other, but there are some differences in fact. For \(R^2\), it can be seen from the analysis of figure 7 and itself is similar. And \(I\) varies from 0.8384 to 0.9777, 3D MCNN is second only to 3D CNN in five methods. The difference between the two methods is small, but they both get better performance than other methods. The summary results of \(RMSE, NRMSE, MAE\) are obviously different. We can see that for \(RMSE\) and \(NRMSE\) indicators, EOFWNN gets the maximum value, \(MAE\) indicator shows that FRULER acquire the biggest value. The statistical results of the three indicators all express that the minimum value of the 3D CNN is obtained. We know that the smaller the error value, the better the prediction performance. Therefore, the 3D CNN obtains the best performance in wave height prediction. In addition, we observed that the \(RMSE\) difference between 3D MCNN and 3D CNN is only 0.65%, \(NRMSE\) is 1.19%, and \(MAE\) is 0.0124%. Therefore, we can see that their performance is relatively close.
2) PERIOD COMPARISON OF OCEAN WAVE

Similarly, scatter plots of wave periods for each model are as above. As shown in Figure 9, it can be clearly compared that the goodness of fit of wave period values estimated by EOFWNN, FRULER and HM-CCD is not as good as that estimated by 3D CNN and 3D MCNN. The error between the estimated and measured wave periods calculated by 3D MCNN is small, and its correlation coefficient is 0.9945, which is only 0.0039 lower than 3D CNN.

As shown in Figure 10, for wave period, results of $R^2$ and itself $R$ of five model are slightly more obvious than the wave height, and the difference between the best fit and the worst fit is 0.0208 and 0.0104, respectively. However, the fit of 3D CNN and 3D MCNN ranks first and second, respectively, and its square can be seen from the analysis of figure 9 in detail and itself is similar. And the maximum value of $I$ is 0.8041, 3D MCNN is still second only to 3D CNN in five methods and its goodness of fit is 0.7074, which is higher than other methods. Compared with wave height, there are more obvious differences in the results of $RMSE$, $NRMSE$, $MAE$. According to results of $NRMSE$, EOFWNN gets the worst performance, 3D CNN and 3D MCNN are only 0.1335 and 0.1745 respectively. The $MAE$ indicator shows that FRULER and EOFWNN have obtained comparable values, and the performance is not as good as the other three methods. The statistical results of the three indicators all indicate that 3D CNN still performs best, and 3D MCNN performs second, so 3D CNN has obtained the best performance in wave period prediction. Besides, we observed that the difference of $RMSE$, $NRMSE$, and $MAE$ of 3D MCNN and 3D CNN is 0.2006, 0.041, 0.1771, respectively. The difference is slightly higher than the wave height detection result, and the performance is still relatively close.

To further demonstrate the performance improvement of our model relative to other methods, we use $\Delta IMT = ((I_{3D\ MCNN} - I_{\text{base}})/I_{\text{base}}) \times 100\%$ to calculate improvements, where $I_{\text{base}}$ is the index of agreement for the benchmark method, while $I_{3D\ MCNN}$ is the index of agreement of the 3D MCNN. The performance improvements of $R^2$ and $R$ are calculated in a similar way. For other two metrics $NRMSE$ and $MAE$, we use a manner as $RMSE$, which is as follows.

$$\Delta IMT = ((RMSE_{\text{base}} - RMSE_{3D\ MCNN})/(RMSE_{\text{base}})) \times 100\%$$

where $RMSE_{\text{base}}$ is the $RMSE$ for the benchmark method, while $RMSE_{3D\ MCNN}$ is the $RMSE$ of the 3D MCNN.

TABLE 2 and TABLE 3 show a summary of the wave height and cycle results obtained using $\Delta IMT$. From the data in Tables 2 and 3, we can see that our 3D MCNN has a certain improvement over EOFWNN, FRULER, and HM-CCD. In table 2, though there is a small improvement on $R^2$ and $R$, yet our proposed method can acquire up to 66.34% improvement on the $RMSE$, up to 66.31% improvement on the $NRMSE$, up to 16.61% improvement on the $I$, up to 74.15% improvement on the $MAE$. In table 3, though there is a small improvement on $R^2$ and $R$, yet our proposed method can be up to 40.07% improvement on the $RMSE$, up to 40.08% improvement on the $NRMSE$, up to 67.16% improvement on the $I$, up to 54.17% improvement on the $MAE$.

D. MODEL CHOICES

TABLE 2 and TABLE 3 show that our algorithm has different degrees of improvement compared with EOFWNN,
TABLE 2. Height performance improvements attained by our model.

| EVALUATION METRICS | WAVE PARAMETER | WAVE HEIGHT |
|--------------------|----------------|-------------|
| R²                 | R              | RMSE        | NRMSE     | I         | MAE        |
| EOFWNN             | 1.11%          | 0.55%       | 66.34%    | 66.31%    | 16.61%     | 68.52%     |
| FRULER             | 1.04%          | 0.52%       | 65.51%    | 65.27%    | 12.42%     | 74.15%     |
| HM-CCD             | 0.43%          | 0.22%       | 50.48%    | 50.40%    | 6.90%      | 49.44%     |

TABLE 3. Period performance improvements attained by our model.

| EVALUATION METRICS | WAVE PARAMETER | WAVE PERIOD |
|--------------------|----------------|-------------|
| R²                 | R              | RMSE        | NRMSE     | I         | MAE        |
| EOFWNN             | 1.74%          | 0.86%       | 40.07%    | 40.08%    | 67.16%     | 54.01%     |
| FRULER             | 0.93%          | 0.46%       | 28.52%    | 28.51%    | 8.08%      | 54.17%     |
| HM-CCD             | 0.53%          | 0.26%       | 19.52%    | 19.51%    | 12.73%     | 36.37%     |

TABLE 4. depthwise separable: 3D MCNN vs full convolution 3D CNN.

| MODEL             | 3D MCNN | 3D CNN |
|-------------------|---------|--------|
|                   | HEIGHT  | PERIOD | HEIGHT | PERIOD |
| R² (%)            | 99.86   | 99.06  | 99.90  | 99.45  |
| R (%)             | 99.93   | 99.53  | 99.95  | 99.72  |
| RMSE              | 0.0409  | 0.8531 | 0.0344 | 0.6525 |
| NRMSE             | 0.0747  | 0.1745 | 0.0628 | 0.1335 |
| I(%)              | 97.77   | 70.74  | 98.23  | 80.41  |
| MAE(%)            | 0.0227  | 0.4529 | 0.0103 | 0.2758 |
| Million Multi-Adds| 31.5    | 286.8  |
| Prediction time(ms)| 8.6     | 78.5   |
| Thousand Parameters| 21.7    | 375.9  |

FRULER, HM-CCD. We can see from the evaluation results in Figure 8 and figure 10 that the performance of our proposed model is close to that of 3D CNN. In order to make a detailed comparison, we put the information of the two in a table and the results are as follows.

According to TABLE 4, it shows the comparison of the 3D MCNN model using depth separable convolution with the 3D CNN model using full convolution. In TABLE 4, the results of both statistics indicate that, compared with 3D CNN, 3D MCNN is slightly inferior to 3D CNN in the measurement metrics of wave height and period for actual measurement. The intuitive data representation directly shows the difference in metrics R², R, MAE, which is within a range of less than 1% for height and period. For NRMSE, the difference of both is not more than 4.1%. For RMSE, height data is still relatively small, and the difference in period data is slightly larger, and so is index of agreement indicator. However, we observed that for computational overhead, 3D MCNN significantly saves tremendously on mult-adds, prediction time and parameters, greatly reducing the computational cost.

Due to the complexity of the marine environment, devices deployed on facilities such as mobile observation points are usually edge devices with limited resources, and their memory and computing power are limited. As the scale of the parameters in the model is larger, the demand for hardware resources is higher, which is not conducive to the deployment of the model on the observation equipment. In order to realize that the trained learning model can be well deployed on mobile devices to complete inference tasks and obtain reliable and faster inference results, it is the ideal method to select fewer model parameters. In addition, the lighter the model, the lower the equipment cost. Therefore, we selected 3D MCNN as our final best model.

V. CONCLUSION

In this paper, we develop a low-cost, high-accuracy detection scheme named iOceanSee for ocean wave parameters using 3D Mobile Convolutional Neural Network. iOceanSee is a light-weight implementation version of 3D Convolutional Neural Network specifically to detect wave period and height in the marine environment. Firstly, depth separable convolution is adopted to construct lightweight networks to reduce computational overhead. Then, a mobile terminal composed of low-cost measuring equipment and non-interference (except light) devices-an RGB camera is used to collect video data in real time on vessels named YUMING and XIANGYANGHONG81. Secondly, we extract features from both space and time dimensions, thus capturing the encoding motion information in multiple adjacent frames in the video using our model. Finally, we compare our scheme with other models. The experimental results show that this
method gets performance comparable to 3D CNN and obtains higher measurement accuracy than existing methods such as EOFWNN, FRULER and HM-CCD, for both height measurement and periodic measurement.

In future work, we intend to use the ocean data collected from other observation stations to pre-train our model to pursue of higher accuracy and lower cost. In addition, we will improve our architecture, such as the use of attention mechanisms, in hope of achieving other tasks in the marine field.

REFERENCES

[1] R. A. Arinaga and K. F. Cheung, “Atlas of global wave energy from 10 years of reanalysis and hindcast data,” Renew. Energy, vol. 39, no. 1, pp. 49–64, Mar. 2012.

[2] X. Yin, X. Zhao, and W. Zhang, “A novel hydro-kite like energy converter for harnessing both ocean wave and current energy,” Energy, vol. 158, pp. 1204–1212, Sep. 2018.

[3] T. Xia, H. Yu, R. Guo, and X. Liu, “Research on the field-modulated tubular linear generator with quasi-Halbach magnetization for ocean wave energy conversion,” IEEE Trans. Appl. Supercond., vol. 28, no. 3, pp. 1–5, Apr. 2018.

[4] E. Alexandre, L. Cuadra, J. C. Nieto-Borge, G. Candil-García, M. del Pino, and S. Salcedo-Sanz, “A hybrid genetic algorithm—Extreme learning machine approach for accurate significant wave height reconstruction,” Ocean Model., vol. 92, pp. 115–123, Aug. 2015.

[5] S. Salcedo-Sanz, J. C. Nieto Borge, L. Carro-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.

[6] Z. Ti, M. Zhang, L. Wu, S. Qin, K. Wei, and Y. Li, “Estimation of the significant wave height in the nearshore using prediction equations based on the response surface method,” Ocean Eng., vol. 153, pp. 143–153, Apr. 2018.

[7] X. Cheng, G. Li, R. Skulstad, S. Chen, H. P. Hildre, and H. Zhang, “Modeling and analysis of motion data from dynamically positioned vessels for sea state estimation,” in Proc. Int. Conf. Robot. Autom. (ICRA), May 2019, pp. 6644–6650.

[8] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[9] A. K. Bhunia, A. Konwer, A. K. Bhunia, A. Bhowmick, P. P. Roy, and U. Pal, “Script identification in natural scene image and video frames using an attention based convolutional-LSTM network,” Pattern Recognit., vol. 85, pp. 172–184, Jan. 2019.

[10] K. Brkić, T. Hrkać, Z. Kalafatić, and I. Sikirić, “Face, hairstyle and clothing colour de-identification in video sequences,” IET Signal Process., vol. 11, no. 9, pp. 1062–1068, Dec. 2017.

[11] E. E. Geertsema, R. D. Thijs, T. Gutter, B. Vledder, J. B. Arends, F. S. Leijten, G. H. Visser, and S. N. Kalitzin, “Automated video-based detection of nocturnal convulsive seizures in a residential care setting,” Epilepsia, vol. 59, pp. 53–60, Jun. 2018.

[12] M. Ye, J. Li, A. J. Ma, L. Zheng, and P. C. Yuen, “Dynamic graph matching for unsupervised video-based person re-identification,” IEEE Trans. Image Process., vol. 28, no. 6, pp. 2976–2990, Jun. 2019.

[13] D. Chen, Z. Yuan, J. Wang, B. Chen, G. Hua, and N. Zheng, “Exemplar-guided similarity learning on polynomial kernel feature map for person re-identification,” Int. J. Comput. Vis., vol. 123, no. 3, pp. 392–414, Jul. 2017.

[14] F.-C. Chen and M. R. Jahanshahi, “NN-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion,” IEEE Trans. Ind. Electron., vol. 65, no. 5, pp. 4392–4400, May 2018.

[15] Y. Ding, W. K. Wong, Z. Lai, and Z. Zhang, “Bilinear supervised hashing based on 2D image features,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 2, pp. 590–602, Feb. 2020.

[16] J. H. Bappy, C. Simons, N. Nataraj, B. S. Manjunath, and A. K. Roy-Chowdhury, “Hybrid LSTM and encoder–decoder architecture for detection of image forgeries,” IEEE Trans. Image Process., vol. 28, no. 7, pp. 3286–3300, Jul. 2019.

[17] S. Ji, W. Xu, M. Yang, and K. Yu, “3D convolutional neural networks for human action recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 221–231, Jan. 2013.

[18] J. Li, T. Wang, Y. Zhou, Z. Wang, and H. Snoussi, “Using Gabor filter in 3D convolutional neural networks for human action recognition,” in Proc. 36th Chin. Control Conf. (CCC), Jul. 2017, pp. 11139–11144.

[19] X. Ren, L. Xiang, D. Nie, Y. Shao, H. Zhang, D. Shen, and Q. Wang, “Interleaved 3D-CNNs for joint segmentation of small-volume structures in head and neck CT images,” Med. Phys., vol. 45, no. 5, pp. 2063–2075, May 2018.

[20] R. Ceschin, A. Zahner, W. Reynolds, J. Gaesser, G. Zuccoli, C. W. Lo, V. Gopalakrishnan, and A. Panigrahy, “A computational framework for the detection of subcortical brain dystasmaturation in neonatal MRI using 3D convolutional neural networks,” NeuroImage, vol. 178, pp. 183–197, Sep. 2018.

[21] Z. Zhang, P. Jaiswal, and R. Rai, “FeatureNet: Machining feature recognition based on 3D convolution neural network,” Comput.-Aided Des., vol. 101, pp. 12–22, Aug. 2018.

[22] J. Oh and K.-D. Suh, “Real-time forecasting of wave heights using EOF–wavelet–neural network hybrid model,” Ocean Eng., vol. 150, pp. 48–59, Feb. 2018.

[23] L. Cornejo-Bueno, P. Rodríguez-Mier, M. Mucientes, J. C. Nieto-Borge, and S. Salcedo-Sanz, “Significant wave height and energy flux estimation with a genetic fuzzy system for regression,” Ocean Eng., vol. 160, pp. 33–44, Jul. 2018.

[24] J. Yoon and D. Song, “Preliminary study for detecting of ocean wave parameters using CCD images,” J. Coastal Res., vol. 85, pp. 1371–1375, May 2018.

[25] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “MobileNets: Efficient convolutional neural networks for mobile vision applications,” Tech. Rep., 2017.
YUANYUAN ZHANG received the B.S. degree from the School of Electrical and Information Engineering, Jiangsu University of Technology, in 2017. She is currently pursuing the Ph.D. degree with the Merchant Marine College, Shanghai Maritime University. She participated in two National natural science foundation project, one is Study on the 3-D Dynamic Cooperative Localization Mechanism of Ocean Sensor Network based on Wave Shadowing Effect, and the other is Study on the Mechanism of Dynamic Topology and Temporal and Spatial Random Coverage of Wireless Sensor Networks in Maritime Search and Rescue. Her research interests include data processing and deep learning in wireless sensor networks.

JUN WANG (Senior Member, IEEE) is currently a Full Professor of computer engineering and the Director of the Computer Architecture and Storage Systems (CASS) Laboratory, University of Central Florida, Orlando, FL, USA. His group is investigating three U.S. National Science Foundation projects, one DARPA, and one NASA project. He has graduated 13 Ph.D. students who upon their graduations were employed by major U.S. IT corporations, such as Google, Microsoft, and so on. He has authored over 120 publications in premier journals, such as the IEEE TRANSACTIONS ON COMPUTERS and the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, and leading HPC and systems conferences, such as VLDB, HPDC, EuroSys, IPDPS, ICS, Middleware, and FAST. He has conducted extensive research in the areas of Computer Systems and High Performance Computing. His research interests include massive storage and file system in local, distributed, and parallel systems environment, big data, and big compute systems, data-intensive high performance, computing massive storage, and file system, I/O architecture. His group has secured multi-million dollars federal research fundings in last five years. He received the IEEE Transactions on Cloud Computing Editorial Excellence Award, in 2019. He is a General Executive Chair of the IEEE DASC/DataCom/P/Com/CyberSciTech 2017. He has co-chaired technical programs in numerous computer systems conferences, including the 2018 IEEE International Conference on High Performance Computing and Communications (HPCConf18), the 10th IEEE International Conference on Networking, Architecture, and Storage (NAS 2015), and the 1st International Workshop on Storage and I/O Virtualization, Performance, Energy, Evaluation and Dependability (SPEED 2008) held together with HPAC. He has been serving on the Editorial Board for the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS and the IEEE TRANSACTIONS ON CLOUD COMPUTING.

WEIJUN WANG received the M.S. degree from the Merchant Marine College, Shanghai Maritime University, in 2015, where he is currently pursuing the Ph.D. degree. He has participated in two National natural science foundation project, (1) Study on the Three-Dimensional Dynamic Cooperative Localization Mechanism of Ocean Sensor Network based on Wave Shadowing Effect and (2) Study on the Mechanism of Dynamic Topology and Temporal and Spatial Random Coverage of Wireless Sensor Networks in Maritime Search and Rescue. His research interests include machine learning and objects targeting in wireless sensor networks.

JINGFENG XIAN received the B.S. degree from the College of Transport and Communications, Shanghai Maritime University, in 2015. He is currently pursuing the Ph.D. degree with the Merchant Marine College, Shanghai Maritime University. He has participated in two National natural science foundation project, (1) The Three-Dimensional Dynamic Cooperative Localization Mechanism of Marine Sensor Networks Based on Wave Shadowing Effect Model and (2) Study on the Mechanism of Dynamic Topology and Temporal and Spatial Random Coverage of Wireless Sensor Network in Maritime Search and Rescue, respectively. He has published four articles indexed by SCIE in peer-reviewed journals. His research interests include data fusion and target detection in ocean wireless sensor networks.

JING CHEN received the bachelor’s degree from the China University of Petroleum, the master’s degree from the East China University of Science and Technology, and the Ph.D. degree in marine transportation engineering from Shanghai Maritime University. From 2018 to 2019, he was a Visiting Scholar with Shanghai Jiao Tong University. He is currently a Lecturer and a Master’s Tutor with the Merchant Marine College, Shanghai Maritime University. He has enthusiasm and research expertise in areas of ship attitude law and maritime navigation monitoring, and so on. He has published more than ten articles in peer-reviewed journals, such as the Journal of Coastal Research and the Journal of Ship Mechanics, most of which are indexed by SCI and EI.

XIANGYI ZOU received the M.S. degree from the Merchant Marine College, Shanghai Maritime University, in 2014. He currently works with the National Deep Sea Center to study the management of scientific research ships. His research interests include vibration and noise control of ships and objects targeting in wireless sensor networks.

PRASANT MOHAPATRA (Fellow, IEEE) received the Ph.D. degree from Penn State University, in 1993. From 2007 to 2013, he was the Department Chair of computer science. From 2016 to 2018, he was the Dean and the Vice-Provost of Graduate Studies. He has also held Visiting Scientist positions with Intel Corporation, Panasonic Technologies, the Institute of Infocomm Research (I2R), Singapore, and the National ICT Australia (NICTA). He is currently the Vice Chancellor for Research with the University of California at Davis, Davis, CA, USA. He is also a Professor with the Department of Computer Science. His research has been funded through grants from the National Science Foundation, the U.S. Department of Defense, the U.S. Army Research Labs, Intel Corporation, Siemens, Panasonic Technologies, Hewlett Packard, Raytheon, and EMC Corporation. He has published more than 350 papers in reputed conferences and journals on these topics. His research interests include wireless networks, mobile communications, cybersecurity, and the Internet protocols. He is a Fellow of AAAS. He was a recipient of the Distinguished Alumnus Award from the National Institute of Technology, Rourkela, India. He received the Outstanding Engineering Alumni Award, in 2008, the Outstanding Research Faculty Award from the College of Engineering at the University of California at Davis, and the HP Labs Innovation Awards, in 2011, 2012, and 2013.