Received October 1, 2020, accepted October 15, 2020, date of publication October 23, 2020, date of current version November 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033455

An Enhanced Extreme Learning Machine for Dissolved Oxygen Prediction in Wireless Sensor Networks

LIANG KUANG[1,2], PEI SHI[3,4], CHI HUA[1], BEIJING CHEN[2], AND HUI ZHU[5]

[1]School of IoT Engineering, Jiangsu Vocational College of Information Technology, Wuxi 214153, China
[2]School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China
[3]Binjiang College, Nanjing University of Information Science & Technology, Wuxi 214105, China
[4]Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, Wuxi 214081, China
[5]School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China

Corresponding author: Pei Shi (njxk_sp@sina.cn)

This work was supported in part by the High-Level Backbone Specialty Construction Project of Jiangsu Province under Grant 2017[17], in part by the Natural Science Foundation of Jiangsu Province under Grant BK20131097, in part by the Jiangsu Province Higher Vocational Education Industry Education Integration Platform Construction Project under Grant 2019[26], in part by the scientific research team in Jiangsu Vocational College of Information Technology, and in part by the Natural Science Research in Jiangsu Universities in 2020.

ABSTRACT Water quality monitoring using Wireless Sensor Networks (WSNs) is essential in aquaculture water quality management. In the field of water quality monitoring, dissolved oxygen (DO) is a key parameter, and its prediction can provide decision support for aquaculture production, thereby reducing farming risk. However, it is difficult to build a precise prediction model, and existing methods of DO prediction neglect the importance of analyzing DO content. To address this problem, this study proposes a hybrid DO prediction model, named KIG-ELM, which is composed of K-means, improved genetic algorithm (IGA), and extreme learning machine (ELM). This model is based on edge computing architecture, in which data acquisition, processing and dissolved oxygen prediction are distributed in sensing nodes, routing nodes and server respectively. Sensing technique and clustering operation are applied in the process of data acquisition and processing. Meanwhile, an optimized extreme learning machine is implemented for DO prediction. We evaluate the efficiency and accuracy of proposed prediction approach in a practical aquaculture on massive water quality data. Experimental results show that the hybrid model achieves significant prediction results and can meet the needs of practical production and management.

INDEX TERMS Sensor networks, dissolved oxygen prediction, edge computing, water quality monitoring, extreme learning machine.

I. INTRODUCTION

In recent years, the development of WSNs has promoted the progress of smart fishery. More and more smart devices have already been put in production of smart fishery. To reduce the data transfer volume and bandwidth caused by these smart devices, edge computing provides guidance and strategy [1]–[3]. As a special application of WSNs, the study of water quality monitoring has become an important tool for sustainable development of smart fishery [4]–[6]. Based on edge computing architecture [7]–[9], water quality monitoring using WSNs can monitor and analyze culture environment information and provide guidance for improving fishery efficiency in aquaculture.

Dissolved oxygen (DO) concentration is a critical water quality parameter in aquaculture, which needs to be monitored in real time in intensive aquaculture [10]. Generally, when DO content is less than 3 mg/L, it will has great impact on healthy fish [11], because the prolonged hypoxia will even lead fish to death from suffocate. Thus, it is critical to make DO prediction in the cultivation process, which can help managers making decision. However, the DO prediction is a complex process involving many factors, and it also has the characteristics of dynamic and nonlinearity [12]. Accurate and effective prediction of dissolved oxygen plays an important directional role in production, which can also produce

The associate editor coordinating the review of this manuscript and approving it for publication was Kun Wang.

enormous economic benefits. Hence, DO prediction becomes a prominent issue in water quality monitoring using WSNs.

In technical literature, DO prediction methods can be categorized into two types: single parameter prediction [13]–[16] and multi-parameter prediction [17]–[25]. In the former type, the prediction of DO just considers the change law of historical dissolved oxygen data. Although different methods are used in single parameter prediction, it is different for them to obtain extra information about DO content. Therefore, this kind of prediction is inaccurate and inflexible, especially for dynamic and complex prediction of DO. On the contrary, considering the correlation between DO and related environment parameters, these parameters are operated as input to forecast future DO content in the latter prediction type. Actually, when the weather changes suddenly, water will change due to meteorological effect. Meanwhile, DO content also changes with other water quality parameters, such as pH value and water temperature. Hence, this kind of prediction is reliable and feasible for DO prediction.

Unfortunately, multi-parameter prediction of DO have several drawbacks: (1) Redundant input data. Although there is correlation between DO content and related environment parameters, redundant input parameters also reduce the efficiency of prediction and don't help improving accuracy. (2) Neglect the importance of DO characteristic. Most multi-parameter prediction methods pay more attention to the related environment parameters, and neglect the importance of DO characteristics. Actually, there are many DO characteristics that can be utilized in forecasting DO, such as the dynamics of DO in daytime and nighttime [27]. However, very few works study this issue combining DO characteristics with environment parameters.

To tackle these limitations, in this paper we propose a new multi-parameter DO prediction method, named KIG-ELM, which is based on edge computing to improve prediction accuracy and efficiency. In order to grasp the characteristics of DO varied with weather condition, K-means method is adopted to cluster DO data and capture the characteristics of these nonlinear data according to the weather index. Extreme learning machine (ELM) is a very efficient and simple learning algorithm for overcoming DO prediction problems with low generalization, nonlinear samples, low efficiency and accuracy [28]. The improved genetic algorithm (IGA) is utilized to obtain the optimal initial parameters of ELM, avoiding premature convergence and chattering problems in the process of optimization. Thus, this hybrid prediction model consists of K-means method, IGA and ELM. KIG-ELM is verified and compared with the counterpart algorithms using real monitoring data in aquaculture tanks. The experimental results show that the performance of forecasting model is improved in efficiency and precision.

The rest of this paper is organized as follows: Section 2 presents a brief review on the previous works. The study area, data acquisition, custom weather index, and the construction of the hybrid prediction model are outlined in Section 3. Section 4 provides the steps of developing the hybrid dissolved oxygen prediction model are introduced. The accuracy and efficiency of the forecasting model are demonstrated in Section 5. Section 6 concludes this study and points out our future works.

II. RELATED WORKS

A. SINGLE PARAMETER PREDICTION

Single parameter methods forecast the future DO by capturing the changing trends of historical DO data. Due to the limitation of unstable prediction performance, few studies have used such methods. In [13], the back-propagated (BP) neural network is used to predict the changing trend of DO using historical monitoring data. Liu *et al.* [14] utilized the wavelet analysis method to de-noise and decompose the original DO sequences, and proposed the hybrid WA-CPSO-LSSVR for DO content forecasting. Although this study uses the multi-scale decomposition method to help get high precision, it is also time-consuming. Rahman *et al.* [15] presented an approach that can predict dissolved oxygen multiple time stamps for long term prediction. And this approach requires a higher window of past observations, which may limit the number of time steps ahead prediction it can make. Olyaie *et al.* [16] utilized three different AI methods to predict the DO content in Delaware River located at Trenton, USA. In these methods, the number of input can be single or multiple. However, the comparison of estimation accuracies of various intelligence models indicates that the accuracies of multiple parameters prediction are higher obviously.

B. MULTI-PARAMETER PREDICTION

Multi-parameter methods forecast the future dissolved oxygen using several related parameters as the input, and the dissolved oxygen content as output. These parameters include water quality parameters and meteorological parameters. Many studies have been conducted at present [17]–[27].

Liu *et al.* [17] systematically discussed and compared the application of the attention-based RNN method in dissolved oxygen prediction using two water quality parameters, two soil parameters and six meteorological parameters. Liu *et al.* [18] proposed a prediction model based on support vector regression (SVR) to solve the aquaculture water quality prediction in modern intensive river crab aquaculture management problem. Yu *et al.* [19] developed a new hybrid dissolved oxygen content forecasting model based on the radial basis function neural networks (RBFNN) data fusion method and least squares support vector machine (LSSVM) with an optimal improved particle swarm optimization (IPSO). Wu *et al.* [20] established a new model of dissolved oxygen prediction based on sliding window, particle swarm optimization (PSO) and BP neural network. A dissolved oxygen prediction model based on fuzzy neural networks is proposed in [21]. And genetic algorithm was utilized to optimized fuzzy neural network. Heddam and Kisi [22] attempted to estimate dissolved oxygen concentration using

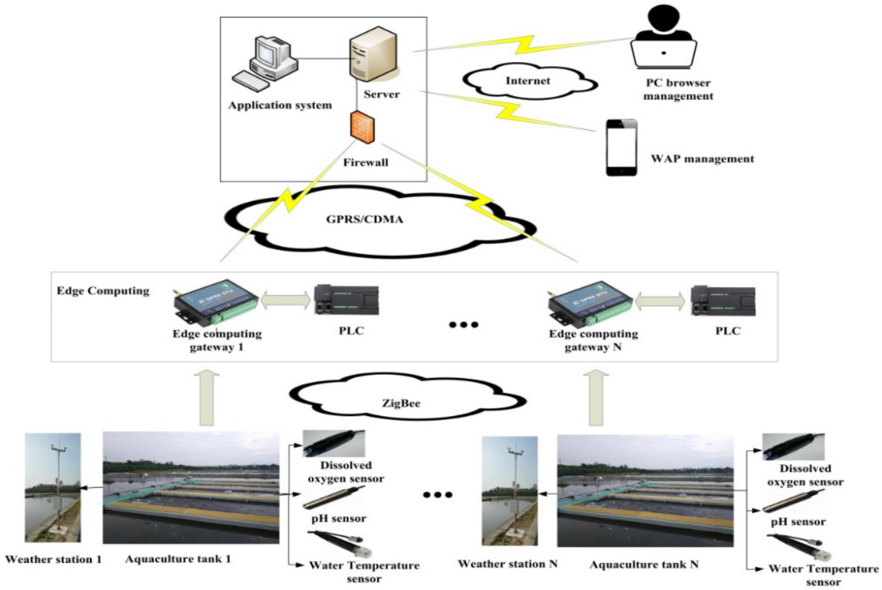

FIGURE 1. System architecture diagram.

four different ELM models without water quality variables. The components of the Gregorian calendar are used as input in these ELM models.

Although these studies have considered the impact of related parameters, they still face some prominent problems, such as lacking efficient processing of DO data and analyzing the changing characteristics of DO content. In [23], the ensemble empirical mode decomposition (EEMD) is adopted to capture the multi-scale features of DO in the proposed hybrid model. Although this study uses the multi-scale decomposition method to help get high precision, it is also time-consuming. Huan *et al.* [24] applied the EEMD in the proposed EEMD-LSSVM prediction model to decompose the DO time series into a group of relatively stable subsequences. This prediction model is good, but there is one problem. If the added white noise and iteration times are not appropriate, the false component will appear after decomposition. Shi *et al.* [25] adopted K-medoids to group the dataset into different clusters according to its characteristics in CSELM dissolved oxygen prediction model, but there exists redundant input of CSELM. Cao *et al.* [26] presented a prediction of dissolved oxygen in pond culture based on K-means clustering and Gated Recurrent Unit (GRU) neural network. In [27], discrete wavelet transforms (DWT) with different wavelet functions are compared in de-noising diel, daytime and nighttime dynamics of DO. Based on de-noising results, multiple non-linear regression (MNLR) models are adopted to predict DO. These works just focus on how to decompose the DO or optimize prediction model whereas leaving a gap between the characteristics of DO that varies with the weather condition in daytime and nighttime. It is necessary to study a new multi-parameter algorithm which can plug this gap and meet the needs in real world aquaculture production.

In this paper, our aim is to develop a DO prediction algorithm based on edge computing architecture in aquaculture, which can capture the characteristics of DO and maximize the prediction accuracy and efficiency.

III. MATERIALS AND METHODS

In this section, we present our materials and methods of prediction. We first present the study area and the architecture of data acquisition system. Then we introduce the weather index variable and utilize it to realize the operation of similar time slot clustering. We finally present the prediction method and optimization procedure.

A. STUDY AREA

The experimental base of this study is located in Wuxi city, Jiangsu province (E 120.18° N 31.34°). There are four tanks (about 9×3 m^2 for each) built in the aquaculture pond (about $110\times\ 45\ \times\ 1.5$m^3). There are 2,000 black basses cultured in each tank with the initial size about 3 cm. Circulating waters cultivation technology is used in this high-density tank. Meanwhile, the micro-pore aeration and airlift water push device are used for aeration.

B. DATA ACQUISITION

In this study, a remote monitoring system platform, which is developed by the freshwater fisheries research center, is used to collect water quality data and weather station data in intensive aquaculture. The system architecture diagram is shown in Fig.1.There are two aquaculture monitoring areas distributed in Wuxi city and Changshu city. We can obtain data from these aquaculture bases including dissolved oxygen, pH, water temperature, humidity, temperature, atmospheric pressure, carbon dioxide, illumination intensity, photosynthetically active radiation, radiance, wind speed and direction through the online monitoring system. Water quality sensors are placed in 0.5 meters underwater and meteorological sensors are integrated in the weather station. All data are collected from sensor nodes and transferred to the cloud server through routing node. Meanwhile, data are processed in routing node and stored in the cloud server. When the dissolved oxygen predicted value is lower than threshold value, the routing node will sent the urgent message to the control center. This helps users make the response strategy in advance and drive the bottom controller.

C. WEATHER INDEX

There are many weather factors related to the dissolved oxygen, including temperature, humidity, atmospheric pressure, wind speed and wind direction, etc. Dissolved oxygen concentration changes periodically during the day and night and varies greatly under different weather conditions. In order to capture the change rule under similar weather conditions, a custom weather index variable is used to characterize the weather conditions at different times.

Factor analysis (FA) is a multivariate statistical analysis method [28]. The basic idea is to classify the multi-variables according to the correlation level, and extract various variables as common factors. Then, the score function of common factor is obtained through the cumulative variance contribution rate of common factor and the variance of common factor. Thus establish the comprehensive evaluation index.

Let $X = [X_1, X_2, \ldots, X_i, \ldots X_p]$ be a set of p objects, m common factors F_j $(j = 1, 2, \ldots, m)(m<p)$ are extracted from X. $e_i (i = 1, 2, \ldots, p)$ is the other factors. And the ith index X_i can be expressed with factor analysis method as (1).

$$X_i = a_{i1}F_1 + a_{i2}F_2 + a_{i3}F_3 + \ldots + a_{in}F_m + e_i \quad (1)$$

where a_{ij} denotes the component matrix, which reflects the dependence of X_i on common factor F_j.

Based on (1), we calculate the score coefficient of i th common factor as (2).

$$F_j = \sum b_{ji}x_i \quad (i = 1, 2, \ldots, p; j = 1, 2, \ldots, m) \quad (2)$$

where b_{ji} is the score coefficient of ith index in jth common factor, x_i represents the data after normalization.

D. SIMILAR TIME SLOT CLUSTERING BASED ON WEATHER INDEX

To avoid the problem of low accuracy and efficiency in prediction caused by the difference between characteristics of the data stream, we use the weather index to cluster the historical samples at different time slots firstly. Then, the dissolved oxygen prediction model is constructed in each sub-cluster.

The K-means method is a widely used clustering method [29], where K represents the clustering number of data streams, and means represents the clustering center of cluster. Euclidean distance is used to measure the similarity between different elements. Error square is used to evaluate the clustering effect. The detail process of standard K-means can be summarized as follows [30].

Step 1: Randomly initialize the K cluster centroids in the given K objects.

Step 2: Assign each object to the group with the closet centroid. Use Euclidean distance to measure the minimum distance between the data object and each cluster centroid.

Step 3: Recalculate the cluster centroid vector using (3).

$$c_j = \frac{1}{n_j} \sum_{\forall\, data_p \in S_j} data_p \quad (3)$$

where c_j is the centroid vector of the cluster j, n_j is the number of data vectors in cluster j, S_j denotes the subset of data vectors from cluster j, and $data_p$ represents the p_{th} data vector.

Step 4: Repeat step 2 until these centroids do not change any more in the predefined number of iteration or a maximum number of iteration has been reached.

In this paper, the samples in one day are divided into two data sets of day and night. Each data set contains 72 samples, and each weather index value evaluates a data set. Thus, the matrix of each data set is shown as

$$f_i = [f_{1i}, f_{2i}, \ldots, f_{ti}]^T \quad (4)$$

where f_i is the ith data set, and t is the monitoring time.

The Angle Cosine is used to replace Euclidean distance and calculate the similarity between data set and the center. The clustering results of K-means are obtained based on the similarity of weather index. The similarity calculation formulas of Angle Cosine can be defined as

$$D_{f_i f_j} = \frac{1}{T} \sum_{t=1}^{T} \frac{f_{ti} f_{tj}}{\sqrt{f_{ii}^2 f_{tj}^2}} \quad (5)$$

where f_{ti} and f_{tj} reprensent the values of sample f_i and sample f_j at time t, respectively, $T = (1, 2, \ldots, t, \ldots, 72)$.

E. IMPROVED GENETIC ALGORITHM

The Genetic algorithm (GA) is a random and parallel search optimization method based on natural selection and natural genetic mechanism [31].Through the genetic operations of selection, crossover and mutation, the population can finish the generational evolution until the terminating conditions of evolution are met. In traditional GA, mutation is operated based on crossover. The crossover and mutation operations utilize the randomness of chaotic sequence, and both operations are performed separately [32]. In order to avoid premature convergence in optimization, we propose the IGA based on new crossover and mutation operations.

1) CROSS OPERATION

Different from the traditional single-point crossover, the crossover operation is based on the principle of ''door-to-account pairing''. The parent individuals are sorted according to the fitness function values. So the small fitness value is paired with the small one, large fitness value is large. We use the randomness of chaotic sequence to determine the location of crossing, thus completing the operation of crossing terms.

Take the logistics chaotic sequence as an example, if we pair the genes on chromosomes $(O_1, O_2)\{O_1 = o_1^1, o_2^1, \ldots, o_p^1$, $O_2 = o_1^2, o_2^2, \ldots, o_p^2\}$, it is necessary to use the logistics chaotic sequence $x(m+1) = 4x(m)[1-x(m)]$ to generate a positive integer a between $2\sim\ p-1$. Then take a as the crossing point to finish the single crossing of genes on chromosomes (O_1, O_2). Finally, we get the new chromosomes (O_1', O_2'), $O_1' = o_1^1, o_2^1, \ldots, o_a^2, \ldots, o_p^1, O_2' = o_1^2, o_2^2, \ldots, o_a^1, \ldots, o_p^2$.

2) MUTATION OPERATION

In this paper, two positive integers c and d between $2\sim p-1$ are randomly obtained according to given mutation rate. Then we mutate the genes at the corresponding positions on chromosomes and use chaotic sequences to replace genes at the positions of c and d with new gene values. Thus, we obtain new chromosomes.

The IGA algorithm keeps the mutation operation and the crossover operation apart, so that the two operations are performed independently and in parallel. In IGA, chaos sequence is introduced to determine the crossing point. The low modification of single point cross operation is utilized to weaken and avoid the chattering problem in GA optimization. The process of multi-gene mutation in chromosome is completed by using chaotic sequences to avoid premature convergence.

F. EXTREME LEARNING MACHINE

The ELM is a feed-forward neural network learning algorithm with nice global searching capability [33]. Once the parameters are set, there is no need to adjust in training. Compared with other machine learning algorithms, ELM has the advantages of high learning efficiency and excellent generalization performance, etc.

Given a data set (x_i, y_i), $i = 1,2,\ldots N$, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{iu}]^{\mathrm{T}} \in R^u$ is the ith sample, $y_i = [y_{i1}, y_{i2}, \ldots, y_{im}]^{\mathrm{T}} \in R^m$ is the actual tag of the ith sample. The ELM neural network with L hidden layer nodes and activation function $g(x)$ are mathematically modeled as follows:

$$\sum_{j=1}^{L} \beta_j g\left(w_j x_i + b_j\right) = y_i, \quad i = 1, 2, \ldots, N \tag{6}$$

where $w_j = [w_{j1}, w_{j2}, \ldots, w_{ju}]^{\mathrm{T}}$ denotes the weight vector between the jth hidden node and the input node, b_j is the threshold value of the jth hidden layer node, $\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jm}]^T$ is the weight vector between the jth hidden node and the output node.

G. EXTREME LEARNING MACHINE

In ELM, the weights and bias vectors are randomly obtained. When the weights and bias vectors are special values, the hidden layer nodes will fail, thus reducing the accuracy in training and testing [34]. In order to reduce prediction error, it is necessary to adopt an intelligent algorithm to optimize the parameters. The main steps can be expressed as:

Step 1: Create a new ELM network and initialize the population. In GA, we need to initialize population firstly, and then encode the weight and bias of ELM neural network. Finally, the structure of ELM network is determined with input variable $X = [x_1, x_2, \ldots, x_u]^{\mathrm{T}}$ and the output variable $Y = [y_1, y_2, \ldots, y_m]^{\mathrm{T}}$.

Step 2: Determine the fitness function. In ELM neural network, set the fitness function of network, the size of initial population and the maximum evolution generations *maxgen*. The fitness function can be expressed as

$$fun = \frac{1}{n}\sum_{t}^{n}\left(y_i(t) - y_i'(t)\right)^2 \tag{7}$$

where n is the number of samples, y_i(t) denotes the real observe value at time t, y_i'(t) is the predicted value at time t.

Step 3: Compute the optimal fitness value. According to individual fitness value, selection, improved crossover and mutation operations are performed on individuals randomly to obtain the global optimal fitness value $Fitness_{best}$.

Step 4: Get the optimal weight and bias. The optimal fitness value is used to obtain the corresponding population individually. Finally, we obtain the optimal input weight matrix a_{best} and bias b_{best} of ELM.

Step 5: Determine the hidden layer output matrix. In ELM neural network, we calculate the output matrix H of hidden layer by a_{best} and b_{best}.

$$\begin{aligned} &H\left(a_1, \cdots, a_L, b_1, \cdots, b_L, x_1, \cdots, x_m\right) \\ &= \begin{bmatrix} g\left(a_1 \cdot x_1 + b_1\right) & \cdots & g\left(a_1 \cdot x_1 + b_L\right) \\ \cdots & \cdots & \cdots \\ g\left(a_1 \cdot x_m + b_1\right) & \cdots & g\left(a_1 \cdot x_m + b_L\right) \end{bmatrix}_{m\times L} \end{aligned} \tag{8}$$

Step 6: Compute the connecting weight between hidden layer and output layer. The least square algorithm is used to calculate the output weight β between the hidden layer node and the output node. Based on the output weight, we can get the predicted value and terminate algorithm. The calculation of output weight is shown as

$$\tilde{\beta} = H^{+}Y \tag{9}$$

where H^{+} denotes the Moore-Penrose inverse of hidden layer output matrix H, $H^{+} = (H^T H)^{-1} H^T$.

IV. THE HYBRID PREDICTION MODEL BASED ON K-MEANS AND IGA-ELM

The dissolved oxygen data streams have different features in day and night. K-means method can explore the periodic change patterns of data streams, and ELM is suitable for capturing data characteristics in nonlinear system with high learning speed. Thus, this paper constructs KIG-ELM prediction model to forecast the dissolved oxygen content in bass culture. The implementation process of prediction is shown in Fig.2. The prediction model consists of three parts: data preprocessing, similar time slots clustering and modeling. Data preprocessing includes data standardization and data reduction. After data preprocessing, the input variables of

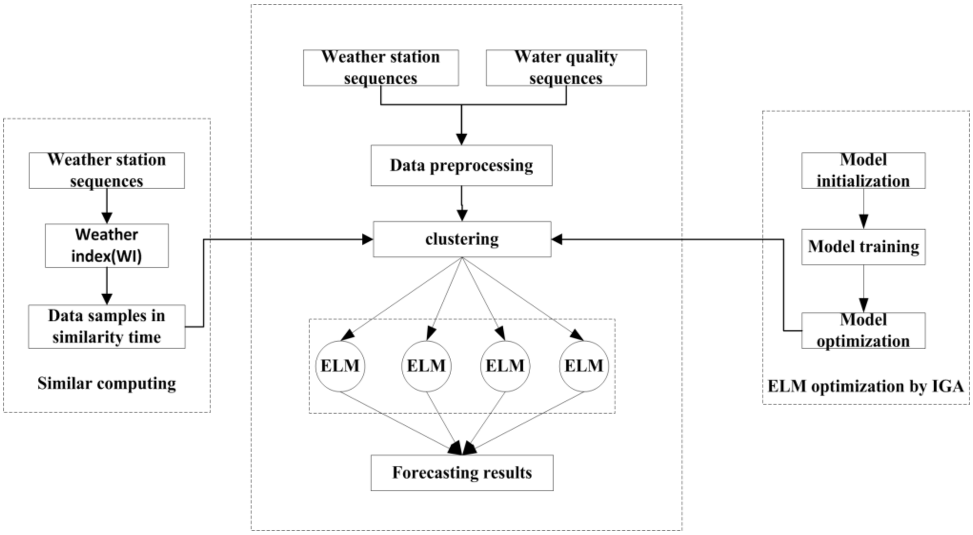

FIGURE 2. Implementation process of dissolved oxygen content prediction.

prediction model are determined. The FA method is used to calculate the weather index. Combined with K-means method, we cluster the data set of similar time slots. Based on these steps, we construct the dissolved oxygen prediction model and obtain the predict values. The implementation process of dissolved oxygen prediction can be described as follows.

Step 1: Data Preprocessing. To eliminate difference between sensor data of different measurement standards, Z-score normalization is performed on the raw data [35]. Meanwhile, there are so many factors which influence on dissolved oxygen. It is necessary to reduce the dimension of factors including pH, water temperature, humidity, temperature, atmospheric pressure, carbon dioxide, illumination intensity, photosynthetically active radiation, radiance, wind speed and direction. We utilize the PCA method to analyze these indicators and determine the input variables of prediction model. Then extract five common factors depending on whether eigenvalue is greater than 1. The results of PCA analysis are shown in Table 1.

From Table 1, we determine the pH, water temperature, temperature, carbon dioxide, illumination intensity, photosynthetically active radiation, radiance, wind speed and direction to represent the five common factors and use these indicators as the input variables of prediction mode.

Step 2: Clustering based on similar time slots. K-means is used to cluster historical samples based on similar time slots. Since dissolved oxygen is affected by multiple weather factors, the custom weather index is used to evaluate weather conditions in aquaculture.

Angle Cosine has the advantages of measuring the difference in different directions, and correcting the problem of non-uniform metric. Considering the periodically change trend of dissolved oxygen in day and night, we choose the Angle Cosine instead of Euclidean distance in K-means and finish clustering data sets of similar time slots. The results

TABLE 1. The results of PCA analysis.

Indictors	Components				
	1	2	3	4	5
dissolved oxygen	.127	.966	-.008	.013	.020
water temperature	-.002	-.076	.914	-.070	-.008
pH	.000	.947	-.019	.117	.168
carbon dioxide	.164	.098	.161	.124	.878
atmospheric pressure	.011	.332	-.529	-.278	.585
temperature	.459	.105	.837	-.009	.155
humidity	-.596	-.346	-.481	.209	-.043
illumination intensity	.984	.039	.105	-.016	.054
radiance	.983	.030	.087	-.016	.046
photosynthetically active radiation	.984	.040	.079	-.015	.070
wind speed	.268	.176	.135	-.723	-.380
wind direction	.093	.246	.020	.877	-.129

TABLE 2. DB values change with k.

Cluster number k	2	3	4	5	6
DB	0.3352	0.2765	0.5471	1.3234	0.8712

of K-means clustering are shown in Table 2. Davies-Bouldin (DB) index is applied to evaluate clustering effect [36]. From Table 2, it is clear that $DB = 0.2765$ is minimum, so $K = 3$ is best.

$$V_{DB}(k) = \frac{\sum_{i=1}^{k} \max_{j,j\neq i} \left\{ \frac{S_i+S_j}{d_{ij}} \right\}}{k} \tag{10}$$

where $S_i = \frac{1}{n_i} \sum_{x\in C_i} ||x - z_i||$ is compactness in cluster C_i, $d_{ij} - ||z_i - z_j||$ represents the dispersion between cluster C_i and cluster C_j.

Step 3: Modeling. Set the parameters of IGA and ELM neural network respectively in three clusters. IGA is used to

TABLE 3. Neural network structures of three clusters.

Serial number of clusters	Number of samples	Number of training samples	Number of testing samples	Structure of neural network	RMSE
1	1008	864	144	9-37-1	0.3062
2	720	648	72	9-27-1	0.2131
3	288	216	72	9-21-1	0.1888

optimize the parameters of each ELM model. Finally, build the hybrid prediction model.

1) IGA parameters. Set the initial population size of IGA algorithm to 10 and the number of iterations to 50. According to the principle of minimum mean squared error for multiple times, the crossover probability is 0.1 and the mutation probability is 0.1.

2) ELM parameters. Set the number of input nodes to 9 and the number of output nodes to 1. The number of hidden layer nodes in each cluster is set to 37, 27 and 21 respectively by using trial and error approach. Thereby obtain the structure of each ELM neural network.

Step 4: Forecasting results. For each new group of dissolved oxygen prediction task, assign new data stream to the corresponding cluster and obtain the forecasting value through KIG-ELM prediction model.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we present the experimental results and discussions of the dissolved oxygen prediction. We first introduce the experimental data. Then we present the clustering results of experimental data sets. When obtaining the best clustering results, we realize the prediction of dissolved oxygen in different clusters. We finally compare the prediction performance and analyze the prediction results.

A. DATA SETS

In this paper, all data are collected every 10min from the aquaculture tank through the water quality monitoring system. All weather data and water quality data (totally 2,016 data sets) are grouped into 28 streams (each for a day or night) from July 1, 2019 to July 14, 2019. And 1,728 data sets (in 12d) are used for training, 288 data assets (in 2d) used for testing. We choose the 9 indexes as the input variables of prediction model, including pH, water temperature, temperature, carbon dioxide, illumination intensity, photosynthetically active radiation, radiance, wind speed and direction. Dissolved oxygen is the output. All of the experiments are implemented by MATLAB 2014a and run on a PC with 3.4GHz Core(TM) processor, 16.0GB memory, and Microsoft Windows 10.

B. CLUSTERING RESULTS

The data sets in 14d (28 days or nights) can be clustered into three clusters by K-means. The structure of IGA-ELM neural network in each cluster is shown in Table 3. Every cluster has different size and structure. The largest cluster contains 828 samples, while the smallest cluster contains only 228 samples. In Table 3, we can find that all night samples are classified into cluster 1. And day samples are classified into two clusters. Cluster results indicate that the different night weather conditions have little effect on dissolved oxygen, while the effects are big with different day weather conditions.

C. PERFORMANCE CRITERIA

It is necessary to use some widely recognized indicators to evaluate the performance of the proposed prediction model. Thus, we utilize different statistical performance evaluation criteria including the mean absolute percentage error (MAPE), the root mean square error (RMSE), Nash Sutcliffe efficiency Coefficient (NSC) [37] (Benyahyaetal *et al.*, 2007) and run time. These indicators are calculated as follows

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

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{12}$$

$$NSC = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{13}$$

where n is the number of samples, y_i and $\hat{y}_i$ are the observed value and predicted value respectively, $\bar{y}$ is the average of all observed values.

D. SIMULATION RESULTS AND ANALYSIS

After K-means clustering, we realize DO prediction of KIG-ELM. The prediction result of the proposed model is shown in Fig.3. To verify effect of K-means clustering and IGA optimizing operations, K means-GA-ELM (GA optimized ELM network after K-means clustering), K means-ELM (ELM neural network after K-means clustering) [38], K means-LSSVM [39], IGA-ELM (improved GA optimized ELM network) and ELM models are constructed for comparison. The total training and testing samples of these contrast models are the same with KIG-ELM. The prediction errors of the six models are shown in Fig.4.

From Fig.3, it is clear that the KIG-ELM prediction models can complete the dissolved oxygen prediction well. The prediction trend of the KIG-ELM is consistent with real value and prediction results are closer to real values with small fluctuations. Fig.4 represents the prediction errors of the above six models. It is obvious that the KIG-ELM is superior to the other models. Its prediction performance is stable. The prediction error of KIG-ELM is less than 0.2 and lower

TABLE 4. Performances comparison of six predictive models.

Comparisons	Our hybrid model	*K*means-GA-ELM	*K*means–ELM	*K*means-LSSVM	IGA-ELM	ELM
NSC	0.9294	0.91	0.8776	0.8256	0.8359	0.7917
MAPE	0.0386	0.046	0.0518	0.0648	0.0605	0.0766
RMSE	0.2591	0.3131	0.3568	0.4201	0.4033	0.4849
Time(s)	1.5375	1.8073	1.0452	1.7949	7.8129	0.1282

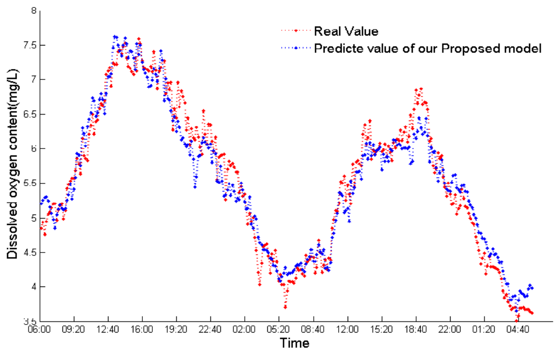

FIGURE 3. Prediction results of proposed prediction model.

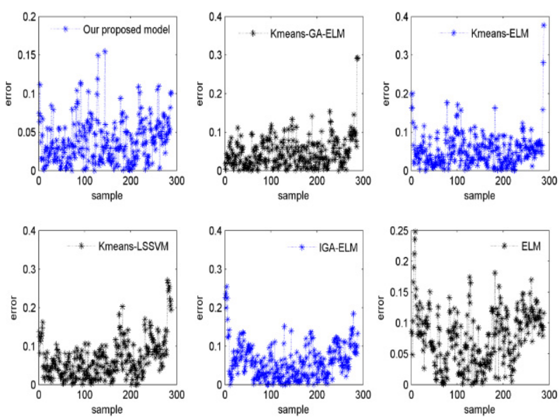

FIGURE 4. Prediction errors of six models.

than the other counterpart models. Meanwhile, the prediction errors of the other counterpart models are unstable.

To verify the superiority of KIG-ELM, we compare the performances of these six models. All contrast models realize with K-means algorithm performed the clustering based on similar time slots. Under the same preconditions, the prediction performance of each model is shown in Table 4.

From Table 4, the prediction accuracy of the six models has reached more than 90% (all MAPEs were less than 0.1). The MAPE, RMSE values of the hybrid model are 0.0386 and 0.2591, respectively. Compared with the *K* means-GA-ELM, *K* means–ELM, *K* means-LSSVM and ELM, the MAPE value is decreased by 16.09%, 25.48%, 40.43% and 49.61% respectively. Meanwhile, the RMSE value is decreased by 17.25%, 27.38%, 38.32% and 46.57% respectively. Meanwhile, KIG-ELM shortens the time by 0.2698s and 0.2574s, when compared with *K* mean-GA-ELM and *K* mean-LSSVM. Compared with IGA-ELM, MAPE, RMSE values of KIG-ELM decreased by 36.20% and 35.76%. This is because dissolved oxygen is susceptible to multiple weather factors such as temperature and illumination. The clustering operation based on weather index can group data streams well, which have similar characteristics. Clustering helps eliminate information interference between data streams with different characteristics, thus improving the accuracy of prediction model. However, the time of IGA-ELM is five times longer than KIG-ELM. The running time of KIG-ELM is slightly longer than *K* mean-ELM. Moreover, KIG-ELM had the best NSC value among these models. The experimental results prove that the improved genetic algorithm could avoid the premature convergence of genetic algorithm and improve the efficiency effectively.

From the prediction results of IGA-ELM and ELM neural networks, we can conclude that parameters optimization process of IGA algorithm helps ELM obtain the best parameters. It dramatically improves the prediction accuracy, but also consumes a little time. The prediction results of *K* mean-ELM and *K* mean-LSSVM show that ELM neural network is more suitable to be applied in forecasting dissolved oxygen in high-density aquaculture than LSSVM. In conclusion, the proposed KIG-ELM has higher prediction accuracy and efficiency than the counterpart models for these indicators.

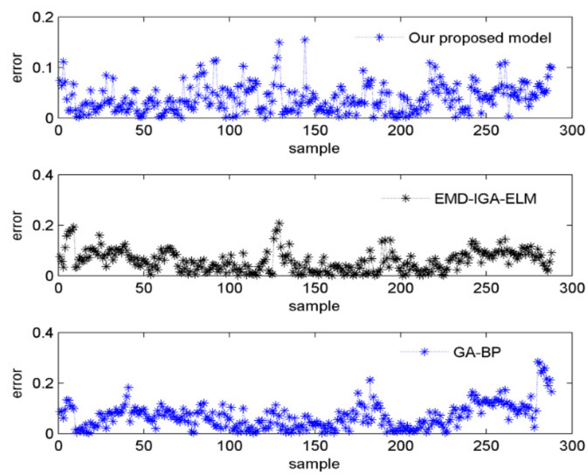

FIGURE 5. Prediction error comparisons with two existing models.

To verify the superiority of KIG-ELM, the existing algorithms EMD-IGA-ELM [40] and GA-BP [41] are selected for comparison. In GA-BP, the structure of BP is 9-44-1, the number of iteration is 1000, and the goal is 0.1. In EMD-IGA-ELM, the number of IMF components is 8. The parameters setting of IGA and ELM are the same as

TABLE 5. Performances comparison with two existing models.

Indexes	Our hybrid model	EMD-IGA-ELM[36]	GA-BP[37]
NSC	0.9294	0.8216	0.7269
MAPE	0.0386	0.059	0.0673
RMSE	0.2591	0.3936	0.4484
Time(s)	1.5375	10.251	2.237

Section 3. Fig.5 represents the prediction error comparisons with these two existing models. It is clear that the error of our proposed KIG-ELM is lower than EMD-IGA-ELM and GA-BP. The prediction error of KIG-ELM is less than 0.2, but the prediction errors of EMD-IGA-ELM and GA-BP both are between approximately 0 and 0.4. To analysis the performance of KIG-ELM, we calculate different prediction indicators of three models in Table 5.

From Table 5, the performance of proposed model is superior to EMD-IGA-ELM and GA-BP. The MAPE of KIG-ELM decreases by 34.58% and 42.65% respectively, when compared with EMD-IGA-ELM and GA-BP. RMSE values of KIG-ELM reduces by 34.17% and 42.22% respectively. Meanwhile, the NSC values of KIG-ELM significantly outperform the other two models. These results prove that KIG-ELM could capture the characteristics of data stream and improve the reliability and accuracy of the prediction model by clustering operation. Besides, the run time of EMD-IGA-ELM is about 7 times of KIG-ELM, and 0.6995s shorter than GA-BP. Although multi-scale decomposition of EMD can reduce the interference at different scales information, it also consumes time. The multi-scale decomposition achieves high accuracy at the expense of time.

In summary, KIG-ELM can solve the problem of random parameters in ELM and explore the characteristics of dissolved oxygen data streams. The combination of K-means, IGA and ELM can effectively improve the accuracy of prediction model, and overcome the problem of low accuracy of single model. We can get the status and trend of dissolved oxygen and make scientific strategy based on this reasonable and reliable method in intensive aquaculture.

VI. CONCLUSION

This paper constructs a new multi-parameter prediction method KIG-ELM to forecast the DO content in aquaculture based on edge computing architecture. In order to capture the characteristics of dissolved oxygen data stream, this paper self-defines the variable weather index and used *K*-means to cluster data streams in similar time slots. The chaotic sequence is introduced into IGA algorithm, which can weaken and avoid the chattering problem in the process of optimization. Therefore, it is used to search the optimal weight and threshold of ELM neural network. The experimental results show that optimized ELM neural network can realize the forecasting process quickly. The combination of ELM, *K*-means and IGA can improve prediction accuracy significantly. Compared with the counterpart algorithms, the proposed hybrid model is more suitable for predicting DO in intensive aquaculture.

Although the proposed model has achieved excellent prediction results in this study, there is still room for improvement in the future. Firstly, additional proper search techniques for obtaining optimum parameters could be combined with ELM to predict DO content. Secondly, prediction accuracy usually decreases rapidly during sunrises and sunsets. It is meaningful to find a proper method to enhance accuracy in above time slots. Finally, how to construct comprehensive indexes as inputs of prediction model is significant. All of above are worth exploring problems in future research.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] X. Liu, T. Qiu, B. Dai, L. Yang, A. Liu, and J. Wang, "Swarm intelligence-based rendezvous selection via edge computing for mobile sensor networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9471–9480, Oct. 2020, doi: 10.1109/JIOT.2020.2973401.

[2] X. He, K. Wang, H. Huang, T. Miyazaki, Y. Wang, and S. Guo, "Green resource allocation based on deep reinforcement learning in content-centric IoT," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 3, pp. 781–796, Jul. 2020.

[3] X. He, K. Wang, and W. Xu, "QoE-driven content-centric caching with deep reinforcement learning in edge-enabled IoT," *IEEE Comput. Intell. Mag.*, vol. 14, no. 4, pp. 12–20, Nov. 2019.

[4] S. Huang, J. Gui, T. Wang, and X. Li, "Joint mobile vehicle-UAV scheme for secure data collection in a smart city," *Ann. Telecommun.*, pp. 1–22, Aug. 2020, doi: 10.1007/s12243-020-00798-9.

[5] Y. Ren, T. Wang, S. Zhang, and J. Zhang, "An intelligent big data collection technology based on micro mobile data centers for crowdsensing vehicular sensor network," *Pers. Ubiquitous Comput.*, pp. 1–17, Aug. 2020, doi: 10.1007/s00779-020-01440-0.

[6] B. Jiang, G. Huang, T. Wang, J. Gui, and X. Zhu, "Trust based energy efficient data collection with unmanned aerial vehicle in edge network," *Trans. Emerg. Telecommun. Technol.*, p. e3942, Mar. 2020, doi: 10.1002/ett.3942.

[7] H. Lu, X. He, M. Du, X. Ruan, Y. Sun, and K. Wang, "Edge QoE: Computation offloading with deep reinforcement learning for Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9255–9265, Oct. 2020, doi: 10.1109/JIOT.2020.2981557.

[8] C. Xu, K. Wang, P. Li, R. Xia, S. Guo, and M. Guo, "Renewable energy-aware big data analytics in geo-distributed data centers with reinforcement learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 205–215, Jan. 2020.

[9] K. Wang, Y. Wang, Y. Sun, S. Guo, and J. Wu, "Green industrial Internet of Things architecture: An energy-efficient perspective," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 48–54, Dec. 2016.

[10] A. Boukerche, H. A. B. Oliveira, E. F. Nakamura, and A. A. F. Loureiro, "Secure localization algorithms for wireless sensor networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 96–101, Apr. 2008.

[11] Z.-S. Hou, H.-S. Wen, J.-F. Li, F. He, Y. Li, and X. Qi, "Environmental hypoxia causes growth retardation, osteoclast differentiation and calcium dyshomeostasis in juvenile rainbow trout (Oncorhynchus mykiss)," *Sci. Total Environ.*, vol. 705, Feb. 2020, Art. no. 135272.

[12] X. Ta and Y. Wei, "Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network," *Comput. Electron. Agricult.*, vol. 145, pp. 302–310, Feb. 2018.

[13] J. X. Yan, L. J. Yu, W. W. Mao, and S. Q. Cao, "Study on prediction model of dissolved oxygen about water quality monitoring system based on BP neural network," *Adv. Mater. Res.*, vols. 912–914, pp. 1407–1411, Apr. 2014.

[14] S. Liu, L. Xu, Y. Jiang, D. Li, Y. Chen, and Z. Li, "A hybrid WA–CPSO-LSSVR model for dissolved oxygen content prediction in crab culture," *Eng. Appl. Artif. Intell.*, vol. 29, pp. 114–124, Mar. 2014.

[15] A. Rahman, J. Dabrowski, and J. McCulloch, "Dissolved oxygen prediction in prawn ponds from a group of one step predictors," *Inf. Process. Agricult.*, vol. 7, no. 2, pp. 307–317, Jun. 2020.
[16] E. Olyaie, H. Zare Abyaneh, and A. Danandeh Mehr, "A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware river," *Geosci. Frontiers*, vol. 8, no. 3, pp. 517–527, May 2017.
[17] Y. Liu, Q. Zhang, L. Song, and Y. Chen, "Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction," *Comput. Electron. Agricult.*, vol. 165, Oct. 2019, Art. no. 104964.
[18] S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, and Y. Wei, "A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction," *Math. Comput. Model.*, vol. 58, nos. 3–4, pp. 458–465, Aug. 2013.
[19] H. Yu, Y. Chen, S. Hassan, and D. Li, "Dissolved oxygen content prediction in crab culture using a hybrid intelligent method," *Sci. Rep.*, vol. 6, no. 1, Jun. 2016, Art. no. 27292, doi: 10.1038/srep27292.
[20] J. Wu, Z. Li, L. Zhu, G. Li, B. Niu, and F. Peng, "Optimized BP neural network for dissolved oxygen prediction," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 596–601, 2018.
[21] Q. Ren, L. Zhang, Y. Wei, and D. Li, "A method for predicting dissolved oxygen in aquaculture water in an aquaponics system," *Comput. Electron. Agricult.*, vol. 151, pp. 384–391, Aug. 2018.
[22] S. Heddam and O. Kisi, "Extreme learning machines: A new approach for modeling dissolved oxygen (DO) concentration with and without water quality variables as predictors," *Environ. Sci. Pollut. Res.*, vol. 24, no. 20, pp. 16702–16724, May 2017.
[23] C. Li, Z. Li, J. Wu, L. Zhu, and J. Yue, "A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features," *Inf. Process. Agricult.*, vol. 5, no. 1, pp. 11–20, Mar. 2018.
[24] J. Huan, W. Cao, and Y. Qin, "Prediction of dissolved oxygen in aquaculture based on EEMD and LSSVM optimized by the Bayesian evidence framework," *Comput. Electron. Agricult.*, vol. 150, pp. 257–265, Jul. 2018.
[25] P. Shi, G. Li, Y. Yuan, G. Huang, and L. Kuang, "Prediction of dissolved oxygen content in aquaculture using clustering-based softplus extreme learning machine," *Comput. Electron. Agricult.*, vol. 157, pp. 329–338, Feb. 2019.
[26] X. Cao, Y. Liu, J. Wang, C. Liu, and Q. Duan, "Prediction of dissolved oxygen in pond culture water based on K-means clustering and gated recurrent unit neural network," *Aquacultural Eng.*, vol. 91, Nov. 2020, Art. no. 102122, doi: 10.1016/j.aquaeng.2020.102122.
[27] F. Evrendilek and N. Karakaya, "Regression model-based predictions of diel, diurnal and nocturnal dissolved oxygen dynamics after wavelet denoising of noisy time series," *Phys. A, Stat. Mech. Appl.*, vol. 404, pp. 8–15, Jun. 2014.
[28] J. Huang, J. Liu, K. Wang, Z. Yang, and X. Liu, "Classification and identification of molecules through factor analysis method based on terahertz spectroscopy," *Spectrochimica Acta A, Mol. Biomol. Spectrosc.*, vol. 198, pp. 198–203, Jun. 2018.
[29] M. Capó, A. Pérez, and J. A. Lozano, "An efficient approximation to the K-means clustering for massive data," *Knowl.-Based Syst.*, vol. 117, pp. 56–69, Feb. 2017.
[30] A. Karami and M. Guerrero-Zapata, "A fuzzy anomaly detection system based on hybrid PSO-kmeans algorithm in content-centric networks," *Neurocomputing*, vol. 149, pp. 1253–1269, Feb. 2015.
[31] H. B. Othman and M. Tagina, "Guided genetic algorithm: A new template concept," *Procedia Comput. Sci.*, vol. 60, pp. 1101–1108, Jan. 2015.
[32] M. Yousefi, M. Yousefi, R. P. M. Ferreira, J. H. Kim, and F. S. Fogliatto, "Chaotic genetic algorithm and AdaBoost ensemble metamodeling approach for optimum resource planning in emergency departments," *Artif. Intell. Med.*, vol. 84, pp. 23–33, Jan. 2018.
[33] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[34] M. Luo, C. Li, X. Zhang, R. Li, and X. An, "Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings," *ISA Trans.*, vol. 65, pp. 556–566, Nov. 2016.
[35] V. R. Apsingekar and P. L. De Leon, "Speaker verification score normalization using speaker model clusters," *Speech Commun.*, vol. 53, no. 1, pp. 110–118, Jan. 2011.
[36] S. Hatamikia, K. Maghooli, and A. Nasrabadi, "The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals," *J. Med. Signals Sensors*, vol. 4, no. 3, p. 194, 2014.
[37] D. Ömer Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Eng. Appl. Artif. Intell.*, vol. 23, no. 4, pp. 586–594, Jun. 2010.
[38] Q. He, X. Jin, C. Du, F. Zhuang, and Z. Shi, "Clustering in extreme learning machine feature space," *Neurocomputing*, vol. 128, pp. 88–95, Mar. 2014.
[39] R. Mall and J. A. K. Suykens, "Very sparse LSSVM reductions for large-scale data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 5, pp. 1086–1097, May 2015.
[40] P. Y. Shi Yuan, L. Kuang, G. Li, and H. Zhang, "Water temperature prediction in pond aquaculture based on EMD-IGA-SELM neural network," *Trans. Chin. Soc. Agricult. Machinery*, vol. 49, no. 11, pp. 312–319, Nov. 2018.
[41] D. Zheng, Z.-D. Qian, Y. Liu, and C.-B. Liu, "Prediction and sensitivity analysis of long-term skid resistance of epoxy asphalt mixture based on GA-BP neural network," *Construction Building Mater.*, vol. 158, pp. 614–623, Jan. 2018.

LIANG KUANG received the master's degree in system analysis and integration from the Nanjing University of Information Science & Technology, China, in 2014, where he is currently pursuing the Ph.D. degree in information and communication engineering. He is also a Lecturer with the School of IoT Engineering, Jiangsu Vocational College of Information Technology. His research interests include distributed artificial intelligence and deep learning systems.

PEI SHI received the master's degree in system analysis and integration from the Nanjing University of Information Science & Technology, China, in 2014, and the Ph.D. degree in control science and engineering from Jiangnan University, China, in 2020. She is currently a Lecturer with the Binjiang College, Nanjing University of Information Science & Technology, Wuxi, China. Her research interests include wireless sensor networks, edge computing, and artificial intelligence.

CHI HUA received the master's degree in computer technology from Shanghai Jiao Tong University, China, in 2009. He is currently pursuing the Ph.D. degree in computer science and techniques from the Nanjing University of Aeronautics and Astronautics, China. He is also a Professor with the School of IoT Engineering, Jiangsu Vocational College of Information Technology. His research interests include big data analysis, artificial intelligence, and software engineering.

BEIJING CHEN received the Ph.D. degree in computer science from Southeast University, Nanjing, China, in 2011. He is currently a Professor with the School of Computer and Software, Nanjing University of Information Science & Technology, China. His research interests include color image processing, image forensics, image watermarking, and pattern recognition.

HUI ZHU received the M.S. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China, in 1989 and 1992, respectively. He is currently a Professor with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. He has published over 20 articles and holds six patents. His research interests include VR technology, smarter learning, and artificial intelligence. He is a member of the Education Information Technology Standards Committee of the Ministry of Education of China.