Intelligent Identification System for Farming Water Quality Based on Deep Learning

Lirong He[1], Jing Su[1], Baogui Tang[2], Hancheng Huang[1], Liang Zeng[1], Chuangwei Wen[1]

[1]Faculty of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang, Guangdong, China

[2]College of Fisheries, Guangdong Ocean University, Zhanjiang, Guangdong, China

Abstract: Traditional aquaculture tends to judge water quality based on the experience of the farmer, failing to make effective use of farming data and lacking an intuitive and easy way to observe farming data. To settle the problem, we have developed a farming water quality intelligence identification system using deep learning and popular front and back-end technologies. The system provides digital management for farming data, makes analysis and prediction of water quality data and water color images using deep learning techniques, and visualizes the output of predictions, which helps to aid aquaculture, such as reducing farming risk, increasing the scale of farming and improving farming efficiency.

1. Introduction

Currently, farmers do not pay enough attention to the testing of water quality, and only rely on their farming experience to judge the water quality, and cannot effectively determine some harmful substances in the water, which leads to the phenomenon of mass mortality or slow growth of aquaculture organisms[1]. Therefore, it is essential to measure and keep track of the condition of aquaculture water quality in a timely manner.

At present, the research by national and international experts on smart aquaculture is mainly focused on the classification and prediction of aquaculture water quality[2,3,4,5], and there are few systems that can predict water quality conditions and visualize water quality data by recognizing water color images. To this end, we have developed a farming water quality intelligence system using deep learning and popular front and back-end technologies. Digital management for farming data and intelligent recognition for farming water color images by using deep learning techniques provides an intuitive and rapid understanding of water quality conditions and helps to improve farming efficiency and increase the farming scale.

2. Main System Functions

The system has five main functions: data management, model training, water quality prediction, water quality warning and data visualization.

The key factors affecting water quality are temperature, pH, dissolved oxygen and ammonia nitrogen [1,4]. Similarly, algae play an important role in aquaculture [6]. The main algae in water are green algae, diatoms and nudibranchs, and predicting the type of algae in water is beneficial for farmers to have a better understanding of water quality. This system predicts and classifies these key factors in water body.

2.1 Data Management

Data management includes user data management, image data management, water quality data management. User data management: the user can manage and maintain the user data, including viewing user information, changing the password and setting thresholds, which will be used for water quality warning. Image data management: users can add, delete, change and check image and the corresponding water quality data. Water quality data management: users can add a single or

batch import multiple water quality data, to be able to query all or specify the conditions of the historical water quality data, and can modify, delete or batch export these historical data.

2.2 Model Training

The system includes two types of model training: water quality model training and water color image model training. For water quality model training, the system employs Long Short Term Memory network (LSTM) for water quality prediction. The system trains the corresponding water quality model according to the water quality factors and model types selected by users. For water color image model training, because Convolutional Neural Networks (CNN) has an excellent performance in the image feature extraction, the system uses CNN as the water color image prediction model. Users can directly train the corresponding CNN model by selecting a certain water quality factor.

2.3 Water Quality Prediction

Water quality prediction includes water quality data prediction, classification and prediction of water color images. Water quality data prediction uses the first three days of water quality data to predict the fourth day of water quality data, and visualizes the predicted results for output. Water color image classification identifies which algae are present in the water color image uploaded by the user. Water color image prediction can directly identify the image to predict the situation of each water quality factor in the image, for example, the user uploads a water color image, choose to predict the pH, the system will predict the image corresponding to the pH data.

These water quality prediction functions can be a good aid to aquaculture, reducing the cost of aquaculture for the user, lowering the threshold of aquaculture and improving economic efficiency.

2.4 Water Quality Warning

When the user predicts the water quality on the system, if the predicted result exceeds the warning threshold set by the user, the system will immediately send alarm information to the user's email.

2.5 Data Visualization

Data visualization [7,8] refers to the data into the form of graphical output, for observing the data more clearly and intuitively. In the data management module, the system will be historical water quality data drawn as a line graph. Through the line graph, the user can clearly and intuitively see the data trend. It is easier to observe the pattern of data changes.

After the training of the model, the system in the same coordinate system with the actual value of the line drawn in red, with the predicted value of the line drawn in blue. The degree of overlap between the red and blue lines can visually reflect the accuracy of the model prediction results, helping users to judge the training results and train the ideal prediction model.

The water quality data visualization output is a line of two colors. The first three days of historical data are shown in black and the fourth day of predicted data is shown in blue. The different color markers help users to see the future trend of the data more quickly.

3. System Design

3.1 Design Patterns

The system is developed using the B/S model[8] with a separation of the front and back ends, simplifying the development, maintenance and use of the system. According to the MVC (model-view-controller) design pattern[9], we divide the system into a control layer, a business layer and a persistence layer.

3.2 Design of the Database

The system uses MySQL database and Redis database together to achieve data persistence storage. Redis database provides caching function[10], to help reduce the pressure of MySQL

database read and write. The system sets up seven database tables, storing user data, image data, water quality data and warning thresholds. As shown in Figure 1.
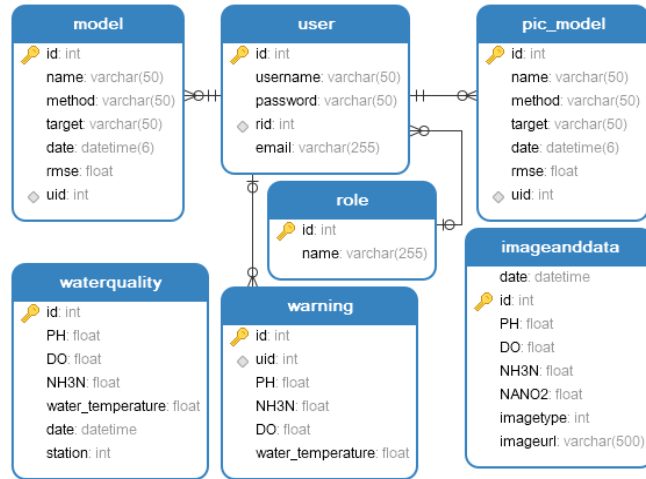

Fig.1 Database Tables

3.3 System Modules

The system is developed in a modular way with the main modules being the front-end, Java back-end and Python sub-back end.

3.4 Front End

The framework used for the front-end part is Vue framework[11], which can bring good human-computer interaction experience to the users. The data visualization part uses the open source visualization library *echarts*, and the front and back-end interaction technology is Axios.

3.5 Java Backend

Java backend is based on SpringBoot framework [11], integrated Spring MVC, Spring Data JPA, Spring Security framework to build the system backend services to ensure smooth deployment and operation of backend services.

3.6 Python Sub-Backends

Python sub-backend using keras, tensorflow , pytorch and other frameworks to quickly build neural network models, using Django framework to train models and predict water quality data. Django has good scalability and can efficiently and quickly build the backend modules of this system, which can pass data to each other with the Java backend[12].

4. Data Analysis

4.1 Data Normalization

We use min-max normalization to process the data, scaling the data between 0 and 1 according to Eq. 1.

$$x^* = (x - x_{max}) / (x_{max} - x_{min}). \tag{1}$$

Data normalization eliminates the inability to analyze multiple data together due to the dimensionality of the data, and speeds up the convergence of the learning algorithm, making the model more accurate[13].

4.2 Model Training Results

We trained the LSTM model to predict pH, and the model training results are shown in Figure 2. The red dash in the figure represents the actual value and the blue dash represents the predicted value. We can see that the blue line and the red line are close to overlapping, indicating that the model is well trained.

Fig.2 Model Training Results

5. Conclusion

We have designed a farming water quality intelligent identification system by using deep learning technology to predict water quality data, which is a more scientific and efficient approach to water quality data management. The system can identify and analyze water quality conditions through the water color image. The data visualization function allows the user to check the changes in the data more intuitively. The system can be a good aid to improve the efficiency and economic benefits of aquaculture.

Acknowledgment

References

[1] Li Zhili, The importance of water quality testing in aquaculture and improvement measures, Agricultural Science and Technology (Lower), 2020, No. 10, pp. 88

[2] Shi Pei, Yuan Yongming, Kuang Liang, Li Guanghui, Zhang Hongyan, Pond culture water temperature prediction method based on EMD-IGA-SELM, Journal of Agricultural Machinery, 2018,49(11)

[3] He Xuxiang, Xiao Richao, Zhang Long, Bit Yaoguang, Design of a water quality prediction and early warning system for factory aquaculture, Anhui Agricultural Science, 2017, Vol. 45(22), pp. 177-179 and 220

[4] Xu Fengdu, Hu Liang and, Zhou Zeqin, Zhong Nana, Wu Siping, Cai Ting, Peng Liyuan, Zhu Shengshan, Research progress on the effects of algal changes on aquaculture, Aquaculture, 2015, No. 01, pp. 48-52

[5] Sun Dongwei, Liu Bin, Data visualization research and technology implementation, Industry and Technology Forum. 2020, Vol. 19(19), pp. 40-42

[6] Qin Mingran, Design and implementation of data visualization system based on B/S model, Beijing University of Posts and Telecommunications, 2020.

[7] Liao Bin, Analysis of MVC pattern in JAVA-based development of Web applications, Electronic Technology and Software Engineering, 2020, (21), pp. 49-50

[8] Qin Li, Ai Qing, Research on page caching based on Redis technology, Science and Technology

Wind. 2021,(12), pp. 74-76

[9] Tian Haiqing, Design and implementation of a shared operation management platform based on SpringBoot and Vue framework, Shandong University, 2020.

[10] Wang Yang, Jiang Xintong, Research on the design of MVC framework under Python and Django, Computers and Information Technology, 2021, No. 01, pp. 55-57 and 63

[11] Weijun Deng, Research on computer graphics description algorithms based on deep learning, Microcomputer Applications, 2021, No. 02, pp. 53-55

[12] Wang Xin, Wu Ji, Liu Chao, Yang Haiyan, Du Yanli, Niu Wensheng, Fault time series prediction based on LSTM recurrent neural network, Journal of Beijing University of Aeronautics and Astronautics 2018,44(04),772-784

[13] Singh Dalwinder; Singh Birmohan, Investigating the impact of data normalization on classification performance, [J] Applied Soft ComputingVolume 97, Issue PB. 2020.