Journal of Agriculture and Food Research 9 (2022) 100357

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Development of smart aquaculture farm management system using IoT and AI-based surrogate models

Min-Chie Chiu [a], Wei-Mon Yan [b], Showkat Ahmad Bhat [c,*], Nen-Fu Huang [d,**]

[a] Department of Mechanical and Materials Engineering, Tatung University, Taipei, Taiwan, ROC
[b] Department of Energy and Refrigerating Air-Conditioning Engineering, National Taipei University of Technology, Taipei, 10608, Taiwan, ROC
[c] College of Electrical Engineering and Computer Science, National Tsing Hua University, Hsinchu, 300044, Taiwan, ROC
[d] Department of Computer Science, National Tsing Hua University, Hsinchu, 300044, Taiwan, ROC

ARTICLE INFO

Keywords:
Smart fish farming
Deep neural networks
Internet of things
Aquaculture
Big data
Bayesian optimization

ABSTRACT

Due to low labor participation by young adults and an aging agricultural population, Taiwan and the rest of the world are facing labor shortages in agriculture, which will affect aquaculture production. The proposed system is intended primarily for solving the problems faced by the aquaculture farming sector in Taiwan by designing a smart IoT-based fish monitoring and control system equipped with different IoT devices to enable real-time data collection; so that fishpond water-quality conditions and other system parameters can be readily monitored, adjusted, and assessed remotely. To predict the growth of the California Bass fish, this study also develops a deep learning model (DL) that correlates the different parameters of the smart aquaculture system. Bayesian optimization-based hyper-parameter tuning was employed to find the optimal DL model configuration to produce accurate predictions on the given experimental data set. The optimal model produces an R^2 value of 0.94 and a mean square error of 0.0015, demonstrating the applicability of the model to predict the desired output. Based on the results of the experiments, the DL model can be incorporated into the autonomous feeding system, reducing the amount of leftover feed. Thus, aquaculture based on the artificial intelligence of things (AIOT) can assist fish farmers in intelligently controlling and managing different fishpond equipment remotely and assist aquaculture operators in performing professional aquaculture, lowering the industry's entry barrier, and promoting aquaculture.

1. Introduction

Internet of Things (IoT) and other advancements have contributed to the innovation of conventional agricultural practices over the past few years [1]. The labor required by conventional fish farming techniques increases the cost of production because workers are needed to supervise the farms. Aquaculture production will be challenged by a manpower shortage since agriculture workers' average ages are rising in various parts of the world. To resolve this issue, major changes are required; automated operations should be managed remotely. With the Internet of Things, labor costs can be reduced and productivity can be boosted. The technology will have a significant impact on monitoring and analytics in the future [2]. Organizations can integrate countless IoT devices to collect huge quantities of data that can be stored and analyzed [3].

Technological advancements in the realm of IoT can help it realize its full potential; intelligent devices and analytics frameworks are being produced every day throughout the world. Overproduction has resulted from modern aquaculture practices, leading to recurrent outbreaks of fish illnesses and worse seafood quality. A solution is required to eliminate the danger of asphyxia caused by poor oxygen levels, water contamination, parasites, or disease transmission. The data collected by the sensor can help determine and solve challenges in the fish farming industry, such as enhancing fish health. The advantages of integrating IoT therefore in the industry are enormous. It aids in effective monitoring by providing a vast coverage of data from numerous locations, allowing for real-time remedial steps to be implemented. Data redundancy and scalability are key characteristics of cloud platforms used with IoT. Data collected over time may be used by artificial intelligence (AI) and machine learning (ML) technologies to create massive predictive models that can be used for correct decision-making, process

* Corresponding author.
** Corresponding author.
E-mail addresses: showkatbhat1994@gmail.com (S.A. Bhat), nfhuang@cs.nthu.edu.tw (N.-F. Huang).

https://doi.org/10.1016/j.jafr.2022.100357
Received 26 June 2022; Received in revised form 4 August 2022; Accepted 7 August 2022
Available online 17 August 2022
2666-1543/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

automation, and timely warnings. Using AIoT, aquaculture is incorporating more and more artificial intelligence to create smart aquaculture systems.

Aquaculture systems that are AIoT-based can simultaneously monitor water quality, microclimate, and provide warning functions while enhancing early warnings and response times, and their equipment can be linked with an intelligent electric box to control water wheels, pumps, feed machines, and other equipment. Moreover, the data gathered from such an AIoT system can contribute to more accurate and intelligent aquaculture through the use of sensors and feed data. This will enable labor savings, water quality stability, energy savings, and accurate feeding. As well as reducing catastrophe and damage risks, it can make aquaculture more productive and maximize its benefits.

Furthermore, overall productivity and nutrition costs are also strongly influenced by the degree of fish feeding technique. As a result, considerable economic advantages can be realized by optimizing the feeding mechanism [4]. Furthermore, as the demand for high-quality aquatic products grows, an increasing emphasis is being placed on fish health throughout aquaculture. In addition to consuming oxygen, unconsumed feed contributes to the production of ammonia and other poisonous compounds, affecting the fish's health and development [5,6]. There are a few key factors that need to be considered when determining the creature's actual feeding requirement. This can be extremely challenging with automation machines. To resolve this issue, a smart monitoring system for fish farming with artificial intelligence frameworks that enable forecasting and determining feeding tactics for animals is needed.

This research proposes a smart aquaculture monitoring system to address existing concerns regarding aquaculture through the design of a prototype of a smart fish farm system based on the internet of things and artificial intelligence. The proposed system is equipped with various sensors controlled by the Arduino Mega2560 with an integrated Wi-Fi module to assist in making different management decisions. The objectives of this study are as follows: 1) To enable real-time data collection; so that fishpond water-quality conditions and other system parameters can be readily monitored, adjusted, and assessed remotely; 2) a DL model is developed to correlate the different parameters of the fish-feeding system and predict the growth of the California Bass fish; 3) to enable the determination of the feed quantity on water using a turbidity sensor, the temperature of the water, underwater oxygen level, and pH value to reduce production costs and enhance fish production by having real-time monitoring of different system parameters; 4) a correlation study was undertaken to determine the impact of input and output parameters that have a greater impact on the prediction effectiveness of the model; 5) a mobile application has been developed to monitor and control the system remotely. As a result, an aquaculture sector of the highest quality and competitiveness could be formed. The economic rewards may influence a change in fish farming. Smart management systems can assist farmers in gathering real-time data, achieving optimal fishery performance, and rejuvenating the metabolic processes associated with finite fisheries.

2. Related works

By using an IoT-based automation solution, it is possible to constantly monitor pH, temperature, and dissolved oxygen in a fish pond. Using the data gathered by the system can result in more optimal resource use and profit maximization, resulting in long-term analyses and informed decisions. The researchers [7,8] developed and deployed an aquaculture real-time monitoring platform using a wireless sensor network (WSN) that monitored water-quality metrics, dissolved oxygen concentration, pH, and level of water. The authors of [9] proposed energy-saving solutions for aquaculture that use wireless sensor networks (WSN). A real-time fish farming surveillance system was designed and deployed by Ref. [10] which employed ZigBee and the GPRS protocol for connectivity among sensor devices and the server-side, therefore boosting transmission reliability. Idachaba et al. [11] created a pond controller that uses suitable sensors for monitoring the pond quality of the water and can be operated remotely through CCTV. Cario et al. [12] developed an effective underwater communication network for long-term environmental surveillance in fish farming. Hu, Wu-Chih et al. [13] developed a deep learning computer vision-based fish feeding system. It determines the feed quantity by determining the waves caused by the fish-eating feed. Table 1 summarizes some of the other recent available literature on the topic.

Table 1
Aquaculture monitoring systems.

References	Description
[1]	An automated fish feeding system was designed based on schedules and levels of need based on pH and TDS monitoring in aquaponics. Through an Android application, pH and TDS are monitored and fish feeding is automated.
[6]	To overcome inefficiency in artificial feeding control of grass carp, an adaptive neural fuzzy inference system is proposed that uses water quality parameters to provide automatic feeding decisions.
[14]	Discussed how the Amazon WEB Services (AWS) can be used for the aquaculture monitoring system by using MQTT protocol for communication with AWS IoT core.
[15]	Fuzzy logic control is used in this paper to design a pH control system in an aquaponics system. Using two pH sensors, one in the aquarium and another in the hydroponics system, it measures pH levels in both.
[16]	An aquarium water quality measuring system for ammonia and pH levels has been created useful for measuring the water quality for fish cultivation.
[17]	Implemented an Internet of Things-based water quality monitoring system. With this system, Ph, Temperature, and Turbidity are monitored in real-time from the Web. In order to display the measurement data on the Web, a database is used to process the data from the microcontroller.

Although some smart fish farm management systems have been proposed in literature, limited work have been done in order to integrate the factors that accurately determine the behavior, growth, water quality, feed to be released by the feeding system. In addition, none of the aforementioned solutions can analyze and interpret real-time data because they only monitor the system in real-time. To have a real-time control on the system parameters, smart data analytics should be integrated with the edge devices. As a result, the gap addressed in this paper is relevant for the study.

The proposed system in this study collects the data about the changing trend for different parameters and manages the operations of various devices of aquarium based on the measured values to preserve the stability of the various fishpond's parameters and thus enhance the production. Growth metrics are significantly affected by stocking density [18]. Data collection, analysis, and processing of water temperature, electrical conductivity, water level, pH, turbidity, and dissolved oxygen [19,20] are crucial when it comes to determining fish growth; therefore, data on these parameters should be collected, analyzed, and processed.

3. Methodology and materials

3.1. Experimental setup

To condense the manpower for the management of the aquaculture environment, automation devices like water quality monitoring and controlling system, and smart fish management systems are compulsory. The AIoT system for smart aquaculture management was designed via a system model. Fig. 1 represents the AIoT-based fish pond system architecture. The proposed smart California Bass fish pond is equipped with multiple sensors such as pH sensor, temperature sensor, dissolved oxygen sensor, and turbidity sensor connected with controller Arduino Mega2560 with integrated Wi-Fi module, multiple actuators such as a heater, limit switch, water pump, agitator, windproofing device, and smart feeding device, and IPCAM for real-time monitoring. All these

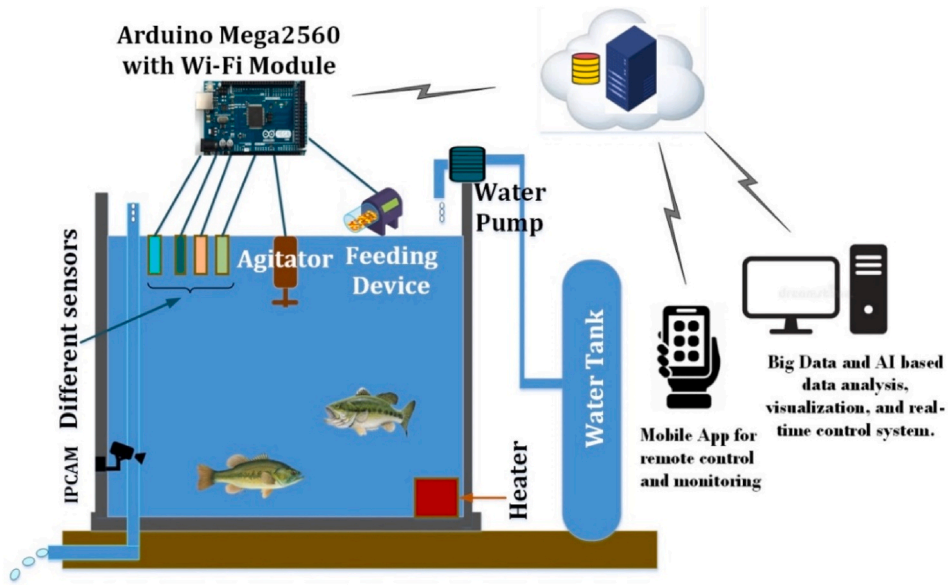


Fig. 1. An IoT-based intelligent fish pond.

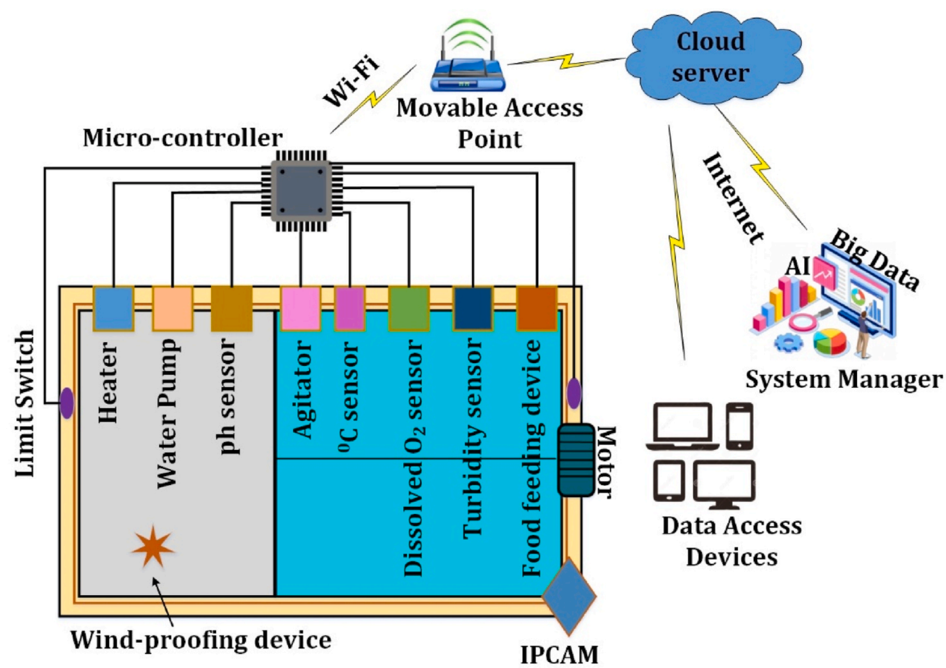


Fig. 2. A fish pond equipped with automatic control of multiple sensors, multiple actuators, and IPCAM.

sensors are used to collect the data of their respective parameters and transmit it to the cloud server using Wi-Fi communication protocol for further analysis. The data is then collected from the cloud server and analyzed at the server system using AI and big data techniques to extract the features and their impact on the system's productivity. A mobile application has been developed to monitor and control the system remotely.

To avoid the heat convection effect caused by a cold wind, a wind-proofing device driven by a C.W. and C·C.W. motor is established. The wind-proofing device will be extended and the heater will be triggered when the water temperature detected by the thermal sensor is below the targeted temperature threshold. As indicated in Fig. 2, the agitator installed within the pond will be started up when the dissolved oxygen value detected by the DO sensor is lower than the targeted value. Furthermore, the water pump will be activated if either the ph or turbidity reading exceeds the set threshold. Fig. 3 shows the water quality control flow diagram for a fish pond. Finally, the system will be equipped with an IP CAM underwater to monitor the growth of aquatic creatures and the baiting condition in compliance with water quality criteria to accomplish intelligent feeding.

A Wi-Fi module is connected with the smart fish aquarium to gather its environmental data and operate switches for system automation. Wi-Fi is based on the 802.11 radio standard, which allows data to be sent across limited distances at high frequencies. According to the type, 802.11 runs at 2.4 GHz or 5 GHz. It enables the system for high-speed data transmission and reception. Using the wireless Wi-Fi protocol, four kinds of water quality parameters (dissolved oxygen value, pH value, temperature, and turbidity value) detected from the near port

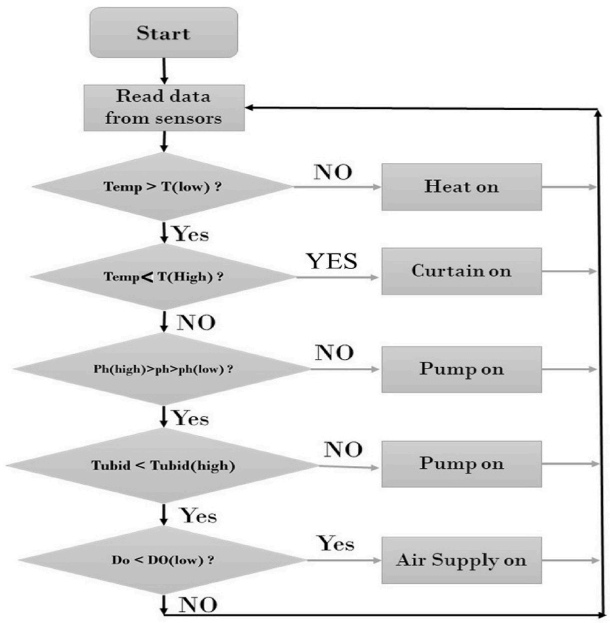


Fig. 3. The algorithm of the water quality controlling process for the fish pond.

(fish pond) will be sent to the cloud via the internet and wirelessly transmitted to the remote port (user's PC or cell phone). Here, an Arduino Mega2560 based on ATMega2560 is adopted as the microcontroller. Data communication enables remote monitoring and control of the different system parameters. The system thus permits the farm owner to collect the data at the cloud server and perform data analysis using the DL model.

3.2. Data collection

An automatic control fish pond equipped with multiple sensors, multiple actuators, a Wi-Fi protocol for data communication, and IPCAM is developed and used to study the effect of different variables on the growth of the fish. Currently, we are using a cloud server as our data collection center. Data is collected through the IoT system, transferred to the cloud for AI calculations, and received by the feedback system (Fig. 4). The data gathered from the feeding system, fish behavior monitoring, and autogiro AI computation is stored in the cloud to speed up AI data analysis. In this way, an effective AIoT smart aquaculture management system is achieved. Thus, the dissolved oxygen (ppm), pH, temperature (^{o}C), turbidity (NTU), bait quantity (grain/week), and length increment (cm/week) data were collected for straight 52 weeks. To search for appropriate conditions of water quality and bait quantity that can be helpful in fish growth, the correlation between the water quality, the bait quantity, and the increment of fish length will be explored. We have developed a DL model to establish the correlation between input parameters such as dissolved oxygen (ppm), pH, temperature (^{o}C), turbidity (NTU), bait quantity (grain/week), and an output parameter (length increment (cm/week)). The abstract of the fish pond's environmental parameters, the quantity of bait, and the corresponding length variation of fish are listed in Table 2. Table 2 summarizes the data parameters by their counts, standard deviation, mean, minimum, and maximum values to easily understand the data used for this study.

3.3. DNN model development and training

Deep learning (DL) is a breakthrough in AI that has surpassed prior

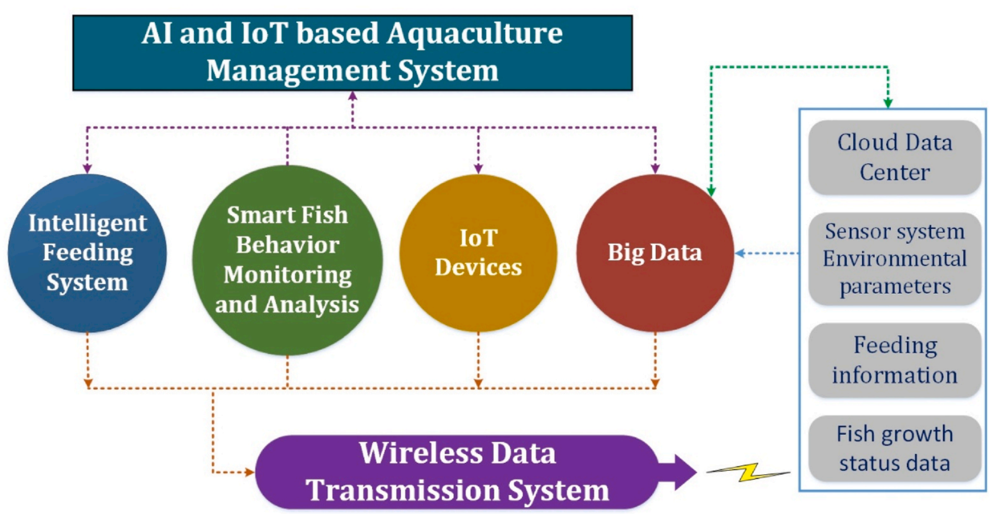


Fig. 4. Flow chart of system operations.

Table 2
The dataset input and output parameter range.

	Dissolved oxygen (ppm)	pH value	Temperature (^{o}C)	Turbidity (NTU)	Bait Quantity (grain/week)	Length Increment (cm/week)
count	312.000000	312.000000	312.000000	312.000000	312.000000	312.000000
mean	2.170673	6.829647	24.103205	5.313782	132.304487	0.354747
std	0.252274	0.158249	2.872755	1.864636	19.646889	0.139755
min	1.770000	6.000000	17.300000	3.200000	90.000000	0.100000
25%	1.880000	6.700000	22.200000	3.600000	120.000000	0.300000
50%	2.200000	6.800000	25.300000	4.000000	140.000000	0.340000
75%	2.400000	6.900000	26.200000	7.300000	140.000000	0.360000
max	2.600000	7.200000	28.100000	8.300000	170.000000	0.920000

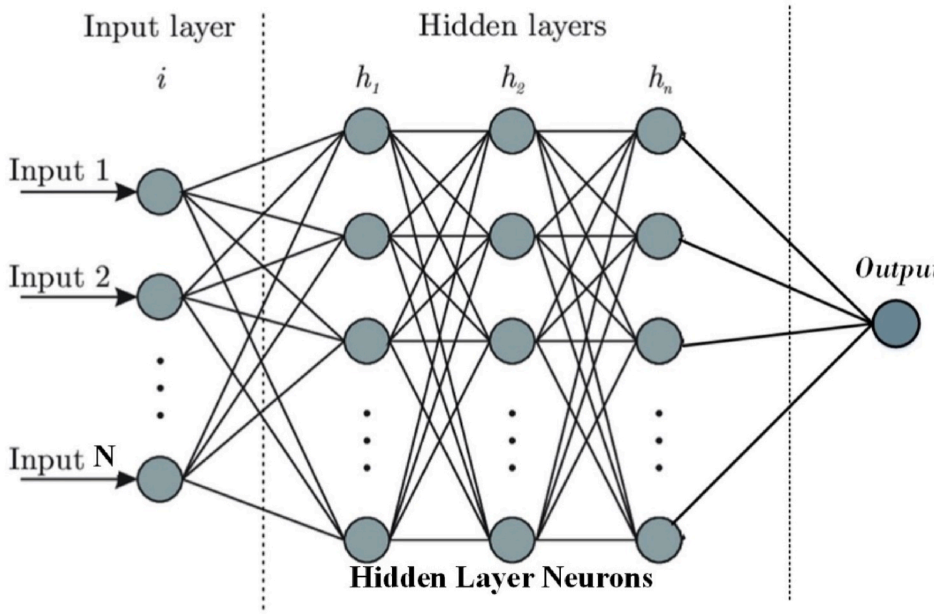


Fig. 5. Artificial neural network structure.

limitations. Artificial intelligence (AI) has been extensively used in various fields [21,22], including agriculture [23–26], bioinformatics, robotics, IoT, medicine, etc. DL is part of the machine learning (ML) area; however, it enhances the data preprocessing by automatically retrieving extremely nonlinear and complicated features via numerous layers, instead of needing manual parameter optimization for a specific data type depending on domain expertise. DL provides superior statistical tools for discovering, measuring, and comprehending the massive amounts of information in big data to assist smart fish farming, due to its automated feature learning and greater modeling capabilities.

A deep learning time-series technique such as LSTM may not be stable enough and, thus, insufficient for predicting all real-time dynamics. There is a problem because predictive processes can either be unique to a single target quality parameter and its dynamics over time or collective, including relationships between multiple predictors and their respective variables. Specifically, these involve creating Machine Learning models (such as Bagging and Random Forest, Support Vector Machines, or Neural Networks) as well as Time Series forecasting (such as Exponential Smoothing models, Autoregressive Integrated Moving Average models, or Vector Autoregression processes) and, when appropriate, combining these approaches to improve forecast quality. Thus, in this paper, it is desirable to use approaches without strong linear assumptions, such as neural networks that can cope with noise and take factors affecting the growth of the fish into account.

Artificial neural networks (ANNs) [24,27] include neurons that are connected unidirectionally so that they mimic the brain's capabilities for recognizing patterns and learning associations among data [28–30]. For each neuron n_i there is an activation function α_i, and each connection between two neurons n_i, n_j has an impact weight S_{ij} that determines the impact of neuron n_i on neuron n_j. By modeling complicated interactions between neurons, which form the basic processor units of an ANN [31], weighted connections are formed between them. Neurons are generally layered, with each layer's neurons connected directly to its successive layers Fig. 5. The first layer is referred to as the 'input layer,' and the last layer is referred to as the 'output layer,' with all layers in between called 'hidden layers.' A data set from the input layer is sent to the first hidden layer, where it is aggregated and modified by the following steps:

$$N_j = \sum_{i \in Q_j} S_{ij}\, Out_i \tag{1}$$

where S_{ij} represents the connection weights between n_i and n_j, Out_i denotes the output of n_i, Q_j represents the neuron set having an outward connection to n_j. The n_i output is determined as:

$$Out_i = \alpha(N_j) \tag{2}$$

where activation function of n_i is represented by α. The hyperbolic tangent function is defined as $g(y) = \frac{e^y - e^{-y}}{e^y + e^{-y}}$ is popular activation function is particularly useful due to its continuous and differentiable characteristics that are essential for the calculation of network error gradient. In a subsequent layer, the output of each neuron is transferred to the following layer's neurons. Until the network's output layer is reached, this technique is repeated for each layer beneath it. Ultimately, the network output can be seen as the output of the output layer.

To simulate nonlinear relationships, an ANN's connection weights have to be changed. There is typically a two-step process involved in this process. The first step of backpropagation involves determining each neuron's error signal for a given observation. The error signal is determined by the error function $Er = \frac{1}{2}\sum_{1}^{m}(V_i - Out_i)^2$, where V_i denotes the target value, and m represents the target value numbers. Based on this error function, the error signal is calculated:

$$\varepsilon_j = \begin{cases} \alpha'(N_j)\,(Out_j - V_j) & where\ j\ is\ an\ output\ neuron \\ \alpha'(N_j)\sum_k \varepsilon_k S_{ij} & otherwise \end{cases} \tag{3}$$

where α' and k represents the activation function and error signal for node k respectively.

Gradient descent is used to adjust the connection weights in the step second as:

$$\Delta S_{ij} = -\eta \frac{\partial Er}{\partial S_{ij}} = -\eta \varepsilon_j Out_i \tag{4}$$

Until a termination condition is met, both of these processes are repeated repeatedly.

To select the input parameters with high impact for building an ideal DL model efficient to find the correlation between different input parameters and forecast the growth of the fish in the smart pond, proper feature engineering was done. Dissolved oxygen (ppm), pH, temperature (^{o}C), turbidity (NTU), bait quantity (grain/week), and length increment

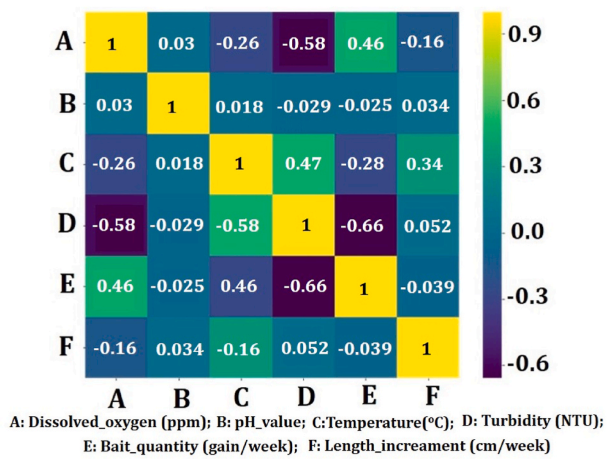


Fig. 6. Correlation heat map for input and output parameters.

(cm/week) were the parameters of the smart fish pond system and were taken into account. The numbers of input and output neurons were changed according to the findings of the feature engineering. To visualize the data dependency of the system input and output parameters, a heat map from the Seaborn library was utilized as the visualization technique as shown in Fig. 6. The correlation heat map shows the interrelationship among various system input parameters. From Fig. 6, it can be seen that the dissolved oxygen has strong correlations with all the input parameters temperature, turbidity, bait quantity, and length increment but has a less significant correlation with pH value. The temperature and bait quantity also show strong correlations with all the system parameters except pH value. Moreover, turbidity is also an important parameter for the aquaculture system and has shown significant interrelationship with other parameters.

For this research, we have utilized Gaussian process-based Bayesian optimization surrogate models [32,33] to find the best DL model that will produce the best R2 and mean square error. By optimizing an objective function that is expensive to evaluate, Bayesian optimization minimizes the number of actual function evaluations required. This technique, based on Bayesian inference and Gaussian processes [34], is suitable for situations where the objective function's closed-form expression is unknown but where it is known that observations (potentially noisy) at sampling values exist of the function.

In BayesOpt, a probabilistic proxy is created for the goal using the results of previous trials as training data. We can evaluate the underlying objective function with a proxy model (e.g., Gaussian process) at a lower cost and still get enough information to determine where we should evaluate the underlying objective function to achieve a decent result. Consider a vector $H_p = \{H_1, H_2, \ldots, H_n\}$ for a set of n hyper-parameters needs to be tuned. Given a collection of training settings $\{(r_i, s_i)\}_{i=1}^{m}$, we require to determine $H^* = argminf_H(H|\{(\{(r_i, s_i)\}_{i=1}^{m})$, where g denotes a cost function. The process of optimization is guided by a suitable acquisition function (AF), which decides the next point be assessed (i.e., the next collection of hyper-parameters). A balance between exploration and exploitation must therefore be established within each acquisition function. The Bayesian optimization with the Gaussian process is illustrated in Fig. 7.

Thus, by using the modified hyper-parameter [24] to estimate the following hyper-parameters, the Bayesian optimization considers the past evaluations. It is possible to concentrate on the sections of parameter space that have a high likelihood to produce the best validation score by picking a set of input parameters in an educated manner. To get the best set of hyper meter readings, this method usually takes fewer

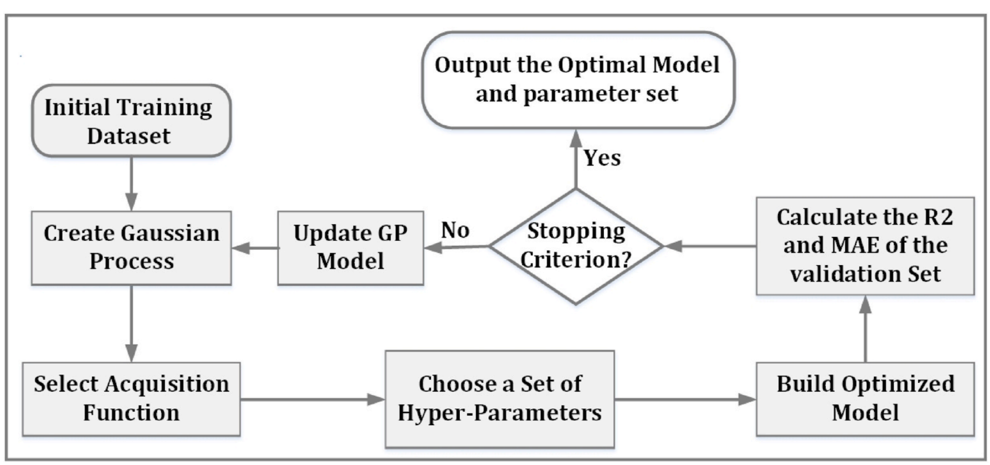


Fig. 7. Bayesian optimization surrogate models based on Gaussian process.

Table 3
Nine best-optimized models developed from the Bayesian optimization process.

Models	M1	M2	M3	M4	M5	M6	M7	M8	M9
Learning Rate	0.0988	0.01351	0.09911	0.01338	0.00024	0.00021	0.09379	0.00053	0.00014
Hidden Layer	2	8	1	7	10	10	1	10	8
Hidden Nodes/Layer	478	10	196	316	485	498	247	38	414
Activation	tanh	softsign	softsign	softsign	softsign	tanh	softmax	softsign	tanh
Acquisition	he_ normal	he_ normal	glorot_ normal	he_ normal	normal	he_ normal	glorot_ normal	he_ normal	he_ normal
Optimizer	Adadelta	Adamax	RMSProp	Adagrad	Adamax	Adamax	RMSProp	RMSProp	Adam
Decay Rate	0.00069	0.00923	0.00888	0.00113	0.00434	0.00044	0.00310	0.00284	0.00035
Batch size	5	43	136	189	298	289	302	297	48
Epochs	245	750	616	319	525	720	636	660	503
R_Score	0.882	0.804	0.858	0.885	0.880	0.891	0.874	0.917	0.895

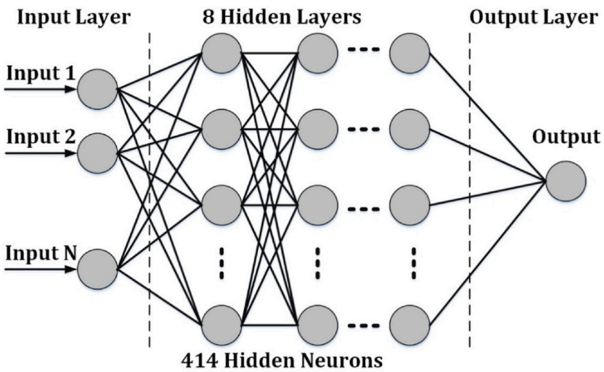


Fig. 8. The structure of the optimized DL model.

rounds. Specifically, it disregards parts of the variable space that have little effect on tables. As a result, the number of times that a model must be trained for validation is reduced since only parameters that provide a better validation score are taken into account. In this study, we have used Bayesian optimization for the random forest DL model with acquisition functions like EI and R^2 values that were used to assess them. Besides, activation functions, including tanh, softsign, he_uniform, relu, etc., were also tested. Varied optimizers were also tested, comprising RMS prop, SGD, Adam, etc., as well as diverse learning rates and batch sizes of (0.0001–0.1) and (1–310) respectively. Table 3 represents the best nine optimized models produced by the Bayesian optimization process. We have used an optimized model (M9) having an R^2 value of 0.89 on optimization and to produce an R^2 value of 0.94 on validation.

Fig. 8 shows the optimal DL model for assessing the growth of the fish in the smart pond system, which featured 5 neurons at the input layer, 8 hidden layers with 414 dense neurons in each layer, and a single neuron at the output layer. Adam and tanh were selected as suitable optimizers and activation functions. The values of 0.00014, 48, and 0.00035 are the best learning rate, batch size, and decay respectively of the optimal DL

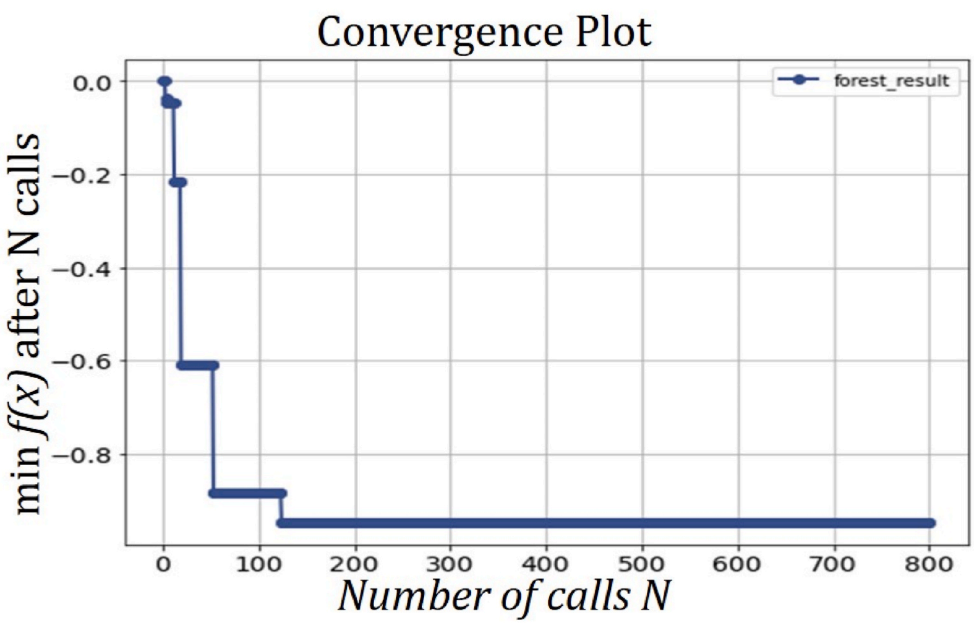


Fig. 9. Convergence plot for the developed models.

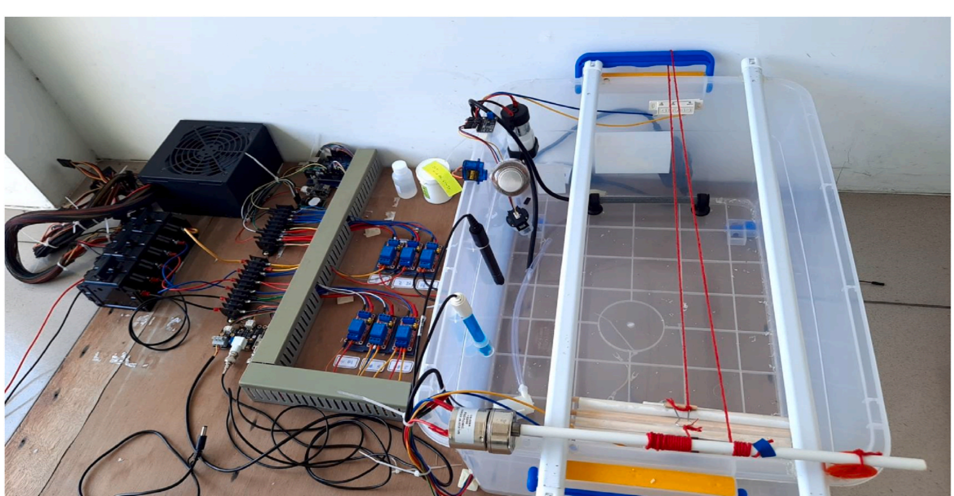
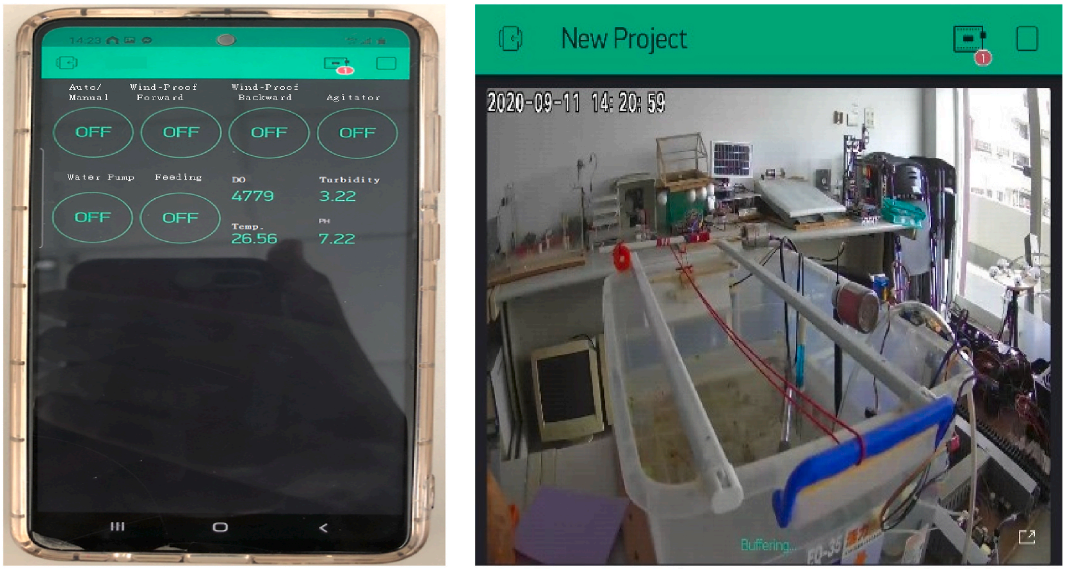


Fig. 10. The prototype of the intelligent fish-feeding system.



(a) (b)

Fig. 11. Screenshots of the developed system prototype (a) An IoT-based water quality monitoring and control shown on the cell phone (remote user), (b) Remote video monitoring.

model (M9) that we have used for the evaluation of the smart pond system. Throughout the training stage, the number of iterations was raised from 100 to 800. As a result, the required accuracy was achieved after 503 cycles, and training was completed. Once the ideal hyper-parameters concerning R^2 values were found, a DNN was built to evaluate the growth of fish in the developed smart fish pond.

Irrespective of the actual operation, the DL model was trained to find the correlations among the key variables of the system. To acquire the optimal DL model, weights and biases were modified in several rounds. To evaluate the accuracy of the predicted results, we have utilized some of the statistical error reduction methods such as coefficient of determination (R^2), mean absolute error (MAE), and mean square error (MSE) in this research. Thus, the goal of performing optimization of the produced ANN model was to reduce the outliers and ambiguity during prediction. The hyper-parameters were modified depending on the number of hidden layers to determine the optimal model and the neurons in every layer were mapped to enhance the produced ANN model. The convergence plot depicts the evolution of the time step along the optimization process. Thus, a convergence graph for different generated random forest models with different activation and acquisition functions

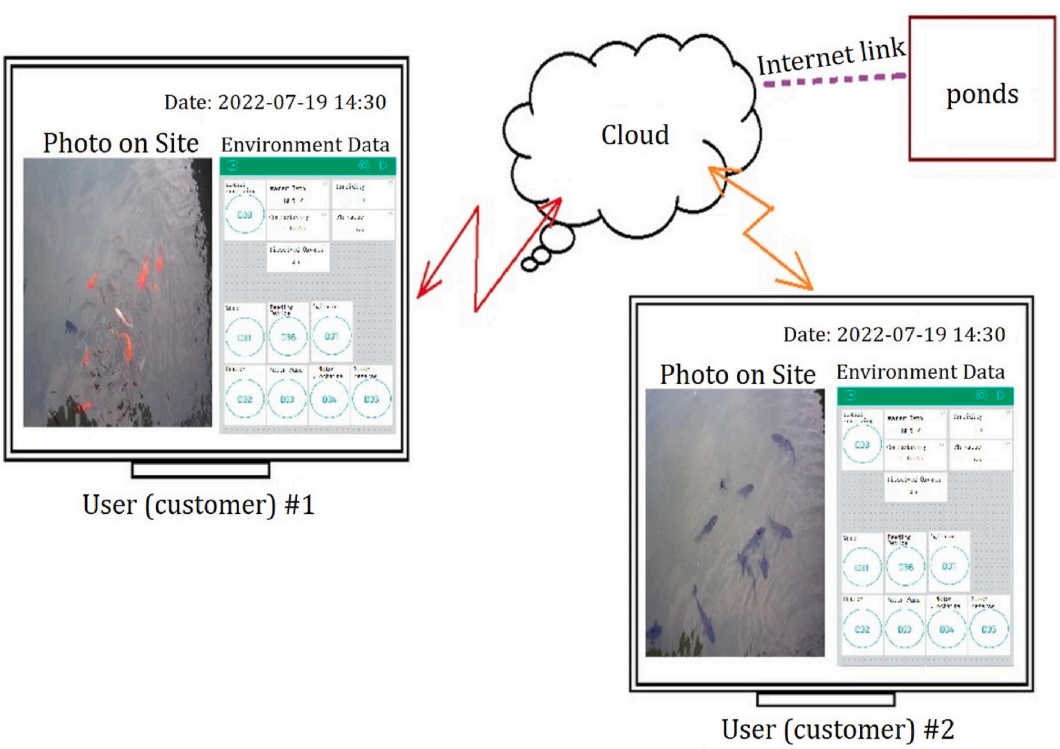


Fig. 12. Both the environmental data and site photo served as a portfolio.

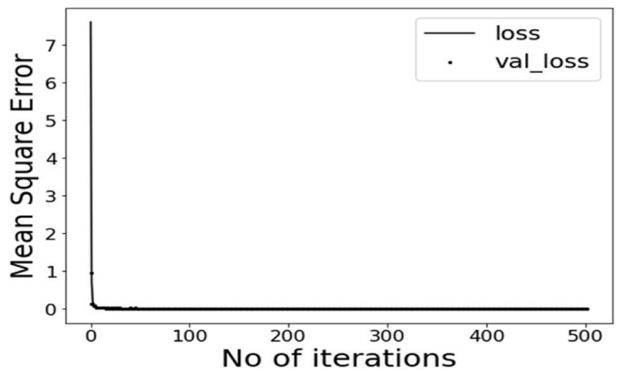


Fig. 13. Developed model training and validation losses.

is represented in Fig. 9. The figure demonstrates that most of the optimized models converge at nearly identical values.

4. Results and discussion

Five sensors (a dissolved oxygen sensor (DO sensor), temperature sensor, turbidity sensor, ph sensor, and a proximity switch sensor used for location identification purposes) and five actuators (a heater, a C·W·/C·C.W. motor used for wind-proofing purposes, a water pump, an agitator, and a feeding device driven by servo motor) have been adopted and built. The prototype of the intelligent fish-feeding system is illustrated in Fig. 10.

We have also developed a mobile application through which remote users can monitor the related water quality parameters measured by the smart pond via the cloud server as illustrated in Fig. 11. As described in Fig. 11(a), two modes (auto mode and manual mode) can be selected in the interface of the cell phone. To online monitor the status of fish baiting and record the fish length, IPCAM is installed onto the fish pond. Fig. 11(b) shows a screenshot of the online remote video monitoring through a cell phone or the remote PC via the IoT system. Since, users cannot stay connected 24 h a day, seven days a week, to monitor the parameters, so we designed an alarm system. By sending a line text message, the user is notified when a parameter for the farm exceeds a specific critical threshold. In this manner, farmers can react quickly and initiate a problem-solving process.

As shown in Fig. 2, the water quality parameters (dissolved oxygen level, water temperature value, turbidity level, and pH value) can be automatically controlled to the targeted region. Consequently, the intelligent fish pond system can record the site photo of fish by using an IPCAM. As shown in Fig. 12, both the environmental data and site photos stored on the cloud can also be served as a portfolio for the potential customer.

The other purpose of this study is to develop a DL generalized model for predicting the outcome of the developed smart fish pond (California Bass) with different variable parameters. Statistical error reduction methods were utilized to obtain the most appropriate parameters for forecasting new data points. The results of the developed deep learning model's learning process are shown to highlight its prediction accuracy.

Fig. 13 shows the iterative outcomes obtained to lower the MSE. When the parameters are applied, as indicated in the prior section, the ANN model was able to learn the data well. As a consequence, the optimal ANN model approached a zero MSE value within 14 cycles; thereafter, a constant error value was recorded until 500 iterations, when the repeated loop was stopped. Fig. 14 shows the results obtained from the trained AI model using tuned hyper-parameters. The output characteristics of smart California Bass fish pond length increment (cm/week) were found to have a linear relationship, showing strong learning and fitting with the provided experimental data. The ANN model precisely anticipates the output properties, according to the findings. Furthermore, the kernel error density charts show how the error approaches zero. Thus, the given optimal model was selected to predict the new datasets. The error density of length increment (cm/week) ranged from −0.15 to 0.20.

Therefore, the developed optimal model demonstrates viability and applicability by providing 0.94% accuracy under different input parameters. The output parameter error density converged to zero, confirming the correctness of the established model for assessing the performance of the smart California Bass fish pond. As a result, the generated optimal ANN model can be utilized for a wide range of tasks.

Using real-time farm information will allow you to identify problems before they occur since the prediction is based on real-time data. There may be a correlation between the evolution of one parameter and the evolution of external variables. Based on the real-time measurements of different parameters, we have used the AI-based Surrogate model to predict the most relevant parameters explaining the growth of the smart fish pond. Water temperature is found to be one of the most influential factors impacting water dissolved oxygen levels in different studies. Dissolved oxygen levels in aquaculture systems are affected by several parameters, and these parameters are also strongly influenced by the type of aquaculture. The extent of these variables varies depending on the type of production. However, most can be quantified and may be taken into account in a prediction model based on the Internet of Things. These parameters include the nutritional quality of feeds and their chemical composition. In addition, the length and feed level of fish in any system, as well as fish biomass and algae biomass in fish systems. Aquaculture recirculation systems depend on the biofilter activity of ponds.

Furthermore, predictive models may provide farmers with the

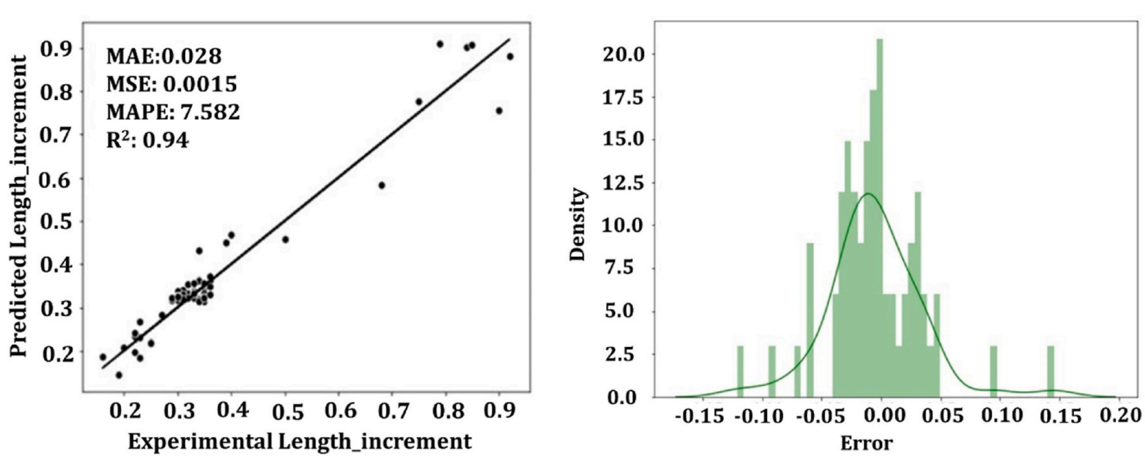


Fig. 14. Optimal model prediction performance and error density analysis.

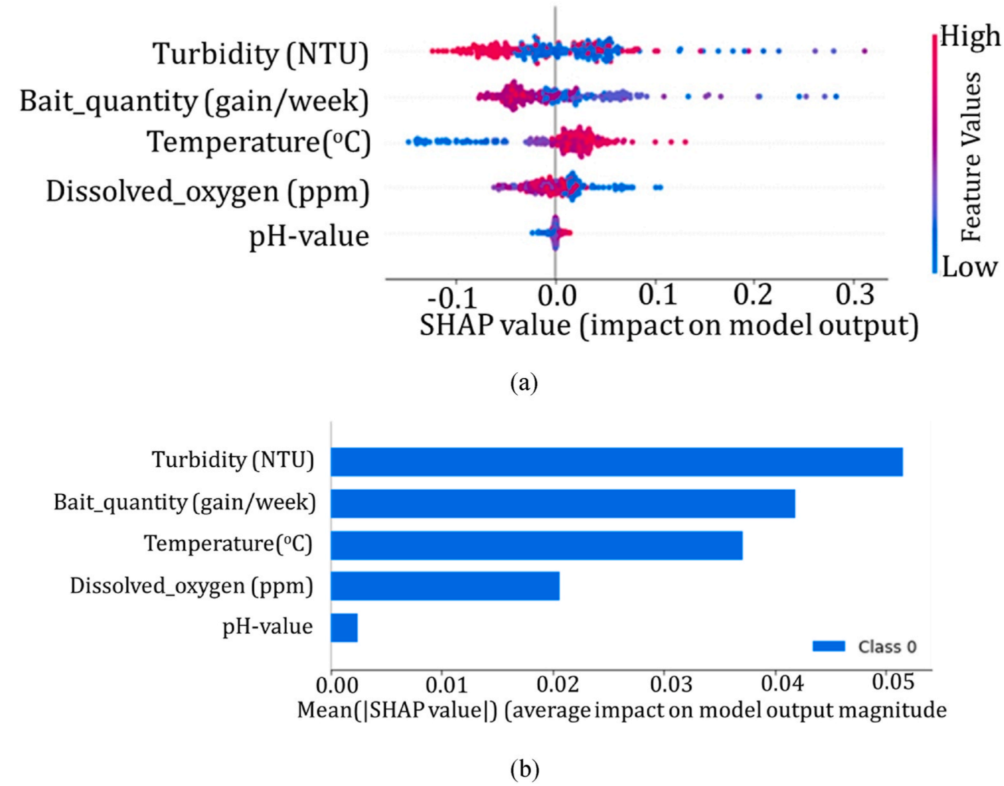


Fig. 15. Impact of each input variable on the model output magnitude.

dissolved oxygen concentration value a day in advance, with a high prediction range, so they can adjust feeding or activate aerators as necessary. As a result, farmers can lose their entire crop if inadequate oxygen management is not made. Our study combined the results of 52 weeks of different parameter monitoring at an hourly level. In this study, the SHAP library is used to depict the impact of each input parameter on the output of the generated optimal model. The input parameters turbidity (NTU) and bait quantity (grain/week) were the most important variables for predicting the output parameter. The input parameter pH has the lowest impact on the output of the optimal ANN model. As shown in Fig. 15, the SHAP library is employed to visualize the impact of the input characteristics on the generated AI model output. Varying the turbidity may result in large variations in the growth of the fish. Because high turbidity reduces the feeding abilities of the California Bass fish. Therefore, this species' avoidance of the turbid waters could correspond to an increase in its feeding ability. Thus, the ability of DNN models to produce accurate assessments of the parameter in question can be demonstrated given the right set of parameters such as turbidity, bait quantity, and temperature.

The results of this research can significantly influence on-farm management. Having water quality monitors, dissolved oxygen levels, and other parameters constantly monitoring the pond is extremely beneficial to the farmer. One of the most significant things to monitor is the oxygen level in the pond. Fishermen are used to taking manual measurements regularly, frequently at the same time every day, but having the measurements continuously reveals patterns that they weren't aware of before. During that time, they have changed the way they feed the fish (in the fish farm) or increased their biomass.

Fish feed represents a significant share of total input costs, accounting for 20 to 50% of total expenditures. This is due to a scarcity of sustainable raw materials. The proposed aquaculture farms should assist in the optimization of feeding for long-term sustainability. Boosting nutrient utilization and reducing feeding costs will increase yields and help the farming industry while decreasing the environmental impact of undegraded feed. With the help of digitizing real-time biological and environmental information, coupled with IoT and machine learning technologies, and new technologies, feeding techniques can be improved to optimize profit while reducing environmental footprint. In this study, an innovative, adaptable, and integrative smart aquaculture farming system featuring smart feeding, monitoring, and control will be developed to combat this problem. The system will enhance feed utilization efficiency. Real-time monitoring and evaluation will be conducted on all factors that may impact fish diets, such as the fish itself, its habitat, and external parameters. It will make it possible to feed them at the right time, in the right quantity, and with the right quality.

Although real-time monitoring capabilities are the first step, they are only one part of the solution. In the future, we plan to integrate water monitoring data with production and meteorological data so that farmers can receive tailored production recommendations. As part of the machine learning algorithms utilized in the data, we also plan to anticipate and forecast crucial occurrences for farmers.

5. Conclusion and future work

This study demonstrates a smart aquaculture monitoring and control system based on IoT devices was developed to collect and analyze data, as well as to develop an artificial intelligence-based DL model to find the correlation between different system parameters and predict the growth of the California Brass fish. The smart fishpond is equipped with multiple sensors such as pH sensor, temperature sensor, dissolved oxygen sensor, and turbidity sensor connected with controller Arduino Mega2560 with integrated Wi-Fi module, multiple actuators such as a heater, limit switch, water pump, agitator, windproofing device, and smart feeding device, and IPCAM for real-time monitoring. This allows for gathering large amounts of data about fish and feeding in a cage environment quickly and in real-time. The designed fish pond (cf. Fig. 1) can monitor automatically the water temperature, the water pH value, the water turbidity, and the dissolved oxygen, and control these parameters in a specific region by triggering the respective actuators through a developed mobile application. To find the appropriate conditions for the growth of California Bass fish, we have developed a DL model to perform the correlation analysis of different input parameters with the output parameter on the data collected for 52 weeks. Upon validation, the optimal model produces an R^2 value of 0.94 and a mean square error of 0.0015 and thus demonstrates the applicability of the model to predict the desired output.

Furthermore, it may be useful in interpreting pictures of underwater aquatic creatures and developing AI feeding systems for AI computations. With underwater image processing technologies, aquaculture operators can analyze real-time pictures and receive current creature statuses non-invasively and without interruption. In addition, for the food safety and quality check, the video and other data collected during farming can be made available to customers and food agencies through QR code tags attached to the product. Thus, based on the findings of our study, aquaculture operators or owners can decrease feed residues, monitor fish growth, and boost fish survival rates, resulting in a higher feed conversion rate.

Author contributions

Conceptualization, S.A.B. M.-C.C., N.-F.H., and W.-M.Y.; methodology, S.A.B. and N.-F.H.; validation, S.A.B. M.-C.C., N.-F.H., and W.-M.Y.; formal analysis, S.A.B, N.-F.H.; investigation, S.A.B; resources, M.-C.C., N.-F.H., and W.-M.Y.; data curation, S.A.B, N.-F.H., M.-C.C., and W.-M. Y.; writing—original draft preparation, S.A.B., and M.-C.C.; writing—review and editing, S.A.B. and N.-F.H.; visualization, S.A.B., M.-C. C., N.-F.H., and W.-M.Y.; supervision, N.-F.H., and W.-M.Y.; project administration, S.A.B., M.-C.C., N.-F.H., and W.-M.Y.; All authors have read and agreed to the published version of the manuscript.

Availability of data and materials

Previously reported data used to support this study are available at DOI or other persistent identifiers. These prior studies are cited at relevant places within the text as references [**].

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Riansyah, et al., Fish feeding automation and aquaponics monitoring system base on IoT, in: 2020 6th International Conference on Wireless and Telematics (ICWT), IEEE, 2020.
[2] F.K. Došilović, M. Brčić, N. Hlupić, Explainable Artificial Intelligence: A Survey. In 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE, 2018.
[3] J. Waterman, H. Yang, F. Muheidat, AWS IoT and the interconnected world–aging in place, in: 2020 International Conference on Computational Science and Computational Intelligence (CSCI), IEEE, 2020.
[4] C. Zhou, et al., Intelligent feeding control methods in aquaculture with an emphasis on fish: a review, Rev. Aquacult. 10 (4) (2018) 975–993.
[5] X. Yang, et al., Deep learning for smart fish farming: applications, opportunities and challenges, Rev. Aquacult. 13 (1) (2021) 66–90.
[6] S. Zhao, et al., Adaptive neural fuzzy inference system for feeding decision-making of grass carp (Ctenopharyngodon idellus) in outdoor intensive culturing ponds, Aquaculture 498 (2019) 28–36.
[7] D.S. Simbeye, J. Zhao, S. Yang, Design and deployment of wireless sensor networks for aquaculture monitoring and control based on virtual instruments, Comput. Electron. Agric. 102 (2014) 31–42.
[8] J.-H. Chen, W.-T. Sung, G.-Y. Lin, Automated monitoring system for the fish farm aquaculture environment, in: 2015 IEEE International Conference on Systems, Man, and Cybernetics, IEEE, 2015.
[9] B. Shi, et al., A wireless sensor network-based monitoring system for freshwater fishpond aquaculture, Biosyst. Eng. 172 (2018) 57–66.
[10] L. Hongpin, et al., Real-time remote monitoring system for aquaculture water quality, Int. J. Agric. Biol. Eng. 8 (6) (2015) 136–143.
[11] F.E. Idachaba, et al., IoT Enabled Real-Time Fishpond Management System, 2017.
[12] G. Cario, et al., Long lasting underwater wireless sensors network for water quality monitoring in fish farms, in: OCEANS 2017-Aberdeen, IEEE, 2017.
[13] W.-C. Hu, et al., A computer vision-based intelligent fish feeding system using deep learning techniques for aquaculture, IEEE Sensor. J. 22 (7) (2022) 7185–7194.
[14] R.K. Kodali, A.C. Sabu, Aqua monitoring system using AWS, in: 2022 International Conference on Computer Communication and Informatics (ICCCI), IEEE, 2022.
[15] M.D. Fadillah, et al., Fuzzy logic-based control system to maintain pH in aquaponic, in: 2021 7th International Conference on Wireless and Telematics (ICWT), IEEE, 2021.
[16] F. Budiman, M. Rivai, M.A. Nugroho, Monitoring and control system for ammonia and ph levels for fish cultivation implemented on raspberry pi 3B, in: 2019 International Seminar on Intelligent Technology and its Applications (ISITIA), IEEE, 2019.
[17] D. Mayrawan, et al., Development internet of things for water quality monitoring system for gouramy cultivation, in: 4th Forum in Research, Science, and Technology, Atlantis Press, 2021 (FIRST-T1-T2-2020).
[18] S. Zahedi, et al., Effect of stocking density on growth performance, plasma biochemistry and muscle gene expression in rainbow trout (Oncorhynchus mykiss), Aquaculture 498 (2019) 271–278.
[19] R.S. Davidson, B.H. Letcher, K.H. Nislow, Drivers of growth variation in juvenile Atlantic salmon (Salmo salar): an elasticity analysis approach, J. Anim. Ecol. 79 (5) (2010) 1113–1121.
[20] K. Lorenzen, Toward a new paradigm for growth modeling in fisheries stock assessments: embracing plasticity and its consequences, Fish. Res. 180 (2016) 4–22.
[21] S.A. Bhat, N.-F. Huang, Big data and AI revolution in precision agriculture: survey and challenges, IEEE Access 9 (2021) 110209–110222.
[22] J. He, et al., The practical implementation of artificial intelligence technologies in medicine, Nat. Med. 25 (1) (2019) 30–36.
[23] S.A. Bhat, et al., On the classification of a greenhouse environment for a rose crop based on AI-based surrogate models, Sustainability 13 (21) (2021), 12166.
[24] H.M. Asfahan, et al., Artificial intelligence for the prediction of the thermal performance of evaporative cooling systems, Energies 14 (13) (2021) 3946.
[25] T.-W. Huang, et al., Artificial intelligence-based real-time pineapple quality classification using acoustic spectroscopy, Agriculture 12 (2) (2022) 129.
[26] U. Sajjad, et al., A deep learning method for estimating the boiling heat transfer coefficient of porous surfaces, J. Therm. Anal. Calorim. 145 (4) (2021) 1911–1923.
[27] U. Sajjad, et al., Determining the factors affecting the boiling heat transfer coefficient of sintered coated porous surfaces, Sustainability 13 (22) (2021), 12631.
[28] H. Kukreja, et al., An introduction to artificial neural network, Int J Adv Res Innov Ideas Educ 1 (2016) 27–30.
[29] S. Shanmuganathan, Artificial neural network modelling: an introduction, in: Artificial Neural Network Modelling, Springer, 2016, pp. 1–14.
[30] Z. Zhang, Artificial neural network, in: Multivariate Time Series Analysis in Climate and Environmental Research, Springer, 2018, pp. 1–35.
[31] U. Sajjad, et al., Boiling heat transfer evaluation in nanoporous surface coatings, Nanomaterials 11 (12) (2021) 3383.
[32] P.I. Frazier, A Tutorial on Bayesian Optimization, 2018 arXiv preprint arXiv: 1807.02811.
[33] U. Sajjad, et al., Liquid-to-vapor phase change heat transfer evaluation and parameter sensitivity analysis of nanoporous surface coatings, Int. J. Heat Mass Tran. 194 (2022), 123088.
[34] M. Ali, et al., On the assessment of the mechanical properties of additively manufactured lattice structures, Eng. Anal. Bound. Elem. 142 (2022) 93–116.