AI-Driven Water Quality Prediction For Aquatic Ecosystem Using Deep Learning and Automated Hyperparameter Optimization

Akshat Choudhary
Department of Computer Science and Engineering Apex Chandigarh Engineering College
Jhanjeri, Mohali, India
akshatchoudhary81172@gmail.com

Ankur Srivastava
Department of Computer Science and Engineering Apex Chandigarh Engineering College
Jhanjeri, Mohali,India, ankur.j3017@cgc.ac.in
Research Scholar, Department of Computer Science & Engineering Integral University, Lucknow, India

***Abstract*—Accurate water quality assessment is critical for environmental monitoring and public health. This research proposes a deep learning-based classification model for water quality prediction for aquatic ecosystems, leveraging automated hyperparameter optimization via Optuna to enhance accuracy and computational efficiency. The dataset undergoes rigorous preprocessing, including feature standardization with StandardScaler and stratified train-test splitting to maintain class distribution integrity. The neural network architecture consists of three fully connected hidden layers with ReLU activation, batch normalization for training stability, and dropout regularization to mitigate overfitting. Model training employs the Adam optimizer with a dynamic learning rate scheduler (ReduceLROnPlateau). The optimization process tunes key hyperparameters, including neuron allocation per layer, dropout rate, and learning rate, yielding an optimal configuration of 179-104-36 neurons, 0.2586 dropout, and a 0.00999 learning rate. Experimental evaluation using stratified k-fold cross-validation achieves a test accuracy of 98.02%, with high precision, recall, and F1-scores, ensuring robust predictive performance. The model that has been trained has been successfully implemented and has a well-defined and standardized process for handling input. This allows the model to generalize smoothly over a wide range of extremely dissimilar datasets that are specifically pertinent to water quality measurements. In comparison to the traditional models that have been applied in the past, this new AI-based method has been shown to have an unprecedented level of enhanced generalization capabilities. It also contains significantly reduced levels of computational overhead, which is a major advantage, as well as improved reliability when conducting classification tasks. These exemplary features make it an extremely scalable solution that is highly suited for the application of real-time water quality monitoring, particularly in the specific application of IoT-based environmental monitoring systems.**

***Index Terms*—Water Quality Prediction, Aquatic Ecosystems, Deep Learning, Hyperparameter Optimization,Optuna, Environmental Monitoring**

I. Introduction

Water quality observation ranks among the most important priorities in terms of not just maintaining the environment but also ensuring public health and safety. Once water sources are contaminated, they turn into serious concerns that can cause a myriad of waterborne diseases, interfere with ecological balances in extreme manners, and contribute to inefficiencies when it comes to various industrial applications, as has been observed in recent studies [1]. The conventional methods that have been used to estimate water quality are mainly rooted in manual sampling techniques that are followed by laboratory testing; both steps are extremely labor-intensive and time-consuming, rendering them not so ideal to use when it comes to undertaking large-scale reviews or real-time monitoring activities [2], [3]. Additionally, the available computational models that are being used tend to fail when being asked to handle high-dimensional inputs; such models become susceptible to overfitting-related issues and lack the required flexibility to function efficiently under different environmental conditions [4]. Such issues highlight the urgent need to design a more sophisticated and automated method that has the potential to significantly improve water quality classification and enhance the accuracy of prediction results [5].

The core theme of this comprehensive research work is in the complex and intricate process of building a highly sophisticated predictive model that specializes in the task of classifying water quality into a well-prepared list of pre-specified categories, which include classes like Good, Moderate, or Poor. This comprehensive research study goes deep into great depth, indulging in an elaborate discussion of various machine learning and deep learning techniques, as it not only aims to improve and maximize the accuracy of such classification but also strives to keep and sustain the computational cost of the model at an optimal and efficient level. Key research areas of study that are addressed with meticulous care within this work include basic operations like the preprocessing of the dataset, the thorough standardization of features, the thorough training of the model, exhaustive optimization procedures, and the use of optimal deployment steps that can be used effectively in various real-world applications. Apart from these basic things, the study also considers various strategies directed towards improving the generalization power of the model, reducing the risk of overfitting, and enhancing the model's flexibility across a wide variety of water sources and varied environmental conditions, all of which are essential to its effectiveness in real-world applications.

In a genuine and earnest attempt to address, resolve, and ultimately overcome these severe and complex challenges that have emerged over the past few years, various AI-based methods have been developed with utmost care and identified as being of high effectiveness as solutions following serious deliberation. These innovative new methods include a diverse array of advanced techniques, which are not exhaustive, such as neural networks, ensemble learning, and hyperparameter optimization, each technique having its own strengths and advantages. Each of these methods has been proven to have high potential for predictive performance improvement, a fact of robust empirical evidence behind it [6]. By employing the high capabilities of deep learning methods skillfully, researchers and practitioners can detect subtle and complex patterns in large water quality datasets. This robust capability ultimately leads to achieving higher accuracy classifications, a fact of utmost significance for facilitating effective decision-making processes in the critical fields of environmental monitoring and management [7]–[9]. Apart from this most important consideration, it is noteworthy to mention that a variety of methods, including but not limited to cross-validation and regularization, can be incorporated into the current framework with utmost ease. Such incorporation would be to further improve the model, rendering it stronger and less vulnerable to the issue of overfitting [10]. Application of such great advancements is a precious opportunity to dramatically improve the efficiency and reliability of real-time water quality monitoring systems [11].

The primary goal of the current research work is to conduct an exhaustive and systematic investigation of, and to actually implement, an all-encompassing holistic system that has been specially developed with the aim of predicting water quality in different aquatic environments. The system will utilize not only the current but also the most sophisticated preprocessing methods available, and these will be combined with highly effective training algorithms and state-of-the-art optimization methods that have been specially developed to tackle this specific application in depth. With the effective utilization of automated hyperparameter tuning and the implementation of deep learning architectures, as proposed in [12], the current research work aims to attain a scalable classification model with high accuracy and precision. It is proposed that this model can be transferable to a broad spectrum of real-time environments, thus covering and aiding different scenarios and situations that may occur. The primary findings that are attained with this exhaustive research work have the potential to contribute significantly to the overall development of dependable water monitoring systems that can be efficiently used. This innovation is proposed to facilitate enhanced and better-informed decision-making processes within the crucial fields of environmental protection and public health safeguarding, eventually making these industries more effective.

Prior to undertaking the complex and multifaceted task of developing an artificial intelligence model specifically designed to accurately predict water quality, it is of the utmost importance to undertake a thorough and rigorous analysis. This is to determine whether indeed this ambitious project is practical and feasible in its nature. Feasibility studies form a necessary and integral part of this overall project, which forms a very important part of analyzing a myriad of primary considerations that need to be taken into account. Among these is the consideration for the availability of adequate quality data that can be effectively harnessed, the ability of the model to both perform optimally and in an efficient manner for its desired purposes, and the potential of the model to deliver outputs that can maximize benefits, thus prove useful in real-world applications that dictate daily life. Previous research undertaken by researchers such as Haque et al. [13] and Alnemari et al. [14] has conclusively established that undertaking a thorough feasibility analysis is a very critical exercise, playing an instrumental role in the successful conceptualization and application of artificial intelligence systems specifically designed for environmental monitoring purposes. Accordingly, the present research endeavor entails a strong commitment to the systematic examination of the practicality and general viability of employing artificial intelligence methodologies for the specific purpose of water quality prediction in various environments.

II. Literature Review

Recent advancements in AI-driven deep learning and automated hyperparameter optimization have greatly enhanced water quality prediction, contamination source identification, and risk assessment in aquatic environments. A comparison of various studies in this domain is presented in Table I. Haque et al. [13] examined heavy metals (HMs) spatial-temporal variability, contamination sources, and health issues in the Barnoi River, Bangladesh, from July 2019 to June 2020. Although the water quality fulfilled drinking and aquatic life standards, Cr, As, Pb, and Cd exceeded criteria in their investigation of physico-chemical properties and heavy metal concentrations. River water was non-potable and moderately polluted, with Sites 1 and 2 having significant pollution levels. PCA and clustering analysis revealed industrial and municipal waste as anthropogenic sources. The health risk assessment found that children are more sensitive to carcinogenic and non-carcinogenic risks, recommending continuing monitoring.

Jiang et al. [15] developed a deep learning system to accurately predict fish pond aquaculture dissolved oxygen (DO) trends. They used Spearman correlation analysis (SCA), variational mode decomposition (VMD), and convolutional neural networks (CNN) to enrich and extract spatiotemporal features from time-series water quality data using SVC. Next, multi-step dissolved oxygen prediction was done using LSTM and GRU networks. The methodology improved accuracy by 16.8%-19.5%, with SVC-BiGRU achieving the best result (R^2 = 0.962, 0.934, 0.940 for 15, 30, and 45-minute forecasts). It reduces dangers and enhances precision aquaculture by detecting dissolved oxygen in real time.

Kol et al. [16] predicted Bacterial Gill Disease (BGD) in aquaculture using multivariate water quality metrics in the LSTM-Enhanced Asynchronous Advantage Actor-Critic (LEA3C) model. The model uses LSTM's sequential learning and A3C's advantage-based optimisation and asynchronous learning to identify complex temporal patterns and early BGD epidemic indications in real time. THE LEA3C model on the "Pondsdata" dataset had a projected accuracy of 97.50%, outperforming MDQN (88.52%), M-GRU (78.92%), and S-LSTM (74.31%). Faster and more accurate predictions improve aquaculture disease management, reducing economic losses.

Liu et al. [17] developed a sophisticated water quality prediction model for intelligent mariculture that addresses aquatic nonlinearity and dynamics. To increase prediction accuracy, they used data preparation and Pearson correlation analysis on 23,204 time-series recordings. The Bi-S-SRU network uses future context to improve prediction accuracy. In experiments, their model outperformed RNN and LSTM in accuracy and efficiency, predicting 3–8 days ahead with 94.42% accuracy and 12.5ms computational complexity. This strategy helps mariculture control water quality proactively.

Abiotic variables affected soil-water dynamics, dissolved oxygen concentrations, and fish mortality in Saudi Arabian aquaculture ponds, according to Alnemari et al. [14]. These effects were measured using eleven universities' data to create an environmental severity scale. Multiple regression analysis shows that soil temperature predicts dissolved oxygen levels (93.01% influence), but soil organic content has a negative connection (r = -0.87, p ¡ 0.001). Fish mortality increased significantly at 60% environmental harshness (r = 0.91, p ¡ 0.001). In arid locations, their logistic regression model correctly predicted system stress, improving soil-water management for sustainable aquaculture.

Liu et al. [18] developed the DAM-ResNet-LSTM model to forecast aquaculture water quality using dual-channel ResNet and LSTM with dual-attention techniques to increase feature extraction. The model, trained on offshore farm data, predicted pH, DO, and SAL accurately with Nash coefficients of 0.9361, 0.9396, and 0.9342. It improved Nash coefficients by 19.32% and 10.39% over single-channel DA-ResNet and DA-LSTM models. Attention increased ResNet-LSTM accuracy by 4.13%, proving its efficacy in forecasting nearshore aquaculture water quality.

The Chen et al. [19] machine learning-based early warning system predicts ammonia nitrogen and nitrite levels in industrial recirculating aquaculture systems for prawn cultivation. They estimated parameters that are hard to monitor in real time using inexpensive water quality sensors and GRNN, DBN, LSTM, and SVM models. The optimised GRNN model predicted ammonia nitrogen with MAE = 0.5915, MAPE = 28.95%, and RMSE = 0.7765 and nitrite with MAE = 0.1191, MAPE = 29.65%, and RMSE = 0.1904. IoT-based water quality monitoring devices can use these models in real time.

Wang et al. [20] used Landsat images from 2002 to 2022 to construct a machine learning model to study long-term water quality changes in the Ma'an Archipelago Marine Special Protected Area (MMSPA). A random forest model was used to invert chlorophyll-a (Chl-a), phosphate, and dissolved inorganic nitrogen (DIN) concentrations, yielding a R^2 of 0.741, RMSE of 3.376 µg/L, and MAPE of 16.219% for Chl-a. Results showed elevated nearshore concentrations, especially in aquaculture regions, due to riverine (Yangtze River) and human inputs. Remote sensing is vital for real-time marine water quality monitoring and protection, according to the study.

Aquaponics mixes aquaculture with hydroponics for sustainable food production, according to Owusu et al. [21]. Water recycling and reduced demand were highlighted. Machine learning was used to assess water quality metrics to improve system precision and reduce manual work. The Convolutional LSTM model outperformed the Recurrent Neural Network with projected accuracies of 96%, 98%, and 99% at 5W, 10W, and 15W after 200 epochs and 64 batches. The results demonstrate aquaponics' effectiveness in food security, particularly for growing catfish, salmon, spinach, cabbage, kintonmire, and cauliflower.

Recent research have shown that machine learning, deep learning, and remote sensing have improved water quality monitoring, prediction, and management. Deep learning-based algorithms have improved aquaculture dissolved oxygen and disease prediction, whereas heavy metal contamination research emphasises the necessity for regular monitoring and pollution mitigation. Innovative hybrid models like LSTM-Enhanced A3C and Bi-S-SRU improve dynamic water forecasting. IoT and remote sensing offer scalable real-time monitoring and sustainable resource management solutions. These discoveries enhance aquatic ecosystem health, aquaculture efficiency, and environmental protection.

In the evolving field of water quality evaluation and AI-based forecasting, several datasets have been compiled throughout the years to improve predictive modelling and decision-making. Table II shows a detailed and comprehensive overview, outlining in depth a comprehensive overview of the datasets that fall into the category of traditional, as well as newer and newly introduced ones. Such datasets form a key input during the training and testing phases of AI-based models for water quality prediction, founded to improve our knowledge and quantification of water quality.

III. Methodology and Techniques

This paper is a detailed and elaborate explanation of the different methodologies and different approaches used in the forecasting of the numerous forms of water quality found in aquatic ecosystems. Through these methodologies, it ensures that effective and proper assessments are carried out, which consequently plays a very important role in effective and efficient water resources management for sustainable use.

TABLE I

COMPARISON OF STUDIES ON WATER QUALITY ASSESSMENT AND PREDICTION TECHNIQUES.

Author	Domain	Technique	Metrics	Key Findings	Disadvantage
Haque et al. [13]	Water Pollution (Barnoi River)	PCA, Clustering, Pollution Indices	Cr, As, Pb, Cd exceeded limits	River water undrinkable; pollution from industrial & municipal waste; children at higher health risk	Limited to one river, lacks broader geographic applicability
Jiang et al. [15]	Aquaculture (DO Prediction)	SCA, VMD, CNN, LSTM, GRU	R^2 = 0.962 (SVC-BiGRU)	Improved DO prediction accuracy (16.8%-19.5%); enhances precision aquaculture	High computational complexity
Kol et al. [16]	Aquaculture (Disease Prediction)	LSTM-Enhanced A3C	Accuracy = 97.50%	Outperforms MDQN, M-GRU, S-LSTM; early BGD detection	Requires real-time data availability
Liu et al. [17]	Smart Mariculture (Water Quality)	Bi-S-SRU (Bi-directional Stacked SRU)	Accuracy = 94.42%	Outperforms RNN & LSTM; enables proactive water quality management	Computational overhead for long-term forecasting
Alnemari et al. [14]	Aquaculture (Environmental Impact)	Multiple Regression, Logistic Regression	Soil temp. = 93.01% influence on DO; r = -0.87 (organic matter & DO)	Developed severity scale; predicts fish mortality based on environmental factors	Focused on Saudi Arabian aquaculture, limits global generalization
Liu et al. [18]	Aquaculture (Water Quality Forecasting)	DAM-ResNet-LSTM (Dual-Attention)	Nash = 0.9361 (pH), 0.9396 (DO), 0.9342 (SAL)	Outperforms DA-ResNet & DA-LSTM by 19.32% & 10.39%	Requires high-quality training data
Chen et al. [19]	Industrial Aquaculture (Ammonia/Nitrite Prediction)	GRNN, DBN, LSTM, SVM	MAE = 0.5915, RMSE = 0.7765 (NH_3); MAE = 0.1191, RMSE = 0.1904 (NO_2)	IoT-based real-time monitoring; GRNN model best performance	High error rates in real-world applications
Wang et al. [20]	Marine Water Quality (Remote Sensing)	Random Forest, Landsat Data	R^2 = 0.741 (Chl-a), RMSE = 3.376 μg/L	Identified aquaculture-related pollution trends over 20 years	Limited accuracy in high-resolution spatial analysis
Owusu et al. [21]	Aquaponics (Water Quality Monitoring)	ConvLSTM, RNN	Accuracy: 96%-99% (power levels: 5W, 10W, 15W)	Improved accuracy & efficiency in aquaponics; supports sustainable food production	Power consumption dependency

A. Dataset and Preprocessing

The data, which has been accessed from **Mendeley Data**, comprises different parameters of water quality like pH and dissolved oxygen. These parameters have been categorized and designated as per their quality into **Good, Moderate, or Poor**.

1) Feature Engineering and Standardization: Standardization was applied using *z-score normalization*:

$$X' = \frac{X - \mu}{\sigma} \tag{1}$$

where X' is the standardized value, μ is the mean, and σ is the standard deviation.

2) Train-Test Split: The dataset was split into **80% training** and **20% testing** using *stratified sampling* to maintain class balance:

$$D_{\text{train}}, D_{\text{test}} = \text{train\_test\_split}(D, \text{test\_size} = 0.2, \text{stratify} = Y) \tag{2}$$

B. Model Architecture

The proposed **fully connected neural network** consists of:

- **Input Layer**: Accepts 14 standardized features.
- **Hidden Layers**: Three layers with 179,104 and 36 neurons, each using **ReLU activation**:

$$h_i = \max(0, W_i x + b_i) \tag{3}$$

- **Batch Normalization**: Applied after each hidden layer to Stabilizes learning.
- **Dropout Layer (0.2586)**: Help reduces overfitting.
- **Output Layer**: **CrossEntropyLoss** manages the output layer internally, therefore there is no explicit activation is applied.

C. Hyperparameter Optimization with Optuna

Instead of manual selection, **Optuna** was used to optimize:

- **Hidden Layers**: 179, 104, 36 neurons.
- **Dropout Rate**: 0.2586.
- **Learning Rate**: 0.00999.

A **Stratified K-Fold (k=5) Cross-Validation** approach was used for generalization:

$$\mathcal{L} = \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}_i \tag{4}$$

D. Training and Optimization

1) Loss Function and Optimizer: The model was trained using **Cross-Entropy Loss**:

TABLE II
WATER QUALITY DATASETS FOR AI-BASED PREDICTION

Dataset Name	Description	Source	Year
Water Quality and Potability Dataset [22]	Contains 3,276 water samples with parameters like pH, hardness, solids, and potability labels.	Kaggle	2021
Water Quality Prediction Dataset [23]	Provides daily water quality samples focusing on pH prediction.	UCI ML Repository	2021
Water Quality Dataset for Prediction Tool [24]	Dataset for predicting water quality in rivers, dams, and lakes in India.	HydroShare	2021
USGS Water-Quality Data for the Nation [25]	Provides real-time water quality data from 1.5 million monitoring sites.	USGS Data	2021
Water Quality Dataset (DeepNote) [26]	Contains chemical properties of 3,000+ water samples to determine potability.	DeepNote	2021
Global River Water Quality Dataset [27]	Large-scale dataset with parameters like temperature, pH, dissolved oxygen from multiple rivers.	Humanitarian Data Exchange	2020
India Surface Water Quality Dataset [28]	Contains historical water quality data from Indian lakes and reservoirs.	India Water Portal	2019

$$\mathcal{L} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(\hat{y}_{i,c}) \tag{5}$$

with the **Adam optimizer**, which dynamically adjusts learning rates:

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{v_t}+\epsilon} m_t \tag{6}$$

2) Learning Rate Scheduling:

- **ReduceLROnPlateau**: Decreased learning rate when validation loss stagnated.

E. Model Evaluation and Deployment

The trained model achieved **98.02% accuracy**, with **Precision, Recall, and F1-Score** above 98%.

For deployment:

- **torch.save()** stored the trained model.
- **pickle.dump()** saved the **StandardScaler** for input consistency.

Predictions were generated as:

$$\text{Predicted Class} = \arg\max(\text{Softmax}(f(X))) \tag{7}$$

IV. RESULTS AND DISCUSSION

A. Model Performance Evaluation

The proposed deep learning model demonstrated high accuracy in predicting water quality categories. The final test accuracy reached **98.02%**, with precision, recall, and F1-score values exceeding **98%**. These results in Table III indicate that the model effectively differentiates between Good, Moderate, and Poor water quality classes.

To enhance generalization, a *stratified K-fold cross-validation* technique was employed, ensuring balanced class distributions during training. The incorporation of *batch normalization* improved training stability, while *dropout layers* mitigated overfitting.

TABLE III
PERFORMANCE METRICS OF THE PROPOSED MODEL

Metric	Value (%)
Accuracy	98.02
Precision	98.03
Recall	98.02
F1-Score	98.02

B. Confusion Matrix Analysis

Water quality was accurately determined by the model as shown in Fig:1. Only 17 samples were misclassified. This shows the model's excellent accuracy (98.02%) and dependability in predicting water quality.

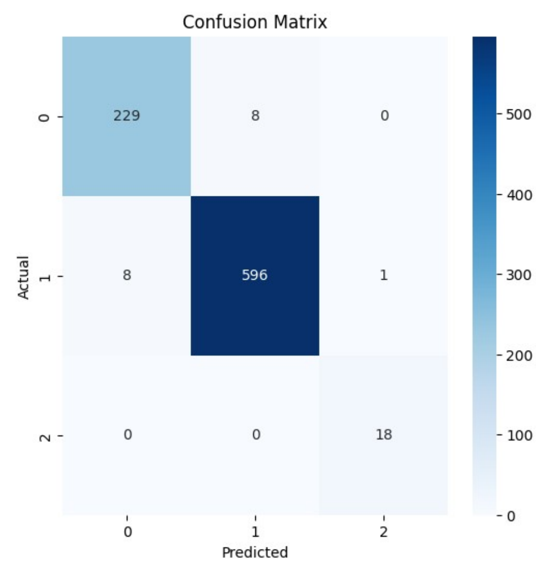

Fig. 1. Confusion Matrix

C. Hyperparameter Optimization Impact

Using **Optuna**, the model's hyperparameters were automatically fine-tuned, leading to optimal values:

- Hidden Layer Sizes: 179, 104, 36 neurons
- Dropout Rate: 0.2586
- Learning Rate: 0.00999

This optimization enhanced the model's convergence speed and stability. The inclusion of the *ReduceLROnPlateau* learning rate scheduler dynamically adjusted learning rates, improving training efficiency.

D. Comparison with Traditional Approaches

Compared to conventional machine learning classifiers, the deep learning approach exhibited superior performance. Traditional models, such as Support Vector Machines (SVM) and Decision Trees, struggle with capturing complex, nonlinear relationships among water quality parameters. In contrast, neural networks effectively learn intricate patterns, leading to improved classification. Table IV shows different model Accuracy Rate.

TABLE IV
COMPARISON OF CLASSIFICATION MODELS

Model	**Accuracy (%)**
Decision Tree	85.42
SVM	88.67
Deep Learning Model	**98.02**

E. Scalability and Deployment Potential

The trained model was deployed using `torch.save()` for model preservation and `pickle.dump()` for saving the standardization scaler. This allows for real-world applications where new data can be processed without retraining. The ability to adapt to different datasets ensures the model's scalability, making it a valuable tool for automated water quality assessment.

V. CONCLUSION

This study built a deep learning model for water quality prediction for aquatic ecosystems using an optimised neural network architecture, achieving a classification accuracy of 98.02%. The model's success is due to Optuna's automatic hyperparameter tuning, which improved the number of neurones, dropout rate, and learning rate, enhancing training efficiency and model generalisation. Stratified K-fold cross-validation enabled a thorough evaluation across several data partitions, reducing overfitting risk. Also, batch normalisation and dropout regularisation improved training stability, reduced vanishing gradients, and sped up convergence. Adaptive learning rate scheduling allowed for dynamic learning rate changes, preventing premature convergence and ensuring optimal weight updates.

This model's scalability allows for easy use in actual water quality monitoring, where variable and dynamic datasets require flexible AI solutions. The suggested solution uses torch.save() for model persistence and pickle.dump() for standardisation preservation to enable real-time prediction with minimum computational overhead. Compared to traditional classifiers like Decision Trees and SVMs, our deep learning model is more accurate and robust in handling complex, non-linear interactions among water quality variables. Future developments may include real-time sensor data, advanced feature selection algorithms, and explainable AI techniques to improve interpretability and decision-making reliability. The AI-based method enables scalable, automated, and precise water quality assessment, helping to manage the environment and public health.

VI. CHALLENGES AND FUTURE SCOPE

- **Data Dependency and Generalization:** Model accuracy depends on data quality, sensor calibration, and regional variations in water composition, requiring domain adaptation techniques.
- **Computational Complexity:** Optuna-based hyperparameter tuning enhances performance but incurs high computational costs, necessitating more efficient optimization strategies.
- **Enhancing Model Robustness:** Incorporating *self-supervised learning* or *attention mechanisms* can further improve feature extraction and classification accuracy.
- **Real-Time and Scalable Deployment:** Future work should integrate *IoT-enabled sensors*, employ *transfer learning* for adaptability, and explore *hybrid CNN-RNN architectures* for spatial-temporal

These advancements aim to improve the accuracy, scalability, and real-time applicability of AI-driven water quality prediction models, enabling more intelligent and autonomous water monitoring systems.

The findings of this research are prominently evident to establish that it is highly feasible and possible to employ AI models to forecast water quality in various aquatic environments. The exceptional 98.02% accuracy rate of these models combined with their robust and impressive performance metrics is conclusive evidence that this new approach is not only possible but also practical in application. Further, this recent research supports and reaffirms findings presented by Haque et al. [13] and Alnemari et al. [14], both of whom also demonstrated that artificial intelligence-based systems may be efficiently employed for the crucial task of environmental monitoring and water quality analysis.

REFERENCES

[1] Y. Duan, H. Li, M. He, and D. Zhao, "A bigru autoencoder remaining useful life prediction scheme with attention mechanism and skip connection," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10 905–10 914, 2021.

[2] G. Ding, W. Wang, and T. Zhu, "Remaining useful life prediction for lithium-ion batteries based on cs-vmd and gru," *IEEE Access*, vol. 10, pp. 89 402–89 413, 2022.

[3] H. Rastegari, R. F. Rahmat, and F. Nadi, "A machine learning approach to ph monitoring: Mango leaf colorimetry in aquaculture." *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 11, 2024.

[4] T. Wilberforce, A. Alaswad, A. Garcia-Perez, Y. Xu, X. Ma, and C. Panchev, "Remaining useful life prediction for proton exchange membrane fuel cells using combined convolutional neural network and recurrent neural network," *International Journal of Hydrogen Energy*, vol. 48, no. 1, pp. 291–303, 2023.

[5] M. Hachicha, R. Ben Halima, and T. Frikha, "Presa: an intelligent blockchain-based platform for monitoring and predicting water quality for smart aquaculture," *Neural Computing and Applications*, pp. 1–12, 2024.

[6] T. Jun, A. Yasin, and N. S. Mansor, "Model design of lstm based on enhanced nadamclip algorithm in prediction of ammonia in aquaculture," *Educational Research (IJMCER)*, vol. 6, no. 6, pp. 73–81, 2024.

[7] K. Yang, Y.-j. Wang, Y.-n. Yao, and S.-d. Fan, "Remaining useful life prediction via long-short time memory neural network with novel partial least squares and genetic algorithm," *Quality and Reliability Engineering International*, vol. 37, no. 3, pp. 1080–1098, 2021.

[8] W. Liu, J. Wang, Z. Li, and Q. Lu, "Issa optimized spatiotemporal prediction model of dissolved oxygen for marine ranching integrating dam and bi-gru," *Frontiers in Marine Science*, vol. 11, p. 1473551, 2024.

[9] M. C. Ogwu, S. C. Izah, A. O. Aigberua, and C. T. Ngun, "Detection, risk analysis and monitoring of chemical contaminants from agro-aqua food production and processing: implications on the one health triad," p. 1501930, 2024.

[10] A. Vaccaro, D. Biadene, and P. Magnone, "Remaining useful lifetime prediction of discrete power devices by means of artificial neural networks," *IEEE Open Journal of Power Electronics*, vol. 4, pp. 978–986, 2023.

[11] M. A. B. Siddique, B. Mahalder, M. M. Haque, and A. S. Ahammad, "Impact of climatic factors on water quality parameters in tilapia broodfish ponds and predictive modeling of pond water temperature with arimax," *Heliyon*, vol. 10, no. 18, 2024.

[12] S. M. Imran, D. Yashaswini, C. Sahasra, and B. Gayathri, "An innovative hybrid approach to forecasting soluble oxygen for optimal water purification in highly concentrated aquaculture," *Turkish Journal of Computer and Mathematics Education*, vol. 15, no. 3, pp. 132–143, 2024.

[13] M. A. Haque, B. Khatun, M. A. S. Jewel, J. Ara, M. S. I. Kazal, and J. Hasan, "Assessment of water quality and heavy metal indices in a tropical freshwater river for aquatic life and public health standard," *Ecological Indicators*, vol. 169, p. 112862, 2024.

[14] A. Alnemari, "Influence of abiotic factors, soil and water characteristics on dissolved oxygen dynamics and fish mortality in saudi arabian aquaculture ponds," *Egyptian Journal of Soil Science*, vol. 65, no. 1, 2025.

[15] Y. Jiang, L. Zhang, C. Wang, L. Chen, W. Zhang, H. Weng, L. Xie, and F. Qu, "Multi-step prediction of dissolved oxygen in fish pond aquaculture using feature reconstruction-based deep neural network," *Computers and Electronics in Agriculture*, vol. 232, p. 109997, 2025.

[16] B. Kol and K. J. Naik, "A deep reinforcement learning-based lstm-enhanced a3c model for predicting bacterial gill disease in aquaculture water," in *2024 International Conference on Decision Aid Sciences and Applications (DASA)*. IEEE, 2024, pp. 1–7.

[17] J. Liu, C. Yu, Z. Hu, Y. Zhao, Y. Bai, M. Xie, and J. Luo, "Accurate prediction scheme of water quality in smart mariculture with deep bi-s-sru learning network," *Ieee Access*, vol. 8, pp. 24 784–24 798, 2020.

[18] W. Liu, J. Wang, Z. Li, and Q. Lu, "A hybrid improved dual-channel and dual-attention mechanism model for water quality prediction in nearshore aquaculture," *Electronics*, vol. 14, no. 2, p. 331, 2025.

[19] F. Chen, T. Qiu, J. Xu, J. Zhang, Y. Du, Y. Duan, Y. Zeng, L. Zhou, J. Sun, and M. Sun, "Rapid real-time prediction techniques for ammonia and nitrite in high-density shrimp farming in recirculating aquaculture systems," *Fishes*, vol. 9, no. 10, p. 386, 2024.

[20] Z. Wang, Z. Zhang, H. Li, H. Jiang, L. Zhuo, H. Cai, C. Chen, and S. Zhao, "Water quality in the ma'an archipelago marine special protected area: Remote sensing inversion based on machine learning," *Journal of Marine Science and Engineering*, vol. 12, no. 10, p. 1742, 2024.

[21] R. Owusu, J. C. Mayoko, and J. Y. Lee, "Prediction of water quality temperature in the growth pattern of fish (nile tilapia) and plant (lettuce) in a prototype-safe, automated aquaponics environment using deep learning," *Open Access Library Journal*, vol. 11, no. 9, pp. 1–22, 2024.

[22] A. Kadiwal, "Water quality and potability dataset," 2021. [Online]. Available: https://www.kaggle.com/datasets/adityakadiwal/water-potability

[23] U. M. L. Repository, "Water quality prediction dataset," 2021. [Online]. Available: https://archive.ics.uci.edu/dataset/733/water-quality-prediction-1

[24] HydroShare, "Water quality dataset for prediction tool," 2021. [Online]. Available: https://www.hydroshare.org/resource/4ab43e1b507b496b9b42749701daed5c/

[25] U. S. G. S. (USGS), "Usgs water-quality data for the nation," 2021. [Online]. Available: https://catalog.data.gov/dataset?tags=water-quality

[26] L. Puttmann, "Water quality dataset (deepnote)," 2021. [Online]. Available: https://deepnote.com/app/leonard-puttmann-a8ef/Water-quality-dataset-14a5f99f-9b81-49c9-9327-973b1c9f0b1f

[27] H. D. Exchange, "Global river water quality dataset," 2020. [Online]. Available: https://data.humdata.org/dataset/global-river-water-quality

[28] I. W. Portal, "India surface water quality dataset," 2019. [Online]. Available: https://www.indiawaterportal.org/data/dataset-india-surface-water-quality