REWAR-Sense: Hepta Sensors Integrated Adaptive Deep Learning Model for Water Quality Monitoring Via Thinkspeak Cloud in Gulshan Lake

Selvi Samiappan [1,*], Asha Mary Ravijothi [2], Revathy Gurumurthy [3], Suganthi Muthuraj [4]

[1] Associate Professor, Department of Artificial Intelligence and Data Science, Builders Engineering College, Nathakadaiyur, Tiruppur Dt. Tamil Nadu, India

[2] Assistant Professor, Department of Information Technology, Adhi College of Engineering and Technology, Sankarapuram, Kanchipuram, Chennai, Tamil Nadu 631605 India

[3] Assistant Professor, Department of Computer Science and Engineering, Srinivasa Ramanujan Centre, SASTRA Deemed University, Kumbakonam, Tamil Nadu, India

[4] Assistant Professor, Department of Information Technology, Kongunadu College of Engineering and Technology, Trichy-621215, India

[*]Corresponding author: Selvi Samiappan

E-mail: selviS75@outlook.com

ABSTRACT

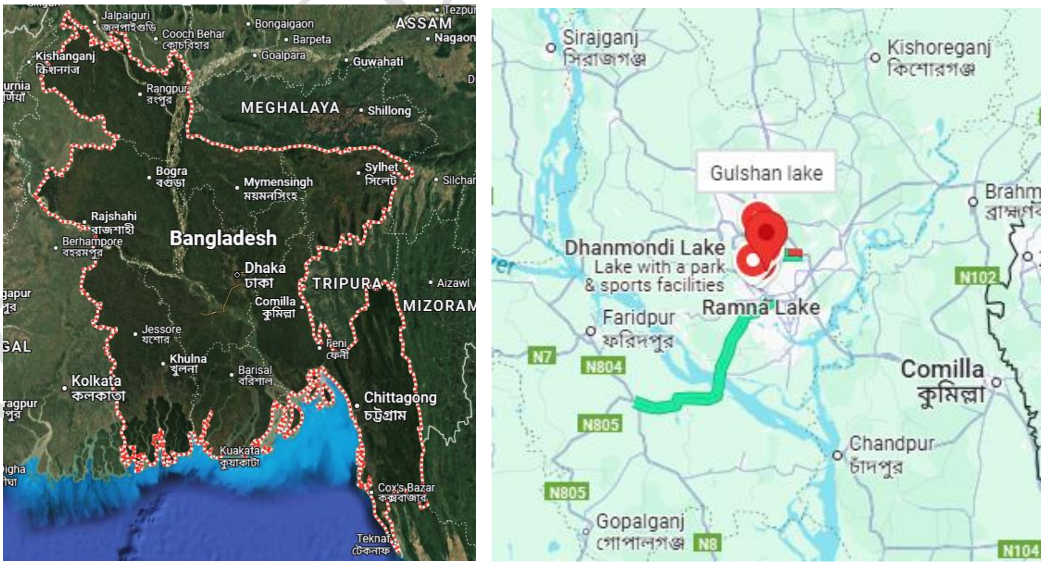
Water Quality Prediction (WQP) plays an essential role in supplying high-quality water to diverse sectors which is dominant for every living organism in the environment. WQP is an important issue that affects both the sustainability of ecosystems and the health of aquatic species. Traditional techniques for determining the quality of water are expensive, time-consuming, and prone to errors. To overcome these issues, a novel REmote WAteR Sensing for quality assessment (REWAR-Sense) methodology is proposed to develop an automated system for the prediction and classification of water quality in Gulshan Lake. Initially, the raw Water Quality (WQ) parameters were gathered from the Gulshan Lake using Hepta sensors and stored them into the ThinkSpeak Cloud for centralized data collection. These gathered data are fed to the preprocessing module to standardize the data. A Deep Learning (DL) Network is employed for feature extraction that identifies the critical patterns of WQ and reduces the data complexity. After feature extraction, a Water Quality Index (WQI) is predicted using an adaptive metaheuristic optimization algorithm that provides a numerical score to indicate the water's condition of the Gulshan Lake. Finally, an attention-based neural network categorizes the WQ into four such categories to enhance the Water Resource Management (WRM) for efficient environmental monitoring. The REWAR-Sense methodology was simulated by using MATLAB and it is validated by Gulshan Lake Dataset. The REWAR-Sense methodology is evaluated based on a number of variables such as accuracy, precision, recall, and F1-score. In comparison, the proposed REWAR-Sense method achieves an accuracy of 93.45%, precision of 92.80%, recall of 93.20%, and F1-score of 93.00% outperforming the existing AutoDL, SOD-VGG-LSTM, and LSTM-CN methods respectively.

Keywords: Gulshan Lake, Water Quality Prediction, Ghost Network, Attention Based Bidirectional Recurrent Neural Network, Adaptive Fish Swarm Optimization.

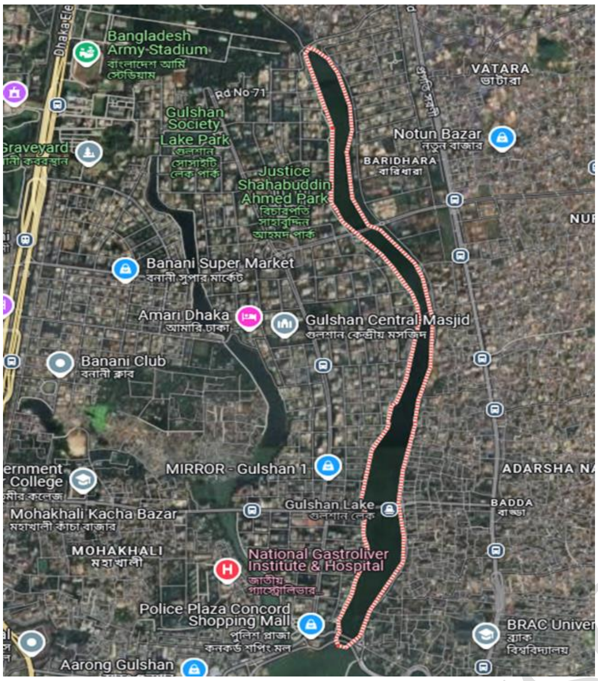
1. Introduction

Water quality monitoring data may play a major role in the efficient management and preservation of water resources (Wu and Wang, 2022; Azrour et al. 2022). Access to safe and clean water is crucial for agriculture, human health, and environmental sustainability which provides long-term benefits that can be achieved by combining Internet of Things (IoT) enabled WQ solutions with modern methods (Deng et al. 2021). In recent years, there has been an increase in interest in using Machine Learning (ML), DL, and the IoT to address environmental challenges (Khan et al. 2022; Iniyan Arasu et al. 2024).

Sensor networks may be installed in water bodies via IoT that gather a number of information on a range of WQ parameters including pH, salinity, temperature, nutrients, and pollutant concentrations (Chen et al. 2021; Zhang and You, 2024). This real-time data collection provides a thorough grasp of the dynamic dynamics of aquatic ecosystems and enables ongoing WQ monitoring (Li et al. 2024; Shams et al. 2024). Decision-makers, managers of water resources, and legislators can use this information to provide the preservation and restoration of a variety of water resources (Venkata et al. 2024; Pang et al. 2024). The Geographical location of the Gulshan Lake is depicted in Figure 1.



(a) (b)



(C)

Figure 1. Geographical Location of Gulshan Lake

Recently used WQP methods mainly include regression analysis, grey system theory, time series forecasting, and Artificial Neural Network (ANN) methods (Nithya et al. 2020; Satish et al. 2024). WQ data exhibit characteristics such as nonlinearity and variability with a strong ability to process nonlinear information are widely used in the field of WQP (Kushwaha et al. 2024; Khullar and Singh 2022). This indicates that fusion models based on DL networks have greater advantages in WQP. However, the aforementioned prediction models do not adequately consider the varying importance of features in long-time series where highly important features often have a greater impact on the model's prediction performance (Rajeswari et al. 2020, Krishna Bikram Shah et al.2023, Santhiya Govindapillai and Radhakrishnan Ayyapazham. 2024).

WQ forecasting research is one such important topic of concern. But a growing global issue is the deteriorating WQ caused by pollution, population growth, and climate change (Díaz-González et al. 2025). The expense, time commitment, and inability to gather data in real-time are the main drawbacks of conventional methods for WQ monitoring. River WQ variations indicate both gradual shifts and unpredictable non-linear processes. As a result, the process of appropriate WQP becomes more difficult (Rajanbabu et al. 2025). WQ forecasting is also essential for planning and managing water resources and their surroundings. Based on the expected outcomes the issue of water contamination can be foreseen and enable an early effort to avert its effects (Geethamani et al. 2023). To overcome these issues, a novel REWAR-Sense method is proposed to forecast the WQ of Gulshan Lake using advanced sensors and DL techniques. The major contributions of the proposed REWAR-Sense methodology are given as follows.

- The goal of the REWAR-Sense methodology is to develop a robust WQ prediction in Gulshan Lake for real-time environmental monitoring in WRM.
- The raw environmental data related to WQ parameters are gathered from Gulshan Lake using Hepta sensors which are stored in the ThinkSpeak Cloud and preprocesses the data by handling the missing values and removing the irrelevant errors to standardize the data.
- The ghost network extracts significant WQ features from the data and the Adaptive Fish Swarm Optimization (AFSO) algorithm produces a precise WQ index by quantifying the conditions through numerical scores.
- The Attention based Bidirectional Recurrent Neural Network (ABiRNN) categorizes the WQ into potable water, palatable water, contaminated water, and infected water for efficient WRM.
- The performance of the REWAR-Sense methodology is validated through metrics such as accuracy, precision, recall, and F1-score.

Motivation

Water quality plays an indispensable role in sustainable WRM and environmental protection. Numerous variables impact water quality and in turn its suitability for human consumption such as mining, industrialization, pollution, and natural occurrences. Current WQ methods gather water samples manually and examine the physical, chemical, and biological agents to identify the type of WQ. These approaches' shortcomings include limited space or time range, time consumption, greater system costs, and insufficient real-time WQ assessment. These factors are the motivation behind this research which is intended to develop a novel REmote WAteR Sensing for quality assessment (REWAR-Sense) methodology to address these complexities. The proposed REWAR-Sense methodology aims to provide an automated system for the prediction and classification of WQ in Gulshan Lake.

Objectives

The major objective of this research is to deploy an innovative automated WQ monitoring system to enhance the sustainability of ecosystems which is tailored for Gulshan Lake. The REWAR-Sense methodology reduces the complexity of the data through feature extraction and the WQI is predicted to identify the water conditions of the Gulshan Lake. The REWAR-Sense methodology categorizes the WQ into such categories to enhance the management of the water resources. These analyses are drawn out by the research through various techniques of WQ analysis on Gulshan Lake employing DL networks and an adaptive metaheuristic optimization algorithm. These research methodology is cost-effective, energy sustainable, reliability of data transmission, less time delay, high network coverage, and sensor accuracy.

The remainder of the research is organized as follows: The related research for detecting WQP is provided in Section 2. The recommended REWAR-Sense methodology for WQP is covered in Section 3. The experiment results of the REWAR-Sense methodology is described in Section 4. Section 5 concludes the REWAR-Sense methodology with future enhancement.

2. Literature Survey

Numerous studies that have examined to efficiently monitor and manage, and forecast WQ have focused on the intersection of IoT, DL, and ML technology. The related research discusses and highlights the following pertinent works and contributions below.

In (Prasad et al. 2022) suggested WQP and assessed using both DL and Auto-DL techniques. For both binary class and multiclass water data conventional DL outperforms AutoDL by 1.8% and 1% respectively. While the accuracy of the traditional model achieves 98% to 99%, the accuracy of the AutoDL approach achieves 96% to 98% respectively.

In (Islam and Irshad 2022) suggested a DL-enabled categorization and WQP model for artificial ecosystem optimization. The suggested AEODL-WQPC method predicts the WQI using an Optimal Stacked Bidirectional Gated Recurrent Unit (OSBiGRU) model and classifies WQ using an AEO with an enhanced Elman Neural Network (AEO-IENN) model. Validated on a WQ dataset, the AEODL-WQPC strategy outperforms more recent state-of-the-art techniques.

In (Wan et al. 2022) suggested a model that tackles WQP caused by pollution from non-point sources using feature extraction and DL methods. When the suggested SOD-VGG-LSTM approach was applied, the Lijiang River watershed showed the largest relative differences between the expected and observed values for DO, CODMn, NH3-N, and TP. It consists of 8.47%, 19.76%, 24.1%, and 35.4% of errors respectively. The SOD-VGG-LSTM's R2 was between 32% and 39.3% greater than that of the ARIMA, SVR, and RNN.

In (Talukdar et al. 2023) suggested lake WQ indicators using DL methods based on sensitivity-uncertainty analysis. The suggested approach forecasts the WQI by combining the models of the Generalized Linear Model (GLM), Neural Network (NN), and Gradient Boosting Machine (GBM). The water samples were found to have poor to very poor quality as indicated by their WQI which varied from 90.75 to 145.29. This model outperformed the existing models with a prediction accuracy of 25.77, RMSE of 5.07, MAE of 3.5, and R2 of 0.98 respectively.

In (Rahu et al. 2023) suggested frameworks for WQ analysis and prediction enabled by ML and the IoT. To gather data from Rohri Canal, SBA, Pakistan, the IoT framework is outfitted with sensors for temperature, pH, turbidity, and Total Dissolved Solids (TDS). According to the data, the SVR model has the lowest R-squared at 0.73, while the MLP regression model has the greatest at 0.93. The Random Forest algorithm has the best accuracy, precision, recall, and F1-score of 0.91, 0.93, and 0.92, respectively among classification techniques.

In (Chhipi-Shrestha et al. 2023) suggested Applications of Artificial Intelligence (AI) and soft computing to assess the quality of drinking water. The adaptive neuro-fuzzy inference system, multilayer perceptron-based ANN, support vector machines, Bayesian networks, and general regression neural networks are some of the AI and SC approaches used in the digital water method to effectively monitor WQ. AI's and SC's primary roles in the suggested digital water were to model physicochemical and microbiological factors and assess the water's quality respectively.

In (Mahesh et al. 2024) suggested WQP effectively manages water by integrating a normalizer with LSTM. While maintaining the intrinsic properties of the data the suggested LSTM-CN model incorporates normalization calculation techniques for adaptive processing of multi-factor data. To learn the properties of the data and produce precise prediction results, the LSTM-CN model works in tandem with the codec. The suggested LSTM-CN approach produces 99.3% accuracy, 95% precision, 18.0% MSE, 11.45% RMSE, and 93.6% recall respectively.

In (Venkatraman et al. 2023) suggested The logistic Giant Armadillo Optimization (GArO) deep differential recurflownet is used to forecast and classify WQ with precision. An Optimization driven Deep Differential RecurFlowNet (ODD-RecurFlowNet) is suggested to predict and classify WQ. The ODD-RecurFlowNet approach produces an overall accuracy of 98.01% and an RMSE value of 0.039 using a standard dataset for WQ.

In (Pavan kalyan et al. 2024) suggested An Analysis of Support Vector Machine (SVM) and Decision Tree (DT) Methods for Predicting Tomato Growth and Yield in Hydroponic Systems Using Deep Water Culture (DWC). In the suggested approach, the efficacy of SVM and DT methods in hydroponic tomato production is assessed using the DWC method. In contrast, the suggested approach provides more accuracy with SVM.

In (Raveena et al. 2024) suggested Coffee crop irrigation systems are continuously monitored and optimized using recycled water and bi-directional RNNs and IoT sensors. The recommended technique collects data on soil moisture, weather, WQ, temperature, humidity, pH, and nutrient value. In terms of irregularity and watering schedules, the recommended method produces an accuracy of 95.66% respectively.

In (Li et al. 2024) suggested an analysis of the modernization and transformation of manufacturing firms using a four-way game and industrial internet platforms. A revenue sharing contract coefficient guarantees the steady growth of the suggested approach and ongoing collaboration. The model gives platforms, manufacturing companies, governments, and developers a theoretical foundation for choosing a strategy.

In (Wang and Ma, 2024) suggested a study on the connection between rising carbon emissions and the expansion of inclusive digital banking. The suggested approach examines carbon emissions and digital inclusive financing are related in Chinese cities between 2011 and 2022. By contrast, the suggested approach shows that carbon emissions can be reduced by 0.311% for every 1% expansion and that China's digital inclusive finance index has increased since 2011.

In (Wang et al. 2024) suggested an examination of the regions in China that produce the most energy in terms of carbon emissions. The suggested approach forecasts carbon emissions from 2021 to 2040 using an open STIRPAT model. The study also highlights the importance of controlling per capita GDP and energy consumption for effective emission reduction strategies.

In (Suresh Maruthai et al. 2025) suggested Real-time monitoring by combining HG-RNN with IoT sensor vision and wastewater recycling. The suggested approach uses IoT sensors to efficiently clean and monitor contaminated ponds and turn them into sources of pure water. To deliver the best possible WQ while avoiding pollution, the HG-RNN algorithm predicts WQ parameters, examines future trends, and incorporates real-time treatment decisions.

In (Zhang et al. 2025) suggested a consideration of heterogeneity and variable interaction in the relationship between artificial intelligence (AI) and digitization (D&AI) and carbon emissions. The suggested Decision Deep and Cross Feature-Transformation Network (DDCFTN) analyzes the carbon impacts of urban emissions. The suggested model works better than the traditional models and demonstrates that the influence of interacting effects exacerbates the overestimated contribution of D&AI to carbon emissions.

In (Wu et al. 2025) suggested the impact of green finance regulations on the ESG performance of construction firms. The suggested approach states that by setting financing caps and promoting the advancement of green technologies, the green credit policy greatly improves ESG performance.

These findings are particularly significant among smaller and non-state-owned firms. The suggested approach enhances the ESG performance and reduces the environmental risks.

In (Zeng et al. 2025) suggested a multi-scale spillover and a tail risk contagion between the top US technology shares and the green finance index. The proposed method identifies Microsoft and the S&P 500 ESG index as the primary risk sources, and the net risk spillover characteristics show fluctuation and cyclicality. According to these findings, volatility connectedness increases in beneficial market conditions and is stronger at extreme tails.

The research evaluation states that people forecast WQ and offer alerts on potential ecological contamination based on past environmental indicators. Determining the quality of water is difficult because of the complicated data. The decline of the surface water ecosystem exacerbates these problems. Predicting and monitoring surface WQ is essential. To overcome these issues, a novel REWAR-Sense methodology has been proposed to predict the WQ of Gulshan Lake using DL techniques.

3. The REWAR-Sense Methodology

In this section, a novel REmote WAteR Sensing for quality assessment (REWAR-Sense) methodology has been proposed to develop an automated system for the prediction and classification of WQ in Gulshan Lake. Initially, the raw environmental data related to WQ parameters are gathered from Gulshan Lake using Hepta Sensors such as Total Dissolved Solids (TDS) Sensor, Dissolved Oxygen (DO) Sensor, Total Organic Carbon (TOC) Sensor, Temperature Sensor, Turbidity Sensor, pH Sensor, and Chlorophyll Sensor to monitor various physical, chemical, and biological parameters in real-time over a specific period. Several Internet of Things protocols and wireless technologies are employed to transmit these data directly to the ThinkSpeak Cloud for storage. These gathered data are fed to the preprocessing module to attain an accurately reliable formatted data by using Handling Missing Values, Data Standardization, and Data Cleaning for feature extraction. The ghost network extracts significant features related to the WQ and reduces data dimensionality to ensure efficient WQ prediction. After feature extraction, the WQI is predicted by an AFSO optimization algorithm that represents whether the quality of water is excellent, fair, or poor through numerical scores. Finally, an Attention based BiRNN categorizes the WQ into respective categories such as potable water, palatable water, contaminated water, and infected water for accurate real-time environmental monitoring in WRM. The overall workflow of the proposed REWAR-Sense methodology is depicted in Figure 2.

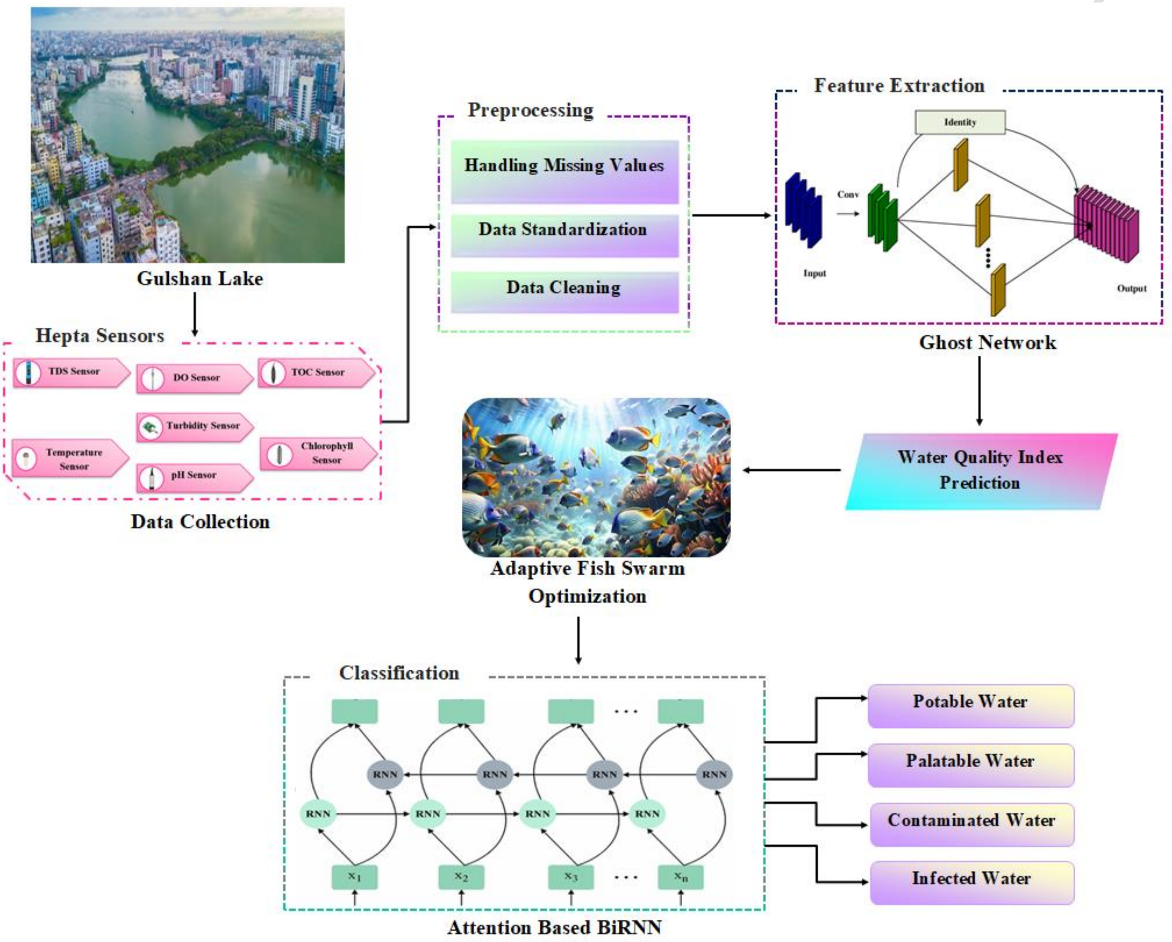


Figure 2. Proposed REWAR-Sense Methodology

3.1. Data Collection (Hepta Sensors)

The REWAR-Sense system utilizes the interconnected devices and Hepta Sensors such as Temperature sensor, Turbidity sensor, pH sensor, TDS sensor, Chlorophyll sensor, Dissolved Oxygen sensor, and TOC sensor deployed along the Gulshan Lake to automate data collection. These Hepta Sensors would continuously measure parameters such as temperature, turbidity, pH, concentration of dissolved solids, chlorophyll concentration, DO concentration, and TOC contents from the lake. The Hepta Sensor measures the temperature level of the water for concerning the ecosystem health and turbidity level for detecting the amount of pollutant levels in the water. In the Hepta Sensor, the alkalinity of water is determined using a pH sensor, and the concentration of dissolved ions is measured by a TDS sensor. The Hepta Sensor monitors the chlorophyll levels as well as oxygen levels in water which is vital for aquatic organisms and detects the pollution from organic matter. The Hepta Sensor quantifies the organic carbon content to provide insights into water pollution and decomposition levels. Several IoT protocols and wireless technologies enable the ThinkSpeak Cloud to store this processed data.

3.2. Data Preprocessing

Those data gathered by using the Hepta Sensor are fed to the preprocessing module to attain accurately reliable formatted data by using Handling Missing Values, Data Standardization, and Data Cleaning for feature extraction.

Handling Missing Values:

Due to regular maintenance of monitoring stations and occasional equipment failures, some WQ data may be missing. To ensure the validity of the experiment, complete data must be provided to the prediction model. It uses linear interpolation to handle missing data. The formula for linear interpolation is represented in Equation (1),

$$y_k = y_\omega + (y_r - y_\omega)\frac{k-\omega}{r-\omega} \tag{1}$$

In the formula, k, ω, and r represent time, y_k denotes the missing value at time k, y_ω represents the known data corresponding to the most recent time ω before y_k, and y_r denotes the known data corresponding to the most recent time r after y_k.

Data Standardization:

The mean of the rescaled features is zero, and their standard deviation is one. Outlier features have the potential to substantially skew distributions. The influence of outliers is lessened by standardization since it focuses on the distribution. Since feature coefficients are all normalized to the same scale. This method can be used to determine a feature's importance. The Mean standardization uses the following Equation (2),

$$X_{Standardized} = (X - m)/sd \quad (2)$$

Where m is the mean, x is the starting value, and sd is the standard deviation. Gaussian normalization which fits a Gaussian distribution, and scaling by interquartile range are two further standardization techniques.

Data Cleaning:

Outliers are eliminated from the data preparation framework through a process known as data cleaning. Data points that exhibit a significant departure from the norm are known as outliers which distort statistical analysis and model training. Finding and managing outliers is necessary to improve the reliability and quality of the data.

3.3. Feature Extraction Using GhostNet

A deep learning based GhostNet framework was implemented to extract the features from the preprocessed data for WQP. A ghost module has been built in the CNN network as shown in Figure 3, that extracts multi-scale bottom-level features to increase the feature utility and reduce the network capacity. In order to identify the features from the inputs while maintaining the correlation between preprocessed data this network initializes random distributions. The layer's input volume which is represented as $M_{lf-1} \epsilon \mathbb{Z}^{K_{lf-1} * K_{lf-1} * N_{lf-1}}$ related to the bias vector and an arbitrary feature $R_{lf} \epsilon \mathbb{Z}^{K_{lf} * K_{lf} * N_{lf}}$ defines the locations of identified features within the input data which is formulated in Equation (3).

$$R_{lf} = D_{lf} * M_{lf-1} + bias_{lf} \quad (3)$$

When applying convolutional kernels to input data Equation (3) indicates that each input element-wise product with the filter weight is dumped into the local receptive field.

In $a = \hat{a} - [K_{lf-1}/2]$ and $= \hat{b} - [K_{lf-1}/2]$, the spectral indexes are represented as $\hat{s_i}$and s_i, and the indexes along the spatial proportions of weights are $a, b, \hat{a}$ and $\hat{b}$ is represented in Equation (4).

$$R_{lf}^{x,y,s_i} = \sum\nolimits_{\hat{x}\hat{y}\widehat{s_{lf}}} D_{lf}^{\hat{x},\hat{y},\hat{s_i}} * \ M_{lf-1}^{x+\hat{x},y+\hat{y},\widehat{s_{lf}}} + bias_{lf}^{m_f} \quad (4)$$

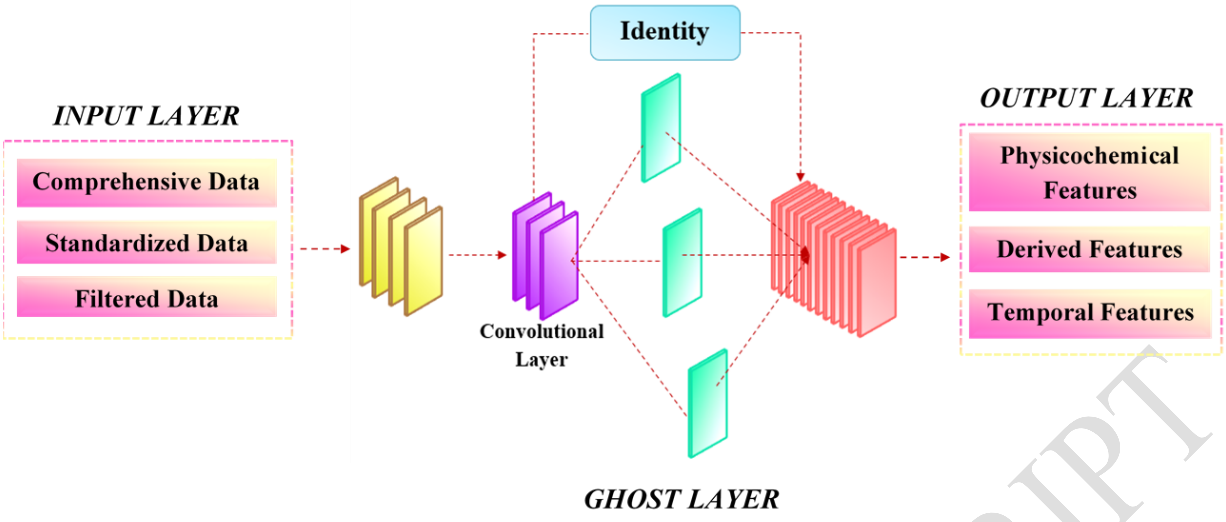


Figure 3. Architecture of Ghost Network

As a result, the nonlinearities of the data are learned using a non-linear activation function A(.) which yields the final output feature maps as $M_{lf} \epsilon \mathbb{Z}^{K_{lf} * K_{lf} * N_{lf}}$ is represented in Equation (5).

$$M_{lf} = A(R_{lf}) \tag{5}$$

Where the ReLU function which is typically employed in backpropagation methods is applied as A. Ghost convolution uses fewer variables and less computational power to produce redundant data. Intrinsic features $\widetilde{M}_{lf}$ are updated in a few simple ways as the output features M_{lf} are developed as "ghosts." $M_{lf} \epsilon \mathbb{Z}^{K_{lf} * K_{lf} * \widetilde{N}_{lf}}$ is the group name for these intrinsic feature maps which are generated by a primary convolution from Equation (5). Furthermore, all of the features are combined and vectorized by the pooling module, which then delivers the result to the WQI prediction module. The ghost network extracts significant features related to the WQ and reduces data dimensionality to ensure efficient WQ prediction.

3.4. Water Quality Index Prediction via Adaptive Fish Swarm Optimization

After feature extraction, the WQI is predicted by an AFSO optimization algorithm to represent the quality of the water. The group of fish is the individual, and the hunting space is the search space. The model begins with a set of populations based on member distribution. There are two types of the suggested routing protocol which are blocks and chasers. Equation (6)-(7) formulates the initialization step,

$$q_j^i = rand.\left(b_i^{high} - b_i^{low}\right) + b_j^{low} \tag{6}$$

$$i = 1,2, \ldots n; j = 1,2, \ldots, m,$$

$$e(c_r) = \sum_{q_{f \in C_r}} \left\| Q_f - \mu_r \right\|^2, f = 1,2, \ldots g; \tag{7}$$

$$r = 1,2, \ldots$$

Using $rand$ to define the random number, which is in the interval [0,1]. Following the aforementioned methods, the entire population Q is divided into discrete groups, or subpopulations, whose behavior can be modeled separately. Equation (8)-(10) displays the mean square error between the cluster μ_r and data points. Fish population Q is the initial data.

$$E(C) = \sum_{r=1}^{0} e(C_r) \tag{8}$$

$$\Phi_r^{s+1} = \Phi_r^s + \alpha \oplus Levy(\beta), 0 < \beta \leq 2 \tag{9}$$

$$\beta = (E(C) \times 0.099) + \frac{0.001s}{s_{max}/10} \tag{10}$$

In order to determine whether the prey has been moved, it will submerge itself in the crack and explore multiple nooks. On occasional walks, the C_r will shift its position and search for any crevices where prey may be hiding. The new position is then determined using Equation (4). Since α = 1 and ⊕ is the entry-wise multiplication, α indicates the step size in this method. An algorithm for WQI prediction via AFSO Algorithm is derived in Algorithm 1.

Algorithm 1: WQI Prediction via AFSO Algorithm

Input: Physiochemical Features, Derived Features, and Temporal Features

Output: WQI Score

1. Initialize random fish population (Q) within bounds
2. Cluster each fish groups (C_r) and evaluate the cluster centroids (μ_r) via MSE
3. Update the positions of each fish using Levy flight

   $$\Phi_r^{s+1} = \Phi_r^s + \alpha \oplus Levy(\beta), 0 < \beta \leq 2$$

4. Assess the fitness function for each fish using WQI
5. Update the best position (Φ_{best}) of the fish according to the low error rate

6. Modify the positions of chaser and blocker fish based on its spiral motion and average movement
7. Return WQI score

The tail regulates the distribution probability, where β is the Levy index. This can be represented using Equation (11)-(14).

$$T = \prod \alpha \oplus levy(\beta) \sim \alpha \left(\frac{u}{|v|^{1/\beta}} \right) (\Phi_r^s - \Phi_{best}^s) \tag{11}$$

$$u \sim M(0, \sigma_u^2) \tag{12}$$

$$v \sim M(0, \sigma_v^2) \tag{13}$$

$$\Phi_r^{s+1} = \sum \Phi_l^s + T \tag{14}$$

Where, T is the randomly selected step. The u and v stand for the normal distribution as per Equation (15). The Equation (16), is used to assess the fitness of the CF at the new sites.

$$\Phi_{best}^{s+1} = \sum \Phi_{best}^s + \prod T' \tag{15}$$

$$T' = \sum_{u=0}^{n} \alpha \left(\frac{u}{|v|^{1/\beta}} \right) \tag{16}$$

$$\varphi_f^{s+1} = Z_f . e^{bp} . cos 2\pi p + \Phi_r \tag{17}$$

The value of T' is provided by Equation (17). A logarithmic spiral represents the Blocker Fish's (BF) movement. They always follow the logarithmic spiral motion of BF, which may be found in Equation (18)-(20).

$$Z_f = \left| l . \Phi_r - \varphi_f^s \right| \tag{18}$$

$$\{\Phi_r, \varphi_f^s\} \in C_r \tag{19}$$

$$q_f^{s+1} = \frac{\Phi_{best} + q_f^s}{2} \tag{20}$$

The number that breaks the distance of Zf in $[-1,1]$ is l. A new location will be chosen to find new prey after the search space is fully occupied. In these situations, the AFSORP model analyzes overexploitation using the λ parameter. The following Equation (21)-(24) has been used to determine the WQI.

$$WQI = \frac{\sum_{i=1}^{N} q_i \times w_i}{\sum_{i=1}^{N} w_i} \quad (21)$$

$$q_i = 100 \times \left(\frac{V_i - V_{Ideal}}{S_i - V_{Ideal}}\right) \quad (22)$$

$$w_i = \frac{K}{S_i} \quad (23)$$

$$K = \frac{1}{\sum_{i=1}^{N} S_i} \quad (24)$$

Where wi is the unit weight for each parameter as determined by Equation (21), N is the total number of parameters used in the WQI computations, and qi is the quality rating scale for each parameter i, as specified by Equation (24). The AFSO algorithm represents that the quality of water is excellent, fair, or poor through numerical scores.

3.5. Water Quality Classification using ABiRNN

An Attention based BiRNN categorizes the WQ for accurate real-time environmental monitoring in WRM. The ability to extract temporal correlations and contextual information from input data makes this method especially suitable for WQ classification. An attention strategy could increase accuracy and reduce noise from irrelevant data by concentrating on the most crucial components is represented in Equation (25). The architecture of the ABiRNN structure is depicted in Figure 4.

$$h_t = f(Ux_t + Wh_{t-1} + b) \quad (25)$$

In the following equation, f is the nonlinear activation function which is used to find the hidden state ht of the RNN at time t. BiRNN uses the forward and backward RNNs which is represented in Equation (26)-(27),

$$\vec{h}_t = \vec{f}\left(\vec{U}_{xt} + \vec{W}\vec{h}_{t-1} + \vec{b}\right) \quad (26)$$

$$\overleftarrow{h}_t = \overleftarrow{f}\left(\overleftarrow{U}x_t + \overleftarrow{W}\overleftarrow{h}_{t+1} + \overleftarrow{b}\right) \quad (27)$$

These represent the trainable parameters such as $\vec{U}$, $\vec{W}$, $\vec{b}$, $\overleftarrow{U}$, $\overleftarrow{W}$, and $\overleftarrow{b}$. The nonlinear activation functions are $\vec{f}$ and $\overleftarrow{f}$. By analyzing the x_t to x_1 series, the reverse RNN generates the backward hidden layers ($\overleftarrow{h}_1,\cdots,\overleftarrow{h}_t$) which is represented by using Equation (28).

$$h_t = \left[\vec{h}_t^T; h_t^T\right]^T \tag{28}$$

After evaluating a set of forward hidden statistics ($\vec{h}_1,\cdots,\vec{h}_t$), the forward RNN analyzes the input series from x_1 to x_t.

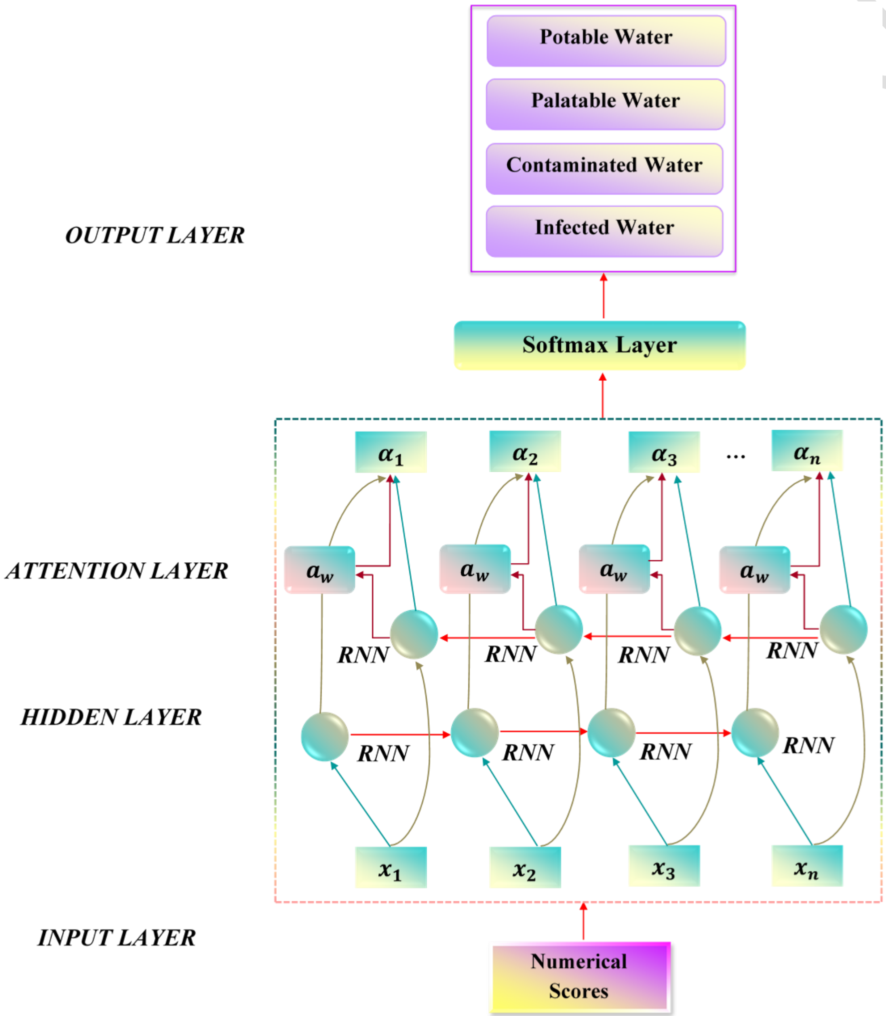


Figure 4. Architecture of ABiRNN Network

By concatenating the backward hidden layers with $\vec{h}_t$ and $\overleftarrow{h}_t$, the ht hidden layer of BiRNN at time t is developed. Using Equation (29)-(30), where ct is the output and $x_t \ldots x_1$ is the input series that finds the attention module's output at time t.

$$C_t = \sum_{k=1}^{T} \alpha_t^k \, h_k \tag{29}$$

$$\alpha_t^k = \frac{exp(\hat{\alpha}_t^k)}{\sum_{j=1}^{T_x} exp(\hat{\alpha}_t^j)} \tag{30}$$

If the weight of the h_k hidden layer is α_t^k, then C_t and αk reflect the weighted total of the RNN's hidden states ($h_1, \cdots, h_t$). Equation (18) represents α_t^k in conjunction with other module components. Finally, the ABiRNN Classifies the quality of the water into respective categories such as potable water, palatable water, contaminated water, and infected water for for efficient environmental monitoring.

4. Result and Discussion

This section discusses the results of classifying the WQ using the proposed REWAR-Sense methodology. In the REWAR-Sense methodology, the data are collected through Hepta sensors including pH, DO, TOC, TDS, turbidity, temperature, and chlorophyll. The REWAR-Sense methodology was simulated by using MATLAB R2023a and it is validated by Gulshan Lake Dataset. The investigation makes use of an Intel i7 processor, 8 GB of RAM, and a Windows 10 OS system. The real-time sensor data are stored and visualized by using the ThingSpeak cloud platform. A comparison is made between the proposed REWAR-Sense methodology and existing methods such as AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024), according to the metrics including accuracy, precision, recall, specificity, F1-score, WQP Time, MSE, RMSE, and Computational Time.

4.1. Dataset Description

Gulshan Lake is located in Dhaka, which is Bangladesh's northernmost city. Gulshan Lake is considered to be one of the primary sources of surface water in these areas. The entire surface area of Gulshan Lake is around 100 hectares, and it is 3.8 kilometers long. The Environment Department (DOE) and Bangladesh's Environment and Forest Ministry provided these samples. Where, the data are gathered from the Gulshan Lake through a Total Dissolved Solids (TDS) Sensor, Dissolved Oxygen (DO) Sensor, Total Organic Carbon (TOC) Sensor, Temperature Sensor, Turbidity Sensor, pH Sensor, and Chlorophyll Sensor. In 2023, the monthly measurements were made of the WQ factors. The dataset used in this investigation contained 108 specimens. Based on the WQI prediction the WQ of the Gulshan Lake is categorized into potable, palatable, contaminated, and infected classes comprising around 25% even distribution among the Gulshan Lake dataset. This distribution across these classification phases employed a stratified partition with a 60:20:20 ratio of training, validation, and testing inputs. A 3-fold cross-validation strategy is employed to effectively assess the model's generalization ability and reduce the risk of overfitting. This operation was repeated three times to ensure that each part had an opportunity to serve as the validation set. Through this approach, all the training data has been used for model training and evaluation thereby avoiding information wastage due to data partitioning. The parameter of epoch was set to 100 and the sample number per batch was set to 20. The Gulshan Lake dataset with epoch 100 was studied and their Val-Loss was calculated which is shown in Figure 6. During transitional climate periods the monthly data may not reflect daily or seasonal fluctuations. Therefore, the data preprocessing phase addresses these potential biases through data cleaning, handling missing values, and data standardization. The data preprocessing phase standardizes the data to balance these seasonal fluctuations. In this context, the Gulshan Lake dataset achieves superior results for predicting and classifying the overall WQ.

4.2. Performance Analysis

The diverse methods used by these models resulted in varying assessments of the results they generated as,

$$MSE = \frac{1}{n}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2 \quad (31)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2} \quad (32)$$

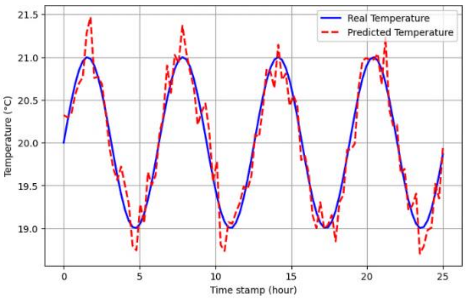
$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \quad (33)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (34)$$

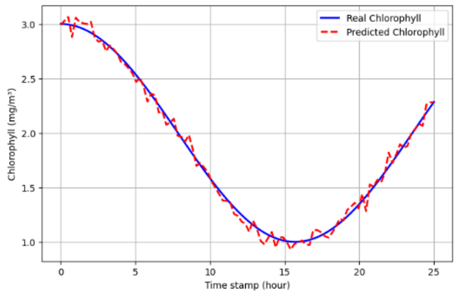
$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (35)$$

$$F1\ score = 2 \times \frac{precision \times Recall}{Precision + Recall} \quad (36)$$

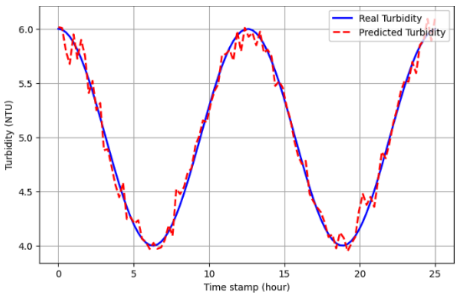
However, by quantifying the mean absolute error between the predicted and actual values using the MSE and RMSE, equations (31)–(32) show the divergence between the expected and actual values that are susceptible to outliers. Equation (33)–(36) shows the metrics of accuracy, precision, recall, and F1 score that were used to assess the models' performance.



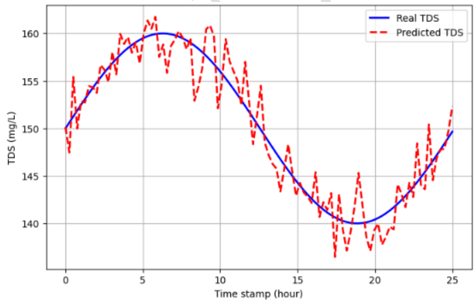
(a) Temperature Prediction



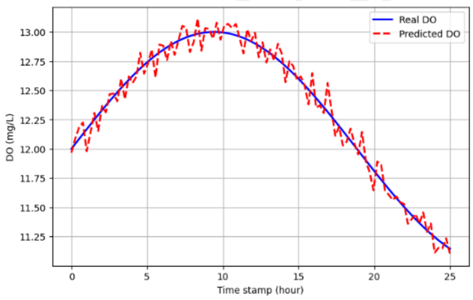
(b) Chlorophyll Prediction



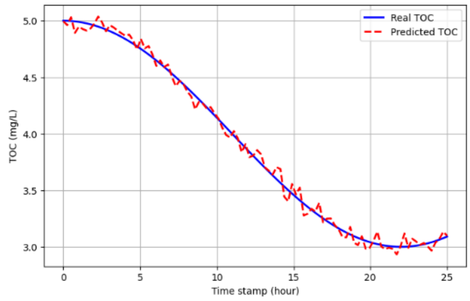
(c) Turbidity Prediction



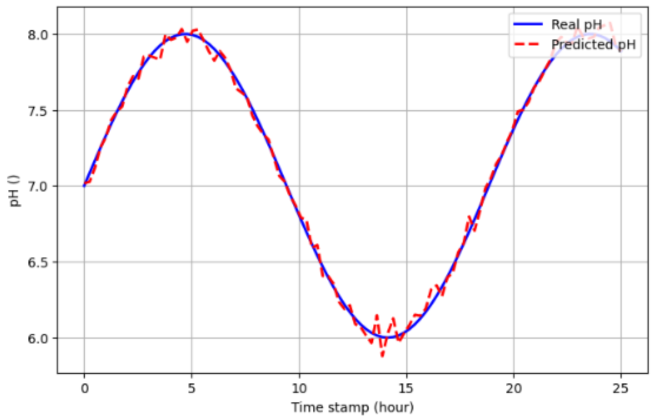
(d) TDS Prediction



(e) DO Prediction



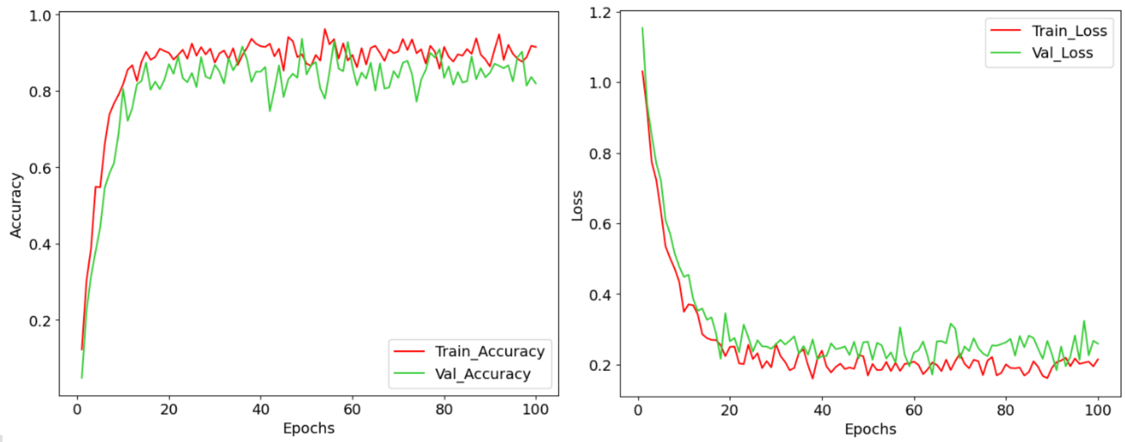
(f) TOC Prediction



(g) pH Prediction

Figure 5. Prediction Effect of Hepta Sensors

The prediction effect of the Hepta sensors for WQP is depicted in Figure 5. Each sensor monitors a specific parameter critical for assessing WQ. The purpose of this prediction using Hepta sensors is to enable real-time and accurate assessment of WQ for environmental monitoring, and public health protection. The advantage of using Hepta sensors for WQP provides continuous real-time data which enhances the responsiveness of monitoring systems. The integration of multiple parameters improves decision-making and supports sustainable WRM to ensure ecological balance and public safety.



(a) (b)

Figure 6. Accuracy and Loss curve for REWAR-Sense Method

The proposed REWAR-Sense system is a model for characterizing the quality of the water. The classification outcomes and accuracy of the validation data set for each model are displayed in Figure 6. The accuracy and loss ratio increase correspondingly when the verification data employs only quality data. The experimental result shows that improved accuracy and loss as well as graph stabilization for the proposed REWAR-Sense model.

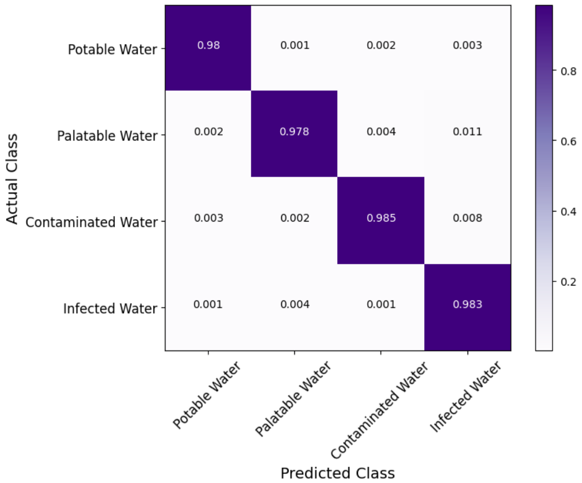


Figure 7. Confusion Matrix

The four-class classification challenge for WQ detection may have resulted in a significant classification error because of the increased inter-class confusion. Most inputs are expected to fall into the category of WQ classification, as the confusion matrix in Figure 7 illustrates. The differences between the closely related forms of WQ, such as potable water, palatable water, contaminated water, and infected water, might not always be clear in contrast to the WQ detection tasks. In line with the results, the confusion matrix shows the most frequently predicted category in WQ detection.

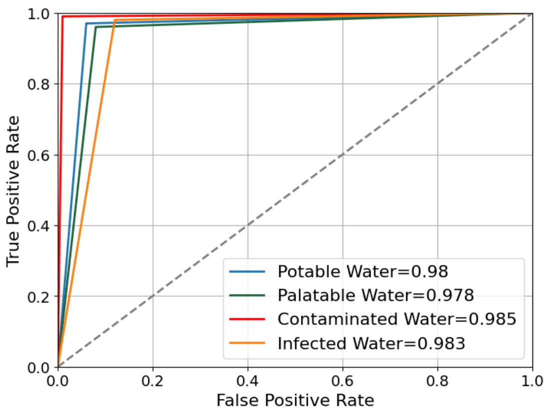


Figure 8. ROC Curve for REWAR-Sense Method

The ROC curves for the REWAR-Sense model are plotted in Figure 8 which further illustrates its classification performance. The ROC curves reveal that the contaminated water has the largest AUC followed by potable water, infected water, and palatable water with AUC values of 0.980, 0.983, and

0.978. This indicates that these REWAR-Sense models have strong classification abilities for the Gulshan Lake dataset and can effectively differentiate between positive and negative class samples.

4.3. Comparative Analysis

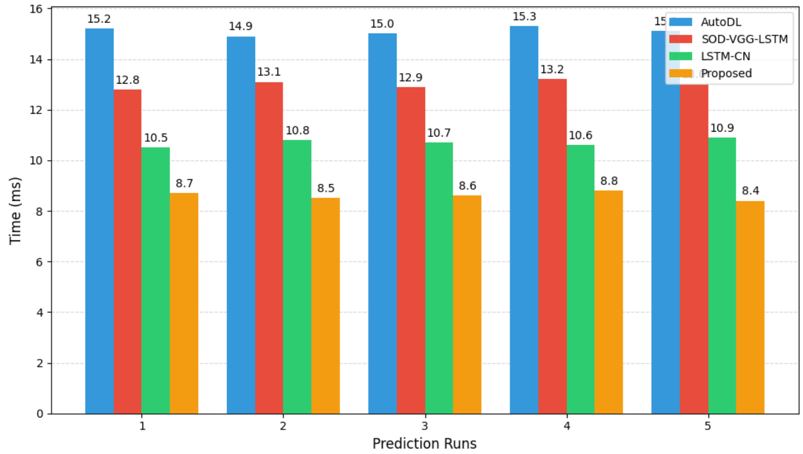


Figure 9. Comparison of WQP Time

The WQP time consumption for different methods of comparison is shown in Figure 9. Using all of these methods, the parallel platform produces the average time after running the experiment times. The Proposed REWAR-Sense approach achieves a WQP time improvement of approximately 43.7%, 34.8%, and 20.4% compared to the existing approaches AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024), respectively on average across all prediction runs.

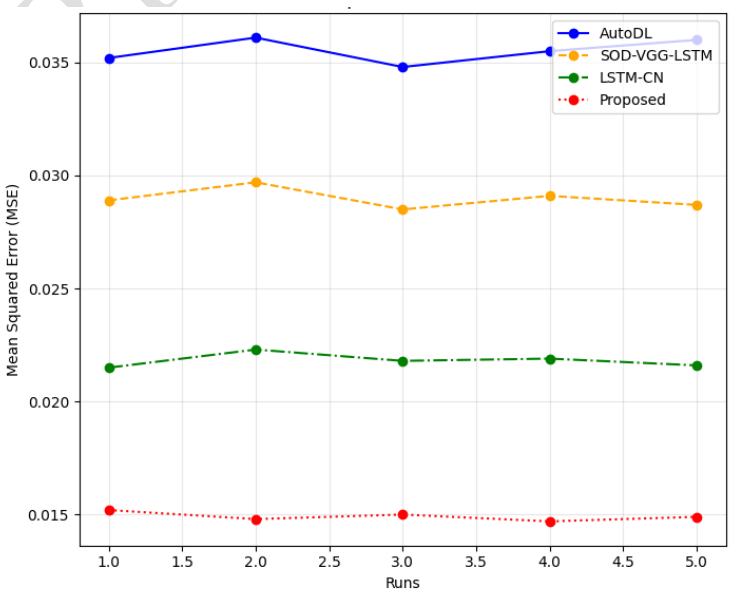


Figure 10. Comparison of MSE

The MSE comparison of REWAR-Sense models is illustrated in Figure 10. For the assigned WQPs, the results of the REWAR-Sense models are assessed in terms of MSE. Differences in the model's performance can be seen by comparing the dataset results. The Proposed REWAR-Sense approach achieves an MSE improvement of approximately 56.7%, 48.7%, and 31.8% compared to the existing approaches AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024), respectively, on average across all runs.

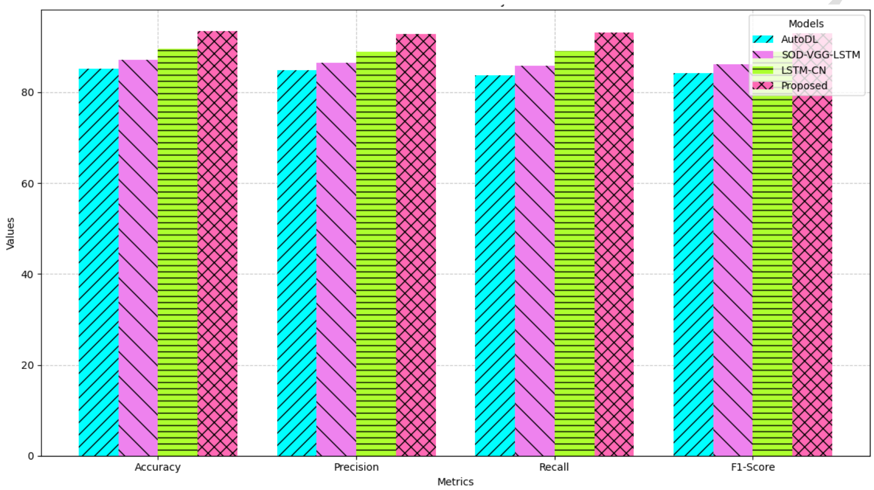


Figure 11. Performance Comparison of REWAR-Sense Method

Figure 11 compares the suggested REWAR-Sense technique's accuracy, precision, recall, specificity, and F1 score to those of current approaches. In the REWAR-Sense methodology, the WQ patterns of the Gulshan Lake are extracted by using the Ghost network, the quality of the water is predicted by using an AFSO algorithm, and the qualities of the waters are classified by the ABiRNN network. The usage of these novel techniques in the REWAR-Sense methodology provides a robust performance for WQ prediction and classification in Gulshan Lake. The REWAR-Sense method performs better than the current AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024) methods according to the metrics including accuracy, precision, recall, and F1-score. Specifically, it achieves 93.45% accuracy**,** 92.80% precision**,** 93.20% recall, and 93.00% F1-score respectively.

Figure 12 shows the typical RMSE accuracy comparison. It displays the overall accuracy of each prediction point. There is minimal variation in the error values between the various methods. The

Proposed REWAR-Sense approach achieves an RMSE improvement of approximately 34.8%, 27.6%, and 16.4% compared to the existing approaches AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024), respectively, on average across all runs.

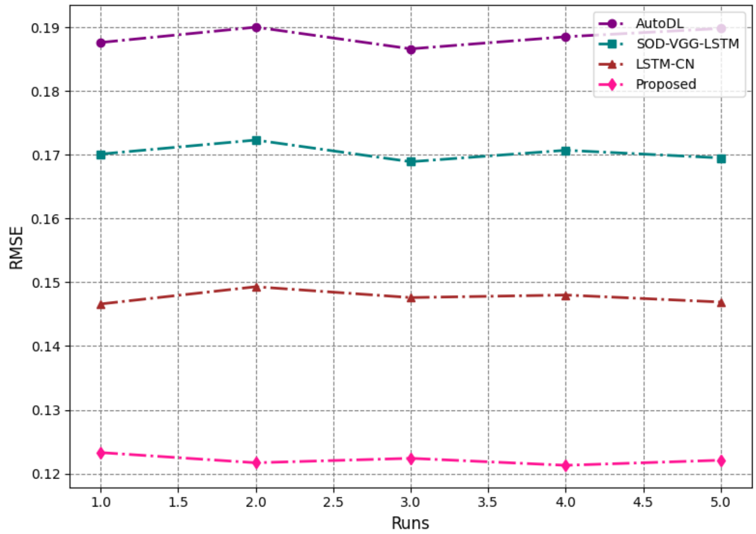


Figure 12. Comparison of RMSE

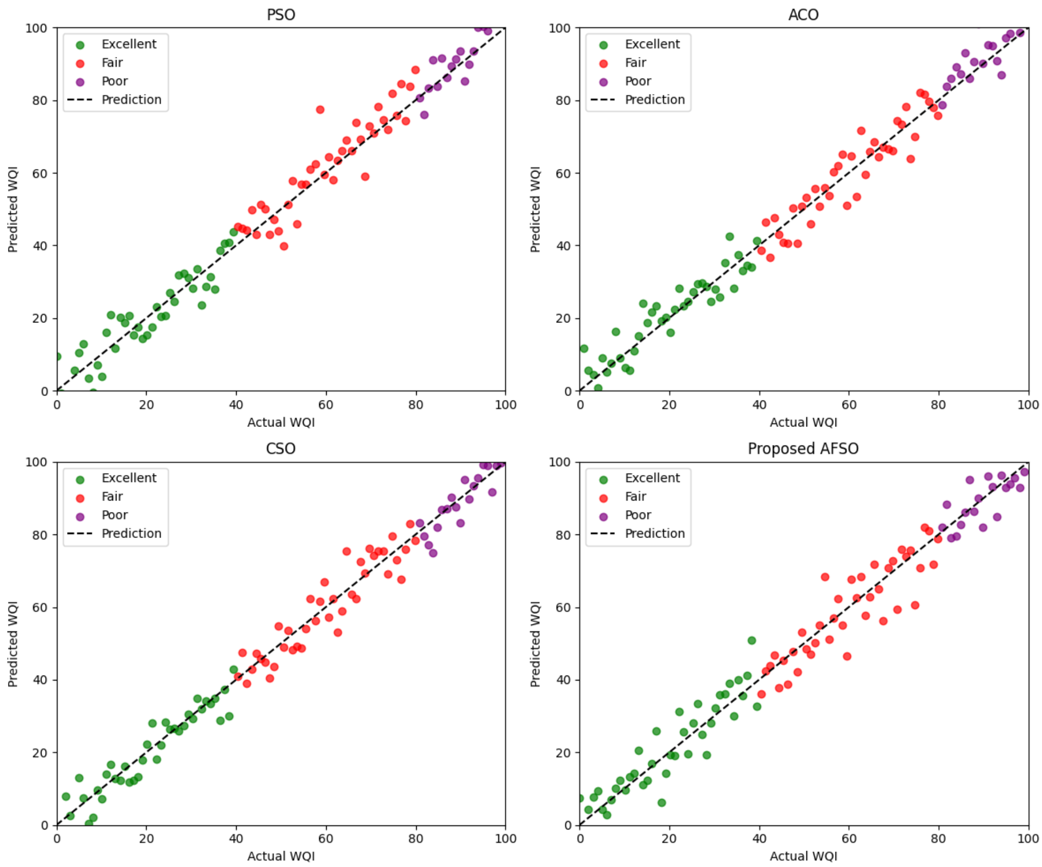


Figure 13. Water Quality Prediction

Figure 13 displays each class's WQI range. Scatter plots that forecast WQ classes, such as excellent, fair, and poor, by taking into account the relevant WQI values are shown in Figure 12. The Proposed AFSO approach achieves a WQP improvement of approximately 46.15%, 35.38%, and 28.81% compared to the existing PSO, ACO, and CSO approaches, respectively. The AFSO algorithm provides an accurate prediction score to identify the quality of the water based on its dynamic WQI values. According to its swarm behavior, the AFSO algorithm dynamically adjusts its swarm search behavior and provides an enhancement in WQ prediction for Gulshan Lake.

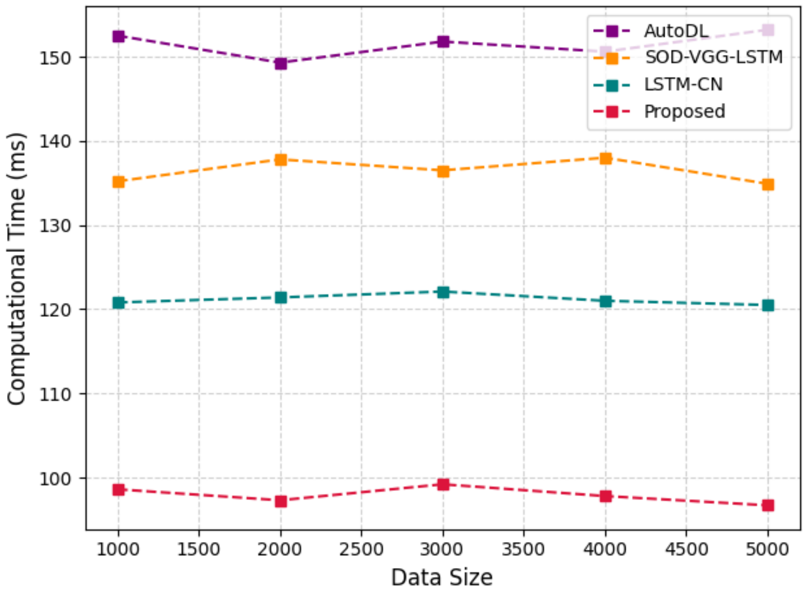


Figure 14. Comparison of Computational Time

Even though the graph topology influences the solutions shown in Figure 14, the run time always stays within the constraints of a stable real-time solution and grows linearly as the number of nodes in the network rises. The Proposed REWAR-Sense approach achieves a computation time improvement of approximately 34.7%, 27.5%, and 19.4% compared to the existing approaches AutoDL (Prasad et al. 2022), SOD-VGG-LSTM (Wan et al. 2022), and LSTM-CN (Mahesh et al. 2024), respectively on average across all data sizes.

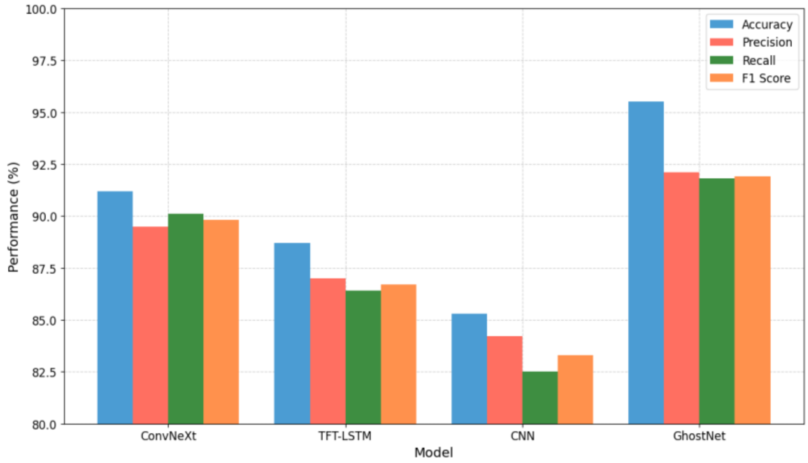


Figure 15. Comparison of Feature Extraction

Figure 15 compares the feature extraction efficiency of the Ghost Network with accuracy, precision, recall, specificity, and F1-score to those of current approaches. Compared to the existing ConvNeXt, TFT-LSTM, and CNN networks, the GhostNet extracts efficient features with less computations. It captures the required WQ patterns from the data without redundant complexity. According to its lightweight framework, the GhostNet framework captures the required features which is highly suitable for WQ monitoring. The GhostNet method performs better than the current ConvNeXt, TFT-LSTM, and CNN methods according to the metrics including accuracy, precision, recall, and F1-score. Specifically, it achieves 95.5% accuracy, 92.1% precision**,** 91.8% recall, and 91.9% F1-score respectively.

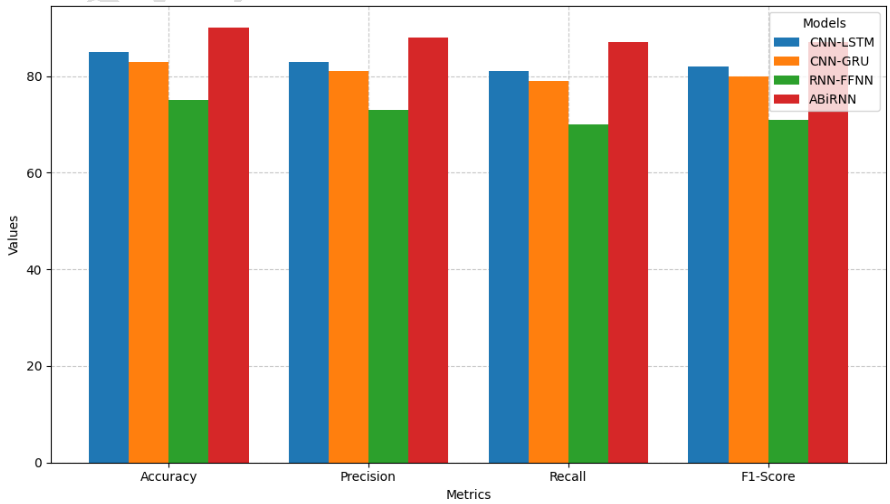


Figure 16. Performance Comparison with DL Techniques

A comparison between the proposed ABiRNN network and DL techniques such as CNN-LSTM, CNN-GRU, and RNN-FFNN is shown in Figure 16. The proposed Attention-based BiRNN approach outperforms the existing CNN-LSTM, CNN-GRU, and RNN-FFNN techniques in terms of F1-score, recall, accuracy, and precision. Specifically, it achieves 90% accuracy**,** 88% precision**,** 87% recall**,** and 87% F1-score respectively. In the REWAR-Sense methodology, the ABiRNN network categorizes the Gulshan Lake's WQ by capturing long-term dependencies in data with its bidirectional recurrent layers. The attention layer in the ABiRNN obtains the relevant patterns from the data and classifies the water according to its respective qualities which enhances the WQ of the Gulshan Lake.

Discussion

In this research, a novel REmote WAteR Sensing for quality assessment (REWAR-Sense) methodology has been proposed to develop an automated system for the prediction and classification of WQ in Gulshan Lake. During real-world applicability, the REWAR-Sense framework provides continuous monitoring of Gulshan Lake through IoT sensors and gathers the data for further processing. Due to this continuous data collection process the sensors may transmit massive amounts of redundant data which may arise computational overhead. These challenges were further addressed by using the given preprocessing and feature extraction techniques. The REWAR-Sense framework processes the gathered data and transforms them into a standardized format through preprocessing techniques. Furthermore, the GhostNet extracts the most relevant WQ patterns from the data with less computations which is highly suitable for WQ monitoring. Also, the REWAR-Sense framework provides an accurate WQI score prediction through the dynamic swarm behavior of AFSOA that ensures accurate prediction on dynamic water qualities of Gulshan Lake. Based on these WQI scores, the ABiRNN network captures the long-term dependencies from the data with its bidirectional recurrent layers and categorizes the Gulshan Lake's water qualities into potable water, palatable water, contaminated water, and infected water. Therefore, while deploying the REWAR-Sense system in different environments, larger water bodies, or different types of pollutants the proposed framework is highly scalable for WQ analysis. However, while deploying the REWAR-Sense system in larger water bodies or different types of pollutants there is a slight variation in its accuracies based on the water conditions and its polluted levels of the Gulshan Lake. However, the proposed REWAR-Sense system achieves superior results for accurate WQ prediction and classification of Gulshan Lake.

5. Conclusion

In this paper, a novel REWAR-Sense methodology is proposed to develop an automated system for the prediction and classification of WQ in Gulshan Lake. The REWAR-Sense methodology was simulated by using MATLAB and it is validated by Gulshan Lake Dataset. A comparison is made between the proposed REWAR-Sense methodology and existing methods such as AutoDL, SOD-VGG-LSTM, and LSTM-CN, according to the metrics including accuracy, precision, recall, specificity, F1-score, WQP Time, MSE, RMSE, and Computational Time. In comparison, the proposed REWAR-Sense methodology achieves a WQP time improvement of approximately 43.7%, 34.8%, and 20.4% compared to the existing approaches AutoDL, SOD-VGG-LSTM, and LSTM-CN respectively. Conversely, the proposed REWAR-Sense method achieves an accuracy of 93.45%**,** precision of 92.80%**,** recall of 93.20%**,** and F1-score of 93.00% outperforming the existing AutoDL, SOD-VGG-LSTM, and LSTM-CN methods respectively. The GhostNet method performs better than the current ConvNeXt, TFT-LSTM, and CNN methods according to the metrics including accuracy, precision, recall, and F1-score. Specifically, it achieves 95.5% accuracy, 92.1% precision, 91.8% recall, and 91.9% F1-score respectively. The REWAR-Sense methodology is currently validated only on Gulshan Lake which may limit its generalizability to other water bodies with different environmental conditions. In the future, the REWAR-Sense methodology will be further enhanced by incorporating real-time alert mechanisms for WQ anomalies and expanding the model to include additional water bodies for broader applicability and generalization.

References

Azrour M, Mabrouki J, Fattah G, Guezzaz A. and Aziz, F. (2022), Machine learning algorithms for efficient water quality prediction, *Modeling Earth Systems and Environment,* **8**(2), 2793-2801.

Chen Z., Xu H., Jiang P., Yu S., Lin G., Bychkov I., Hmelnov A., Ruzhnikov G., Zhu N. and Liu Z. (2021), A transfer Learning-Based LSTM strategy for imputing Large-Scale consecutive missing data and its application in a water quality prediction system, *Journal of Hydrology,* **602**, 126573.

Chhipi-Shrestha G., Mian H.R., Mohammadiun S., Rodriguez M., Hewage K. and Sadiq R. (2023), Digital water: artificial intelligence and soft computing applications for drinking water quality assessment, *Clean Technologies and Environmental Policy,* **25**(5), 1409-1438.

Deng T., Chau K.W. and Duan H.F. (2021), Machine learning based marine water quality prediction for coastal hydro-environment management, *Journal of Environmental Management,* **284**, 112051.

Díaz-González L., Aguilar-Rodríguez R.A., Sansalvador J.P. and Lakouari N. (2025), AQuA-P: A machine learning-based tool for water quality assessment, *Journal of Contaminant Hydrology*, 104498.

Geethamani R., Soundara B. and Kanmani S. (2023), Assessment of Groundwater Quality in and around the Solid Waste Dump Site using Multivariate Plot Analysis, *Global NEST Journal,* **25**(5).

Iniyan Arasu, M., Subha Rani, S., Thiyagarajan, K. and Ahilan, A. (2024), AQUASENSE: aquaculture water quality monitoring framework using autonomous sensors, *Aquaculture International,* 32(7), 9119-9135.

Islam N. and Irshad K. (2022), Artificial ecosystem optimization with deep learning enabled water quality prediction and classification model. *Chemosphere,* **309**, 136615.

Khan M.S.I., Islam N., Uddin J., Islam S. and Nasir M.K. (2022), Water quality prediction and classification based on principal component regression and gradient boosting classifier approach, *Journal of King Saud University-Computer and Information Sciences,* **34**(8), 4773-4781.

Khullar S. and Singh N. (2022), Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation, *Environmental Science and Pollution Research,* **29**(9), 12875-12889.

Krishna Bikram Shah, Visalakshi S. and Ranjit Panigrahi. (2023), Seven class solid waste management-hybrid features based deep neural network, *International Journal of System Design and Computing,* 01(01), 1-10.

Kushwaha N.L., Kudnar N.S., Vishwakarma D.K., Subeesh A., Jatav M.S., Gaddikeri V., Ahmed, A.A. and Abdelaty I. (2024), Stacked hybridization to enhance the performance of artificial neural networks (ANN) for prediction of water quality index in the Bagh river basin, India, *Heliyon,* **10**(10).

Li Z., Liu H., Zhang C. and Fu G. (2024), Real-time water quality prediction in water distribution networks using graph neural networks with sparse monitoring data, *Water Research,* **250**, 121018.

Li, Y., Cong, R., Zhang, K., Ma, S. and Fu, C. (2024), Four-way game analysis of transformation and upgrading of manufacturing enterprises relying on industrial internet platform under developers' participation, *Journal of Asian Architecture and Building Engineering,* 1-22.

Mahesh N., Babu J.J., Nithya K. and Arunmozhi S.A. (2024), Water quality prediction using LSTM with combined normalizer for efficient water management, *Desalination and Water Treatment,* **317**, 100183.

Maruthai, S. et al. (2025), Wastewater Recycling Integration with IoT Sensor Vision for Real-time Monitoring and Transforming Polluted Ponds into Clean Ponds using HG-RNN, *Global NEST Journal,* 27(4).

Nithya, A., Appathurai, A., Venkatadri, N., Ramji, D.R. and Palagan, C.A. (2020), Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images, *Measurement,* 149, 106952.

Pang J., Luo W., Yao Z., Chen J., Dong C. and Lin K. (2024), Water quality prediction in urban waterways based on wavelet packet Denoising and LSTM, *Water Resources Management,* **38**(7), 2399-2420.

Prasad D.V.V., Venkataramana L.Y., Kumar P.S., Prasannamedha G., Harshana S., Srividya S.J., Harrinei K. and Indraganti S. (2022), Analysis and prediction of water quality using deep learning and auto deep learning techniques, *Science of the Total Environment,* **821**, 153311.

Rahu M.A., Chandio A.F., Aurangzeb K., Karim S., Alhussein M. and Anwar M.S. (2023), Towards design of Internet of Things and machine learning-enabled frameworks for analysis and prediction of water quality, *IEEE Access.*

Rajanbabu J., Venkatakrishnan G.R., Rengaraj R., Rajalakshmi M. and Jayaprakash N. (2025), An integrated smart water management system for efficient water conservation, *International Journal of Electrical & Computer Engineering* (2088-8708), **15**(1).

Rajeswari, D., Rajendran, S., Arivarasi, A., Govindasamy, A. and Ahilan, A. (2024), TOSS: Deep Learning based Track Object Detection using Smart Sensor, *IEEE Sensors Journal.*

Raveena, S. and Surendran, R. (2024), Real-Time Monitoring and Optimization of Recycled Water Irrigation in Coffee Plantations using Bi-Directional RNNs and IoT Sensors, In *2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT),* 116-123. IEEE.

Santhiya Govindapillai and Radhakrishnan Ayyapazham. (2024), Segmentation and Classification of Paddy Leaf Disease via Deep Learning Networks, International Journal of Data Science and Artificial Intelligence, 02(05), 161-169.

Satish N., Anmala J., Varma M.R. and Rajitha K. (2024), Performance of Machine Learning, Artificial Neural Network (ANN), and stacked ensemble models in predicting Water Quality

Index (WQI) from surface water quality parameters, climatic and land use data, *Process Safety and Environmental Protection,* **192**, 177-195.

Shams M.Y., Elshewey A.M., El-Kenawy E.S.M., Ibrahim A., Talaat F.M. and Tarek Z. (2024), Water quality prediction using machine learning models based on grid search method, *Multimedia Tools and Applications,* **83**(12), 35307-35334.

Surendran, R. and Krishnan, S. (2024), Comparative of SVM and Decision Tree Techniques for Predicting Hydroponic Tomato Growth and Yield Using Deep Water Culture, In *2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* 705-710. IEEE.

Talukdar S., Ahmed S., Naikoo M.W., Rahman A., Mallik S., Ningthoujam S., Bera S. and Ramana G.V. (2023), Predicting lake water quality index with sensitivity-uncertainty analysis using deep learning algorithms, *Journal of Cleaner Production,* **406**, 136885.

Venkata V. et al. (2024), Water Quality Prediction using Statistical, Ensemble and Hybrid models, *Global NEST Journal,* **26**(2).

Venkatraman, M., Surendran, R., Srinivasulu, S. and Vijayakumar, K. (2023), Water quality prediction and classification using attention based deep differential recurflownet with logistic giant armadillo optimization, *Global NEST Journal,* 27(1).

Wan H., Xu R., Zhang M., Cai Y., Li J. and Shen X. (2022), A novel model for water quality prediction caused by non-point sources pollution based on deep learning and feature extraction methods, *Journal of Hydrology,* **612**, 128081.

Wang, Z. and Ma, S. (2024), Research on the impact of digital inclusive finance development on carbon emissions—Based on the double fixed effects model, *Global NEST Journal*, *26*(7), 06227.

Wang, Z., Wu, Q. and Ma, S. (2024), Research on carbon emission peaks in large energy production region in China—based on the open stirpat model, *Global NEST Journal*, *26*(5).

Wu J. and Wang Z. (2022), A hybrid model for water quality prediction based on an artificial neural network, wavelet transform, and long short-term memory, *Water,* **14**(4), 610.

Wu, Y., Hou, L., Yuan, Y., Ma, S. and Zeng, H. (2025), Green credit policy's influence on construction firm ESG performance: a difference in differences estimation, *Journal of Asian Architecture and Building Engineering*, 1-13.

Zeng, H., Abedin, M.Z., Lucey, B. and Ma, S. (2025), Tail risk contagion and multiscale spillovers in the green finance index and large US technology stocks, *International Review of Financial Analysis*, 97, 103865.

Zhang Q. and You X.Y. (2024), Recent advances in surface water quality prediction using artificial intelligence models, *Water Resources Management,* **38**(1), 235-250.

Zhang, G., Ma, S., Zheng, M., Li, C., Chang, F. and Zhang, F. (2025), Impact of Digitization and Artificial Intelligence on Carbon Emissions Considering Variable Interaction and Heterogeneity: An Interpretable Deep Learning Modeling Framework, *Sustainable Cities and Society,* 106333.