KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 18, NO. 1, Jan. 2024

IoT-Based Automatic Water Quality Monitoring System with Optimized Neural Network

Anusha Bamini A M[1], Chitra R[2], Saurabh Agarwal[3], Hyunsung Kim[4*], Punitha Stephan[5], and Thompson Stephan[6]
[1, 2]Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, India
[3]Department of Computer Science and Engineering, Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida, India
[4]School of Computer Science, Kyungil University, Gyeongsan, South Korea
[5,6]Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun, India
[email : anushabamini@gmail.com[1], chitrajegan5@gmail.com[2], saurabhnsit2510@gmail.com[3], kim@kiu.ac.kr[4], punitharesearch@gmail.com[5], thompsoncse@gmail.com[6]]
*Corresponding Authors: Saurabh Agarwal, Hyunsung Kim

Received September 10, 2022; revised May 2, 2023; revised July 25, 2023; revised September 11, 2023; accepted October 10, 2023; published January 31, 2024

Abstract

One of the biggest dangers in the globe is water contamination. Water is a necessity for human survival. In most cities, the digging of borewells is restricted. In some cities, the borewell is allowed for only drinking water. Hence, the scarcity of drinking water is a vital issue for industries and villas. Most of the water sources in and around the cities are also polluted, and it will cause significant health issues. Real-time quality observation is necessary to guarantee a secure supply of drinking water. We offer a model of a low-cost system of monitoring real-time water quality using IoT to address this issue. The potential for supporting the real world has expanded with the introduction of IoT and other sensors. Multiple sensors make up the suggested system, which is utilized to identify the physical and chemical features of the water. Various sensors can measure the parameters such as temperature, pH, and turbidity. The core controller can process the values measured by sensors. An Arduino model is implemented in the core controller. The sensor data is forwarded to the cloud database using a WI-FI setup. The observed data will be transferred and stored in a cloud-based database for further processing. It wasn't easy to analyze the water quality every time. Hence, an Optimized Neural Network-based automation system identifies water quality from remote locations. The performance of the feed-forward neural network classifier is further enhanced with a hybrid GA- PSO algorithm. The optimized neural network outperforms water quality prediction applications and yields 91% accuracy. The accuracy of the developed model is increased by 20% because of optimizing network parameters compared to the traditional feed-forward neural network. Significant improvement in precision and recall is also evidenced in the proposed work.

Keywords: Arduino model, IoT, neural network, water quality.

http://doi.org/10.3837/tiis.2024.01.004 **ISSN : 1976-7277**

1. Introduction

Water is one of the vital renewable sources for humans and other livelihoods worldwide. Good quality water is scarce because of the high population and pollution. The drinking water of millions of people is polluted or chemically contaminated. Half of the groundwater is used for household activities; a quarter is used for agriculture; one-third is used in industries. Even though the groundwater is not directly polluted, industrialization and climatic change improve the harm to watersheds and aquifers. The presence of chemicals like fluoride and lead in natural groundwater also causes disease in a human. The excess fluoride in water causes bone, thyroid, neurological, and skin diseases. The contaminated water affects green globalization as well as the ecosystem. The water quality is affected by many factors like runoff, sedimentation, temperature, dissolved oxygen, pH value, turbidity, humidity, and so on [1]. Environmental surroundings influence some factors, whereas humans are responsible for most aspects. Suppose the quality of the water is identified at an early stage. In that case, the appropriate measures can be taken to avoid a dangerous situation. The quality of the water depends on the physical, chemical, and biological components diffused in it. Generally, the water quality is checked in a laboratory from the collected water sample. It is a costlier and time-consuming process that involves some human resources. Sensor networks play a significant role in gathering data for environmental applications.

IoT extends its support to predict real-time environmental factors in addition to Industry 4.0. It is important- to create a low–cost and easy-to-use automatic water quality prediction system for providing good quality water to millions of populations. IoT-enabled smart systems are enough to interpret real-world data and are reliable for environmental management; thus, they reduce the risk of human health problems. The abundant growth of VLSI design and Wireless sensor networks possesses many microcontrollers for real-time applications. Arduino is a microcontroller motherboard that is suitable for simple applications. Sensors measured the parameters of pH, turbidity, temperature, and TDS, and the deviations in these water parameters highlighted the quality of the water. Various sensors collect the data transmitted to the internet through the Arduino gateway. In the proposed method, the sensors connected with Arduino measure the water quality, and the values are loaded into the database for a given time interval. The main feature of the developed model is the Arduino is automatically operated from a remote location, and the water quality is predicted on the server side. For adopting the early-level prediction, the optimized ANN is proposed. The system is initially trained with the existing dataset in which nine input parameters predict drinking water quality. Initially, the developed system is trained and tested with various classifiers, and it is noticed that ANN is suitable for the best prediction. But only 70 percent accuracy is archived by this system. Hence, the ANN parameter's hidden layer, momentum factor, and learning rate are optimized with a hybrid GA-PSO algorithm to optimize the system's performance. The entire workflow of our model is divided into two sub-systems. Initially, a real-time Arduino-based water quality prediction system is implemented, and the performance is checked for different kinds of water. In the second stage of the work, the data is stored in the cloud-based server. Thus, the quality is further analyzed by an optimized artificial neural network. This technology fixes the Arduino chip in various water sources, and the data is collected and stored in a particular interval. The optimized ANN periodically analyzes the quality to ensure the quality of the water. This technique's water sources are interconnected, and drinking water quality is guaranteed. Hence, it is unnecessary to go to the place and collect the water for analysis; instead, the water quality is predicted in the remote server itself. The developed water systems continuously monitor water quality. Then, they can test the groundwater or the water from any water body.

2. Literature Review

An IoT-based smart water quality system was proposed by a few researchers [2, 3]. Monitoring water contamination in daily life is becoming more and more crucial nowadays. The developments in sensors, communication, and Internet of Things (IoT) technology play a vital role in water quality prediction. This paper analyzes and discusses the comprehensive assessment of the most recent intelligent water pollution monitoring systems. An IoT-based smart water quality monitoring system with cheap cost and high efficiency that continuously checks quality parameters is provided in this work. The proposed work verifies the built-in model using three water samples. After that, findings are forwarded to a cloud server for additional processing. Three water models are analyzed, and outputs are used to determine whether the water is safe to drink.

A low-cost, highly effective device called Water Quality Monitoring keeps track of the purity of drinking water [4]. The system proposed in that work comprises numerous sensors that measure various parameters, including pH, water turbidity, tank level, temperature, and ambient humidity. This sensor can be connected to a microcontroller unit, and a personal computer was used for additional processing. The gathered data is forwarded to the cloud via the IoT-based ThinkSpeak application for tracking water quality [5]. Several sensors were successfully interfaced with the designed system, which included Arduino Mega and NodeMCU target boards. The pH varies from 6.5 to 7.5 for Hyderabad city supply water and 7 to 8.5 for groundwater. The measured turbidity ranged from 600 to 2000 NTU for both the groundwater supplies for Hyderabad Metropolitan City and surface water. The web server monitored variables such as pH, water turbidity, tank level, temperature, and atmospheric humidity through a web-based application called ThingSpeak. These measured parameters are also tracked via the ThingSpeak smartphone app. In addition, this work must be done to analyze various other factors in the water, such as electrical conductivity, free residual chlorine, nitrates, and dissolved oxygen.

A smart water quality monitoring system for continuously monitoring four different factors, temperature, pH, electric conductivity, and turbidity, was described by Singh et al. [6]. Four sensors were discreetly connected to an Arduino Uno to monitor water properties. The sensor data was received by an application based on the NET platform, and the values obtained were compared to WHO reference values. The proposed SWQM system accurately determined whether the tested water sample could be drunk using a fast forest binary classifier based on the observed data.

The work [7] describes a smart water quality monitoring system that uses the Internet of Things and can continuously monitor four different factors: temperature, pH, electric conductivity, and turbidity. Four sensors are discreetly connected to an Arduino Uno to monitor water properties. The sensor data was collected and compared to WHO reference values in that desktop-based application on the NET platform. Based on the observed data, using a fast forest binary classifier, the proposed SWQM system accurately determined whether the tested water sample was drinkable. The water quality was continuously assessed using Internet of Things devices like NodeMCU. An integrated Wi-Fi module has been implemented in NodeMCU, enabling internet connectivity to forward measured data from sensors to the cloud. Some prototype was developed to observe various contaminants present in the water. Numerous sensors were used to monitor different characteristics to analyze the water body quality. The deep learning algorithms were used to predict the quality of the water, and the findings were recorded in the cloud.

The work proposed in [8, 9] is applied to quantify water's physical and chemical properties. The measured properties of water are temperature, pH, TDS, and turbidity. The core controller processed the sensor data. The Raspberry Pi model is used as a core controller. At last, the ThingSpeak API, the sensor data was transferred to the internet. We aim to provide a water monitoring system with high frequency, mobility, and less power consumption; this work is unique. With the support of a Raspberry Pi and numerous sensors, we could detect TDS, temperature, turbidity, and pH values in water in our suggested system. Conductivity, hardness, chloride, and other parameters will be monitored in the future.

The eyes and skin of pool users may react severely to a high pH level in the water [10,11]. Since all procedures must be carried out manually, the method used to measure the pH level of water at Seberang Jaya Public Swimming Pool is ineffective for maintaining water quality. The depicted Smart Water Quality Monitoring System was evaluated by various impact factors such as pH value and swimming pool temperature using DOE and ANOVA statistical methods. Results of the research show that the time of day, pool vol., and their interface aspects didn't affect the pH value and pool's water temperature. If someone tried to tamper with the system, proximity sensors would have transmitted a message to the officials through the GSM module.

This article described [12] an innovative technique for improving water quality monitoring to avoid polluted water. The well-created method is used to gather quality metrics, particularly chemical markers, and develop typical maps for all parameters of water bodies in an investigation. Real-time data was uploaded to the cloud to keep track of the water body's quality and obtain real-time quality data for several chemical and biological indicators, including pH, dissolved oxygen, total dissolved solvents, turbidity, and so on. pH and temperature are two of the parameters evaluated by the prototype. The suggested approach addressed fundamental flaws in previous systems, increasing water quality monitoring efficiency and reducing pollution.

The development and improvement of real-time low-cost water quality monitoring systems in IoT were mentioned in [13, 14]. The system comprises multiple sensors that identify water's chemical and physical characteristics. The water's temperature, pH, turbidity, and flow sensor qualities can all be measured. The core controller can process the sensors' measured values. One option is to use the Arduino model as the main controller. Ultimately, data collected through the sensor was displayed online via Wi-Fi.

The monitoring water quality system was created to meet the needs of the Department of Environmental Protection in a specific area of water quality criteria. The system comprises a Wireless Water Quality Monitoring Network and a Remote Data Center. The brain of the entire hardware was the wireless microprocessor CC2430. The Zigbee wireless transmission technology was used to build the sensor network. WSN GPRS DTU, having a built-in TCP/IP protocol, was used to get the sample water quality and transmit a result online. The Remote Data Center receives real-time data via the Internet, which it analyses, processes, and records. The Department of Environmental Protection will afford real-time recommendations to businesses. The most significant of this work can be more proficiency and reduced cost. A paper [15] described the role of an optimized neural network for disease prediction from the standard UCI database and the real-time data set. It was proved that the classification accuracy of the ANN was improved with GA and PSO-optimized neural networks.

3. ARDUINO UNO-Based Water Quality Prediction System

The present water quality prediction method with Arduino is proposed in this section. Various sensors identified to characterize drinking water's physical and chemical characteristics are also described here. The data obtained from multiple sensors are processed by the microcontroller and displayed by LED.

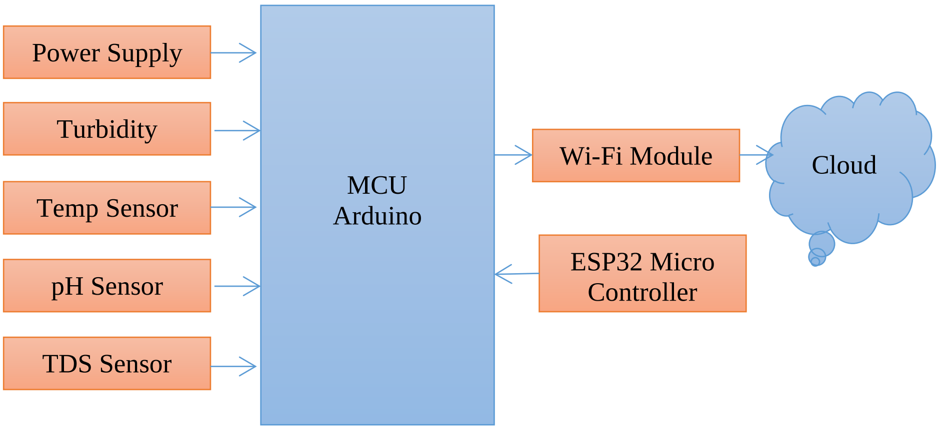

Fig. 1. Proposed System Architecture

The values obtained by various sensors are transmitted into the cloud database through Wi-Fi, and Arduino will act as a gateway for this transmission. The flow diagram of the first stage of water quality prediction is represented in **Fig. 1**.

3.1 SENSOR USED

3.1.1 Turbidity sensor

Turbidity is a measure of pathogenic microorganisms in water, as shown in **Fig. 2**. It will not directly affect the health of humans. Still, it causes illness through the microorganisms present in it. This sensor's output voltage is dependent on a range of temperatures. The cloudiness of water is gauged by its turbidity. The level of water's translucency loss has been expressed by turbidity. It is regarded as a reliable indicator of water quality. The light that submerged undersea flora needs is obscured by turbidity. Due to greased areas suspended close to the face that absorb heat from the sun, it can also cause facial water temperatures to rise over usual.

Fig. 2. Turbidity sensor

3.1.2 Temperature Sensor

The safe limit for drinking water is eight degrees Celsius, more than the human body can withstand. Therefore, The water should be between 28 and 44 degrees Celsius. It could be used to gauge a water's temperature. **Fig. 3** represents the temperature sensor.

Fig. 3. Temperature Sensor

3.1.3 pH Sensor

A water contaminant's pH, which runs from one to fourteen, with seven being the middle point, determines how acidic or basic it is. As the number below seven decreases, the degree of acidity increases, with one being the most acidic. Values above seven are considered alkaline, and the amount of alkalinity increases as the number rises, with 14 being the most alkaline. When water is sufficiently acidic or basic, the H+ or OHion activity in stream water may harm aquatic life. The hydrogen ion attention of the result is implicitly related to the measuring electrode's sensitivity to hydrogen ions. A temperature detector is also required to adjust the voltage change because the discriminational voltage of electrodes changes with the temperature—pH sensor shown in **Fig. 4**.

Fig. 4. pH Sensor

3.1.4 TDS Sensor

TDS (Total Dissolved Solids) are organic & inorganic compounds that are liquefied in water. The lower the level, the better for drinking. TDS levels in drinking water, on the other hand, can range from 300 to 500 mg/liter. If the water is in this range, it is suggested to drink the water Sensor shown in **Fig. 5**.

Fig. 5. TDS sensor

3.1.5 ESP8266 WI-FI

Any microcontroller may access the Wi-Fi network with an ESP8266 Wi-Fi Module in **Fig. 6**, a tone-consist of SOC integrated with TCP/IP protocol stack. ESP8266 has the capacity to host operations or decompress all Wi-Fi networking tasks from a different operation processor. The pre-programmed AT command is set into each ESP8266 module. ESP8266 segments are a lower price panel in small size and have expanding capability.

Fig. 6. ESP8266 WI-FI

To work with the water quality monitoring system, we need technologies such as pH sensor, turbidity sensor, conductivity sensor (TDS sensor), temperature sensor Arduino Uno, ESP8266 (Wi-Fi), and Arduino IDEv (software). The proposed architecture is coupled to the core controller and consists of numerous sensors (temperature, pH, turbidity, TDS). The core controller accesses and processes the sensor values to transmit data via the internet. One of the core controllers is Arduino. On the internet Wi-Fi system, the sensor data can be accessed. After connecting all the sensors with the Arduino and the ESP8266, we must code to get the desired output in the website, such as the blynk; think to speak, we can print the output through PC or Android. We can even print in the form of a graph. The observed values gathered from tested types of water can be added to the cloud database for further processing.

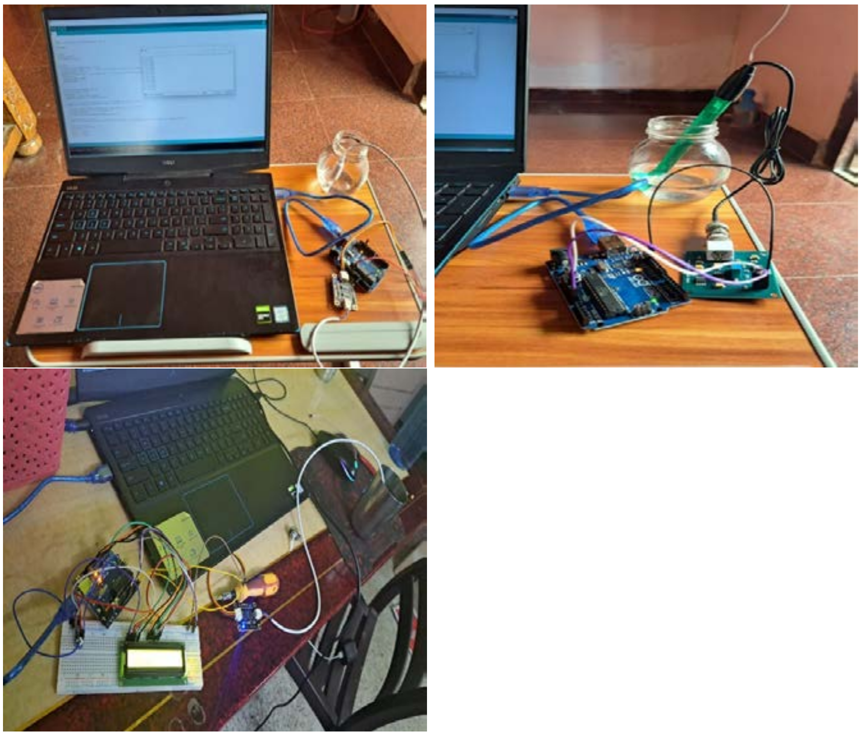

Fig. 7. Working Models

Based on the output values received from the Arduino kit in **Fig. 7**, the graph is plotted for water pH, turbidity, temperature, and TDS, as shown in **Fig. 8**.

Fig. 8. Plots for sensor output

4. Water Quality Prediction Using Optimized Neural Network

The second phase of the proposed work implements soft prediction algorithms for remote monitoring of drinking water. The real-time Arduino-based systems are connected to the server through a cloud environment by the internet. The framework of the connectivity is shown in **Fig. 9**.

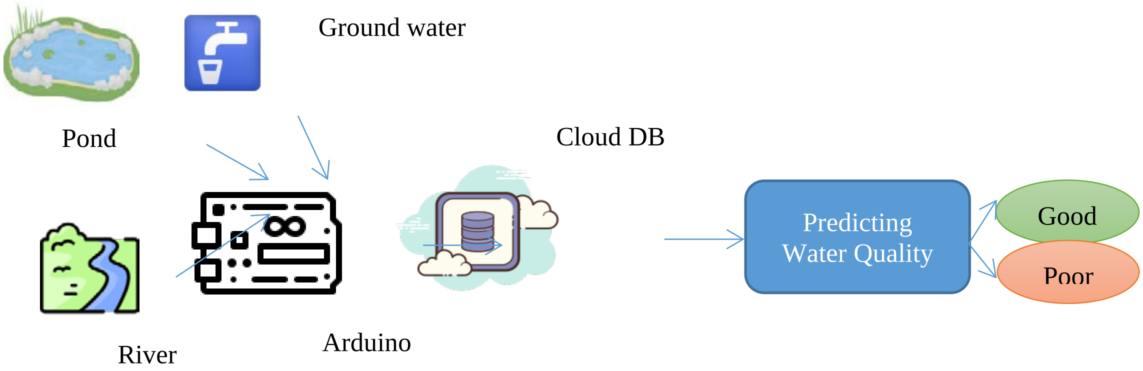

Fig. 9. Framework for water quality

The water quality classifier is designed, and the efficiency is analyzed. The water quality dataset is taken from Kaggle. A total of 3,276 water samples were taken. From these samples, ten variables are taken for making the dataset. Potability is considered a response variable or output parameter from these ten. The remaining nine variables are explanatory variables or input parameters. The parameters listed in the Kaggle dataset are pH, Hardness, Solids, Chloramines, Sulfate, Conductivity, Organic_carbon, Trihalomethanes, Turbidity, and Potability.

4.1 Exploratory data analysis

The exploratory data analysis is conducted for the selected dataset [16] to find the correlation between the variables and the water quality. The heat map of various input parameters is shown in **Fig. 10**. The heatmap is used to check the correlation between the input variables. It is evident that there exists a positive correlation between hardness, sulfate, conductivity, carbon level, and turbidity.

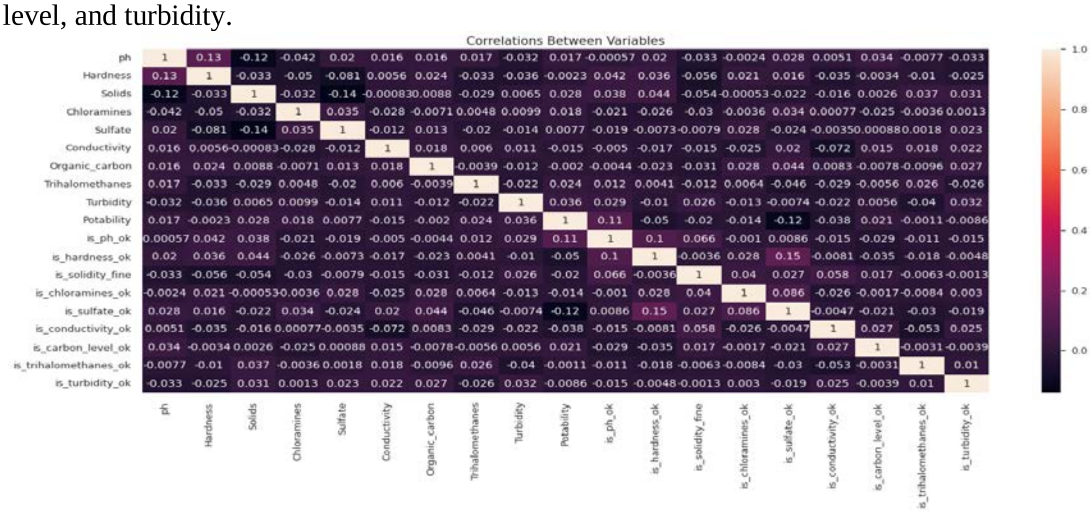

Fig. 10. Heat Map of Input Variables

The range of input variables for water quality is indicated in **Fig. 11**. Potability 1 indicates that the water is safe and suitable for drinking.

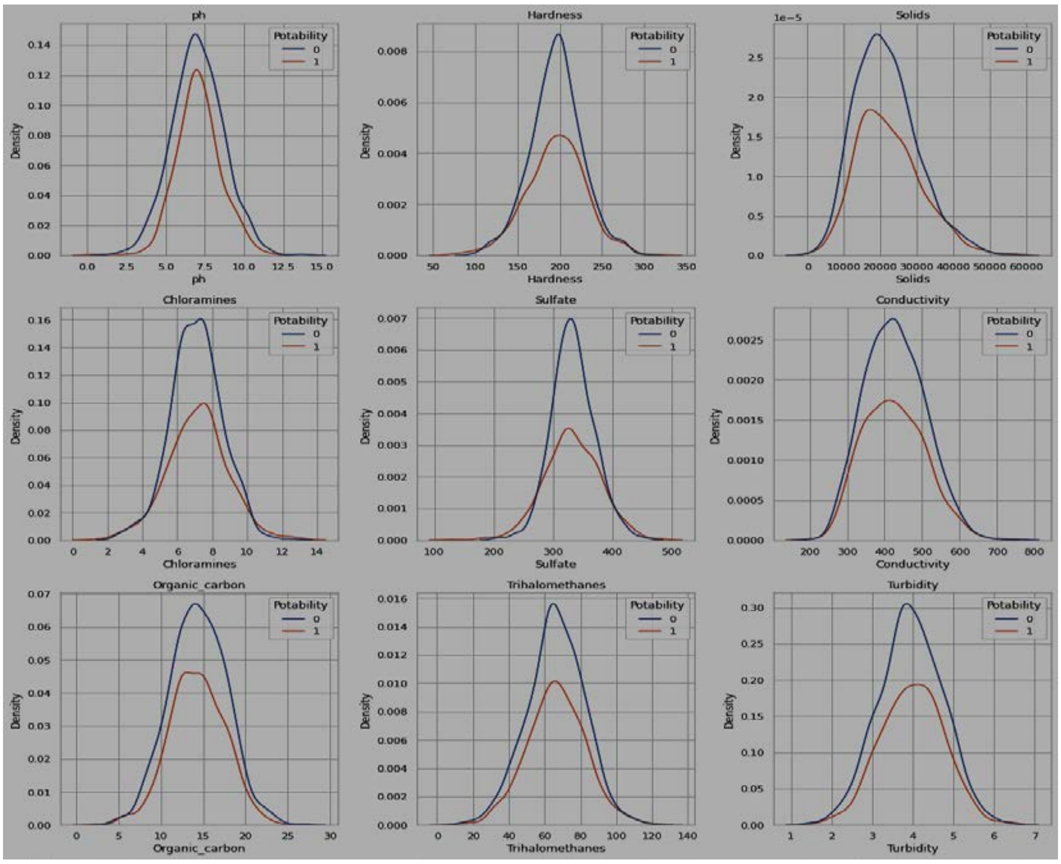

Fig. 11. Range of input variables

4.2 Water Quality Prediction Using ANN

Different classifiers are used for the classification of data. Among the ANN classifiers, the accuracy is high. The dataset values are normalized with min-max normalization before being given as the input layer of the ANN to increase the system's accuracy. In the min-max normalization technique, the value in the dataset is converted to the new value calculated by the given formulas (1) and (2).

$$v_i' = \frac{v_i - min_A}{max_A - min_A}\,(new\_max_A - new\_min_A) + new\_min_A \tag{1}$$

This concept converts all the data in the selected dataset between [0, 1] irrespective of its original range.

$$v_i' = \frac{v_i - min_A}{max_A - min_A} \tag{2}$$

The proposed ANN to predict the water quality is shown in **Fig. 12**.

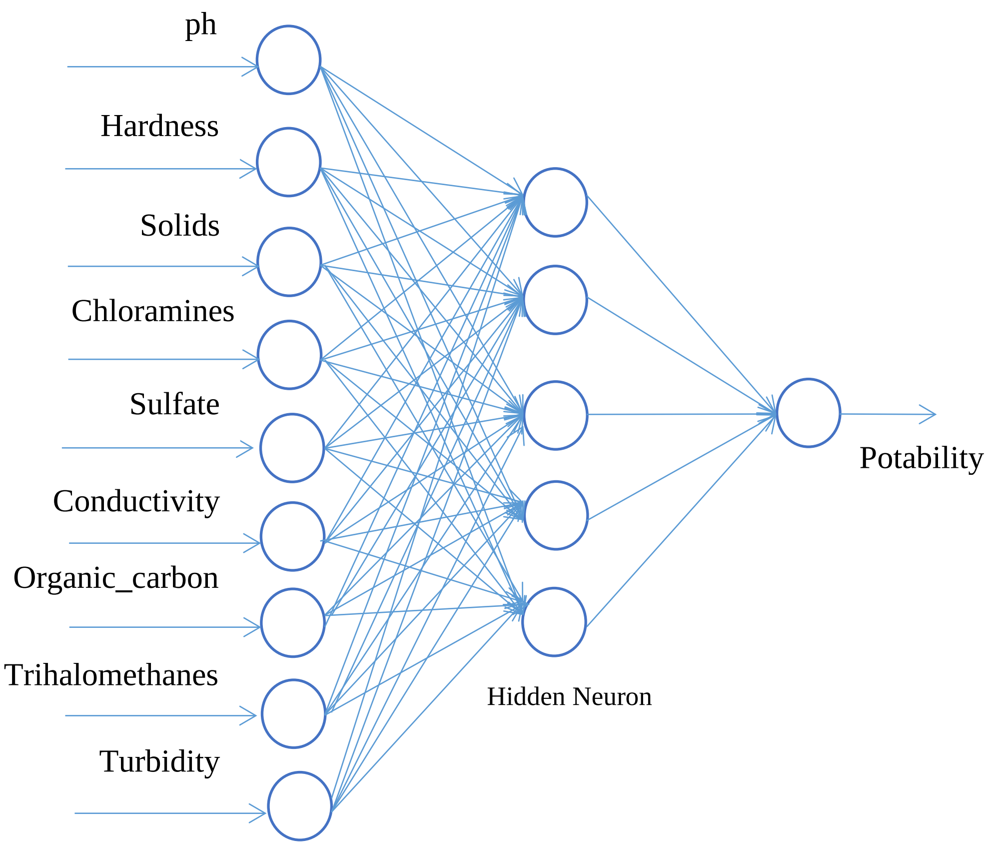

Fig. 12. ANN for predicting water quality

The backpropagation training algorithm is used for training the water quality prediction ANN. The system's accuracy is analyzed with the help of mean square error (MSE). E denotes the MSE, which is calculated by Eqn. 3,

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2 \tag{3}$$

The weights are adjusted using gradient descent to minimize the error function by Eqn. 4.

$$\Delta w_{ij}(t+1) = -\eta\frac{\partial E}{\partial w_{ij}} + \alpha\Delta w_{ij}(t) \tag{4}$$

where E is the error signal, α is the momentum factor, and η is the learning rate.

The input parameters are defined as$X = (x_1, x_2, x_3, \ldots \ldots, x_9)$.

Where 1,2,……………,9 represent the dataset attributes that predict water quality from various water resources. There is only one predominant output variable that predicts the quality of water. The pseudo-code of the backpropagation algorithm used in the water quality prediction ANN is as follows.

Algorithm 1: Backpropagation algorithm for water quality prediction

for $i = 1,2 \ldots\ldots n$ ***do***
begin
fix *the hidden neuron range for ANN*
initialize *learning rate = 0.01, momentum constant = 0.9* ***end***
initialize *the network parameter weights with small random values*
initialize *the* $epoch_{max}=1000$
do
train *the input and compute the network output for each training attribute*
Find *the Mean Square Error*

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2$$

if *MSE is less than or equal to acceptable level*
end *the network training*
else if
$epoch \geq epoch_{max}$
end *the network training*
else
change the weight

$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t)$$

epoch =epoch+1
end
end

The performance of the ANN classifier for predicting water quality is mentioned in **Fig. 13 and Fig. 14**.

Fig. 13. Accuracy

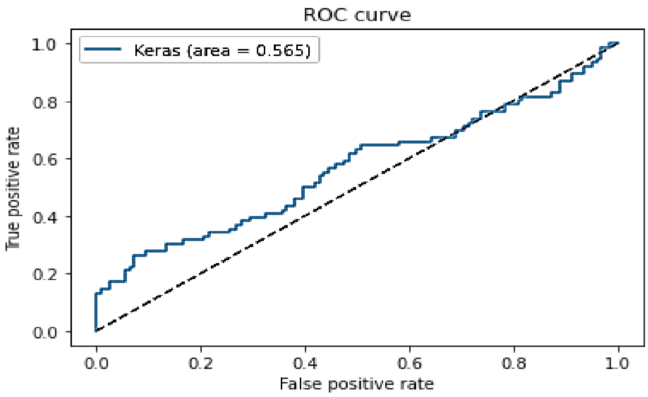

Fig. 14. ROC curve

The accuracy of the ANN classifier for water quality prediction is 70%. In the proposed work, the quality of water quality prediction ANN is improved by an optimization algorithm. The ANN parameters are optimized to increase the accuracy of the automatic water quality prediction system.

5. Water Quality Prediction Using Hybrid GA-PSO Optimized ANN

The Hybrid GA-PSO algorithm is developed to optimize the ANN parameters of the prediction system. In the hybrid approach, the network parameters are optimized by GA, and the PSO algorithm again enhances the GA chromosomes. A flow chart for the hybrid GA-PSO algorithm is pictured in **Fig. 15**.

The upper and lower range of process parameters, such as a hidden neuron, learning rate, and momentum factor, are defined before performing optimization.

$R_{min} = \{Nh_{min}, Lr_{min}, Mc_{min}\}$ and $R_{max} = \{Nh_{max}, Lr_{max}, Mc_{max}\}$

The activation function f is defined in terms of the weighted sum of inputs and the bias value, which is denoted by y_k^p in Eqn. 5.

$$y_k^p = f(s_k^p) \tag{5}$$

Where $s_k^p = \sum_j w_{j,k}\, y_j^p + b_k$, y_j^p . In the selected ANN, the activation functions used in the hidden neuron and output layer are sigmoid and linear, respectively, as denoted in Eqn. 6.

$$f(x) = \frac{1}{1+e^{-x}} \tag{6}$$

For the hybrid GA-PSO optimization process, the MSE of the selected ANN is defined in Eqn. 7,

$$MSE_{GAPSO} = \sum_{p \epsilon T} \sum_{k=1}^{N_O} (t_k^p - y_k^{p,o})^2 \tag{7}$$

Here t_k^p is the expected output, $y_k^{p,o}$ is the actual output from the k^{th} neuron in an output layer for the selected input p in a data set. MSE is taken as the fitness function, and the GA-PSO algorithm yields the optimal solution for the formulated function.

Every particle in the swarm signifies a set of weights for a current epoch or iteration in a PSO-optimised neural network implementation. The quantity of weights is linked to each particle's dimension in the network.

To reduce MSE, the particle moves within the weight space. Changing the position means altering the network's weight to lower the inaccuracy of the present epoch. Every epoch, every particle calculates the new velocity they will utilize to go to the new spot, updating their position accordingly.

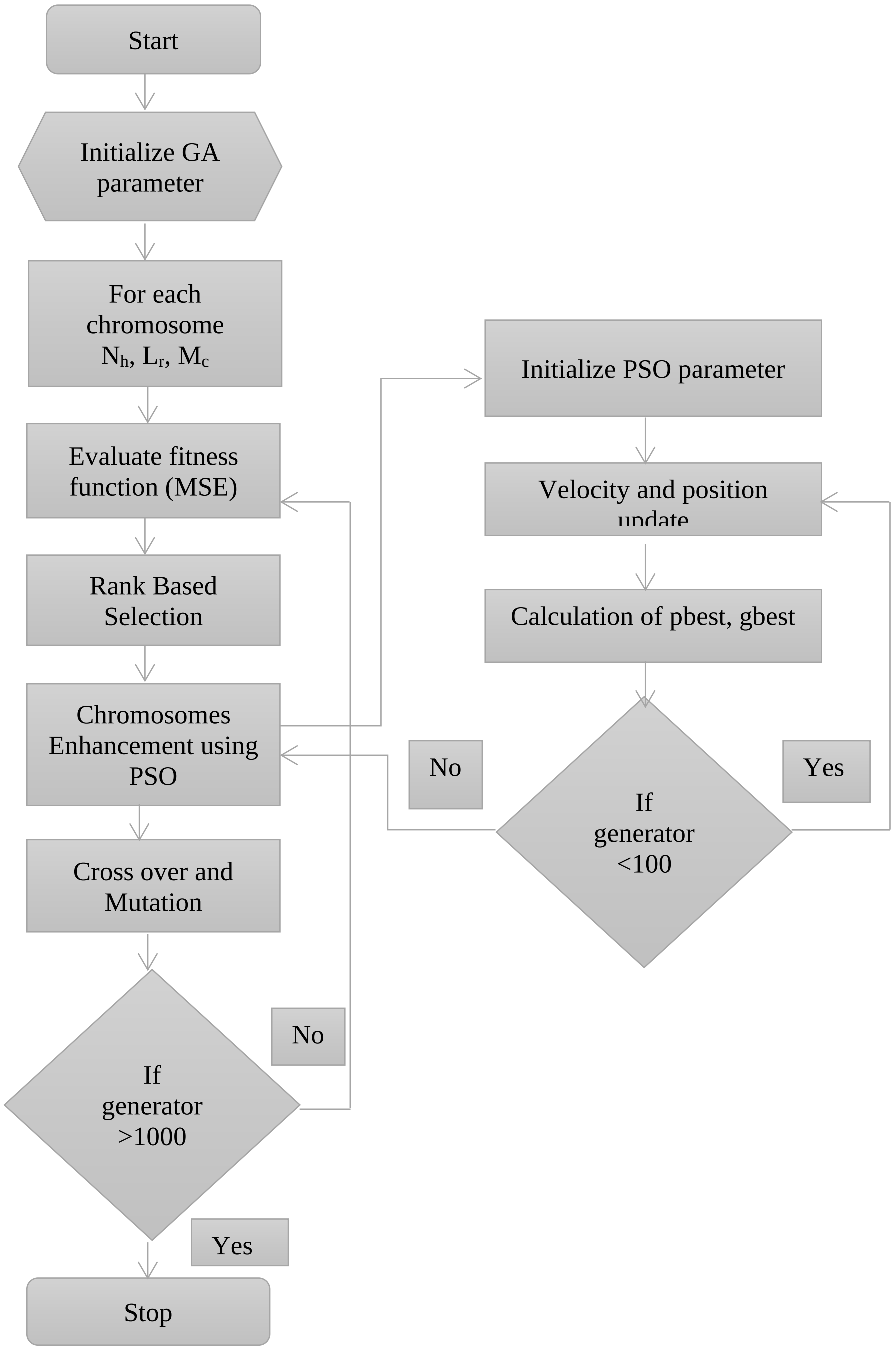

Fig. 15. Flowchart for hybrid GAPSO optimized ANN

A set of new weights called the new position is used to calculate the unknown error. Even though no improvement has been seen, the revised weights are being used for PSO. Each particle is subjected to this process once more. The particles with the highest accuracy are regarded as the best globally thus far. The training continues until the best particle achieves a satisfactory error or until the iteration limit is reached.

When the training is over, the weights are utilized to calculate the classification error for the training patterns. The network is then tested using the test patterns using the same set of weights. PSONN does not use the backpropagation idea, but the feed-forward NN generates a learning error (particle fitness) based on a set of weights and biases (PSO positions). The velocity update equation's best and gbest values produce a value for position adjustment to the best solution or targeted learning error (lowest learning error found throughout the entire learning process thus far). New sets of positions are produced by multiplying the determined velocity value by the current position value and using the movement equation. Then, the ANN uses that new set of coordinates to generate new learning errors (particle fitness). The cycle is repeated until the stop criteria are fulfilled (maximum number of iterations). An acceleration constants c1, c2, the PSO parameters, and r1 and r2 are random integers with initial values set to [0, 1] initially set to 2.

Algorithm 2: The PSO algorithms used to enhance the GA parameters

Each *particle P*
Initialize *particle parameters from the GA algorithm*
Stop
Until *reaches termination Criterion*
All *Particles*
Find *the fitness parameter MSE*

$$MSE_{PSONN} = \sum_{p\epsilon T} \sum_{k=1}^{N_O} (t_k^p - y_k^{p,o})^2$$

New MSE less than the overall Pbest value
Assign new value as (Pbest)
Select the particle *with less MSE among all the particles as Gbest*
Find *the velocity and position of each particle*
Compute *particle velocity*

$$vel_i(t+1) = wvel_i(t) + c_1 rand_1(Pbest_i - pos(t)) + c_2 rand_2(Gbest_i - pos(t))$$

Compute *particle position*

$$pos(t+1) = pos(t) + vel_i(t+1)$$

End

Where

$vel_i(t)$	:	in i[th] iteration velocity of particle t
$pos(t)$	:	position of the particle $in\ the\ iteration\ t$
$Pbest_i$	:	local best of particle
$Gbest_i$	:	Global best of the particle
$rand$	:	a random number between 0 and 1
w	:	weight fn
c_1	:	cognition learning rate
c_2	:	social learning rate

The particle with the G*best* at the optimal dimension achieves the best fitness score (minimum MSE), recorded throughout each PSO iteration. Optimized ANN parameters are produced using the suggested GA-PSO ANN method with N h = 19, L r = 0,0112, and M c = 0.925. Therefore, as long as optimality holds true, the GA-PSO ANN algorithm produces a compact network architecture compared to the complicated network. This newly configured ANN with optimized parameters predicts drinking water quality. It is proved that the optimized

ANN classifier outperforms compared to other classification algorithms. Hence, this technique can be adapted to predict drinking water quality remotely.

Table 1 represents the various classifier results for water quality compared with accuracy, precision, and recall. Compared with six classifiers, logistic regression gives better accuracy.

Table 1. Results obtained for various classifier

Classifier	Accuracy	Precision	Recall
Random Forest	0.68	0.67	0.69
KNN	0.60	0.56	0.58
Logistic Regression	0.50	0.49	0.52
SVM	0.65	0.64	0.67
ANN	0.70	0.68	0.72
GA-PSO optimized ANN	0.91	0.89	0.94

6. Conclusion

An efficient, low-cost IoT model is proposed in this work to predict water quality. Using Arduino, the temperature, pH, turbidity, and TDS sensors are connected. A different type of water is tested in the first phase of the work. The sensors, as mentioned earlier, are utilized to get the output. The second phase of the work continued with the quality test using a different water dataset, which is stored in cloud DB. A total of 3,276 water samples are present in the dataset. From these samples, ten variables are taken for making the dataset: pH, Hardness, Solids, Chloramines, Sulfate, Conductivity, Organic_carbon, Trihalomethanes, Turbidity, and Potability. Using the ANN classifier, the potability parameter is predicted. The performance of the ANN classifier is further enhanced with a hybrid GA- PSO algorithm. Checking the quality of drinking water in the manual is a difficult one. Instead, a remote-operated Arduino gateway connects water resources in a particular region to the cloud database. Thus, the proposed methodology continuously monitors the drinking water quality within a time interval. Further, this technology can be enhanced for remote prediction of seawater quality and analysis of the installation of water distillation centers in coastal areas with IoT technology.

References

[1] He, Dong, Zhang, Li-Xin, "The water quality monitoring system based on WSN," in *Proc. of 2012 2nd International Conference on Consumer Electronics, Communications, and Networks (CECNet)*, pp. 3661-3664, 2012.

[2] Bhattarai, A, Dhakal, S, Gautam, Y, Bhattarai, R, "Prediction of Nitrate and Phosphorus Concentrations Using Machine Learning Algorithms in Watersheds with Different Landuse," *Water*, 13, 3096, 2021. Article (CrossRef Link).

[3] Varsha, Lakshmi Kantha, Anjitha Hiriyannagowda Akshay Manjunath Aruna PattedJagadeesh, Basavaiah, "IoT based smart water quality monitoring system," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 181-186, November 2021. Article (CrossRef Link).

[4] SathishPasika, Sai TejaGandla, "Smart water quality monitoring system with cost-effective using IoT," *Heliyon*, vol. 6, no. 7, p. e04096, July 2020. Article (CrossRef Link).

[5] Ajith Jerom B.; R. Manimegalai; R. Manimegalai, "An IoT Based Smart Water Quality Monitoring System using Cloud," in *Proc. of 2020 International Conference*, 2020.

[6] Singh R, Baz M, Gehlot A, Rashid M, Khurana M, Akram SV, Alshamrani SS, Alghamdi AS, "Water Quality Monitoring and Management of Building Water Tank Using Industrial Internet of Things," *Sustainability*, 13(15), 8452, 2021. Article (CrossRef Link).
[7] Monira Mukta, Samia Islam, Surajit Das Barman, Ahmed Wasif Reza, M Saddam Hossain Khan, "IoT based Smart Water Quality Monitoring System," in *Proc. of 2019 IEEE 4th International Conference*, 2019. Article (CrossRef Link).
[8] Anuradha T, Bhakti, Chaitra R, Pooja D, "IoT Based Low Cost System for Monitoring of Water Quality in Real Time, International Research Journal of Engineering and Technology," e-ISSN: 2395-0056, vol. 05, no. 05, May-2018.
[9] Zessner M, "Monitoring, Modeling and Management of Water Quality," *Water*, 13(11), 1523, 2021. Article (CrossRef Link).
[10] Shabinar Abdul Hamid, Ahmad Mustaqim Abdu Rahim, Solahuddin Yusuf Fadhlullah, Samihah Abdullah, "IoT based Water Quality Monitoring System and Evaluation," in *Proc. of 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 2020. Article (CrossRef Link).
[11] Brinda Das, P.C. Jain, "Real-Time Water Quality Monitoring System using Internet of Things," in *Proc. of International Conference on Computer, Communications and Electronics Manipal University Jaipur, Malaviya National Institute of Technology Jaipur & IRISWORLD*, July 01-02, 2017.
[12] Shriram, K. Vasudevan, Balraj Baskaran, "An improved real-time water quality monitoring embedded system with IoT on unmanned surface vehicle," *Ecological Informatics*, Vol. 65, 101421, November 2021. Article (CrossRef Link).
[13] Vaishnavi V. Daigavane and Dr. M.A Gaikwad, "Water Quality Monitoring System Based on IOT, Advances in Wireless and Mobile Communications," ISSN 0973-6972, Vol. 10, no. 5, pp. 1107-1116, 2017.
[14] Jan F, Min-Allah N, Düştegör D, "IoT Based Smart Water Quality Monitoring : Recent Techniques, Trends, and Challenges for Domestic Applications," *Water*, 13(13), 1729, 2021. Article (CrossRef Link).
[15] V. Seenivasagam, R. Chitra, "Myocardial Infarction Detection Using Intelligent Algorithms," *Neural Network World*, Vol. 26, No. pp. 91-110 (1.77), 2016.
[16] Besseris G, "Micro-Clustering and Rank-Learning Profiling of a Small Water-Quality Multi-Index Dataset to Improve a Recycling Process," *water*, 13(18), 2469, 2021. Article (CrossRef Link).

Anusha Bamini A M received her PhD in Computer Science and Engineering in 2018 from Noorul Islam Centre for Higher Education, Tamilnadu, India. She received a Master of Engineering (M.E) from Anna University Tirunelveli, India 2010, with a University First Rank. She received a Bachelor of Engineering (B.E) with distinction from Anna University, Chennai, India, in 2008. Now, she is an Assistant Professor at Karunya Institute of Technology and Sciences. Her area of interest is Cloud Computing, Scheduling and Security in Cloud Computing, IoT, Image Processing, Deep Learning, and Data Science. She has published papers in conferences, SCIE, SCOPUS, and Peer-reviewed Journals.

Chitra R is currently working as an Associate Professor of Computer Science and Engineering at Karunya Institute of Technology and Sciences, Coimbatore. She has 18 years of experience in Computer Science and Engineering. She received a B.E. degree in Electrical and Electronics Engineering. M.E. degree in Computer Science and Engineering and Ph.D. in Computer Science and Engineering from Manonmaniam Sundaranar University. She has published 14 papers in Data mining, Artificial Intelligence, and Soft Computing. Her research areas include Data Mining, Intelligent Techniques, and Bioinformatics.

Saurabh Agarwal is an associate professor at Amity School of Engineering & Technology, Amity University, Noida, India. He received his Ph.D. in Computer Engineering from the University of Delhi in 2017, India. His current research interests are image forensics, computer vision, and IoT.

Hyunsung Kim received M.Sc. and Ph.D. in computer engineering from Kyungpook National University, Korea, in 1998 and 2002. He has been a Professor in the Department of Cyber Security at Kyungil University, Korea since 2012. Furthermore, he has been a visiting professor at the Department of Mathematical Sciences, Chancellor College, University of Malawi, Malawi, since 2015. He also was a visiting researcher at Dublin City University in 2009. From 2000 to 2002, he worked as a senior researcher at Ditto Technology. He was an associate professor from 2002 to 2012 with the Department of Computer Engineering at Kyungil University. His research interests include cryptography, VLSI, authentication technologies, network security, ubiquitous computing security, and security protocol.

Punitha Stephan received her Ph.D. in computer science and engineering from Pondicherry University, India. Currently, she is working as an Associate Professor in the Department of Computer Science and Engineering at Graphic Era Deemed to be University, India. Her areas of interest are medical image analysis, evolutionary computation, artificial intelligence, and machine learning techniques. She has published her research works in reputed journals, conferences, patents, and book chapters.

Thompson Stephan earned his Ph.D. in Computer Science and Engineering from Pondicherry University, India, in 2018. He currently serves as an Associate Professor in the Department of Computer Science & Engineering at Graphic Era Deemed to be University, Dehradun, Uttarakhand, India. In 2023, he was recognized as one of the world's top 2% most influential scientists, a distinction jointly announced by Elsevier and Stanford University, USA. He achieved academic excellence during his master's degree, earning a university rank. Moreover, he has received prestigious awards, including the Best Researcher Award in 2020 and the Protsahan Research Award in 2023, both presented by the IEEE Bangalore Section, India. His primary research interests revolve around the implementation and application of artificial intelligence techniques in practical scenarios. Thompson has authored numerous technical research papers published in esteemed journals and conferences by organizations like IEEE, Elsevier, Springer, and others. He actively contributes as a reviewer for respected international journals and serves as an engaged book editor. Currently, he is a member of the IEEE.