AI-Driven Smart Hydroponic Monitoring System for Water Quality and Disease Prediction

*Priyanka Kumari[1], Arjun Singh Rawat[2], and Gunjan Gunjan[3]
Department of Computer Science and Engineering
National Institute of Technology Delhi, India
[1*]242210018@nitdelhi.ac.in, [2]arjunsinghrawat@nitdelhi.ac.in, [3]gunjan.rehani@nitdelhi.ac.in

Abstract. Hydroponic farming is increasingly being adopted as a sustainable agricultural technique due to its efficient use of water and ability to support controlled environment agriculture. However, maintaining optimal nutrient solution conditions and detecting plant diseases remain critical challenges that can significantly affect crop productivity. This study proposes an AI-driven smart hydroponic monitoring system that integrates machine learning and deep learning techniques for water quality assessment and plant disease prediction.

Water quality classification is performed using machine learning models trained on a combined dataset consisting of hydroponic and aquaponic environmental measurements, including pH, electrical conductivity (EC), and temperature. Several classification algorithms, including Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting, were evaluated to determine the most effective model for identifying safe and unsafe nutrient solution conditions. Experimental results show that the Gradient Boosting classifier achieved the best performance with an accuracy of 95.58% and a ROC-AUC score of 0.9909.

In addition, a deep learning model based on the MobileNetV2 architecture was employed for lettuce disease detection using a publicly available leaf image dataset. The trained model achieved a validation accuracy of 98.40% in distinguishing between healthy and unhealthy plant leaves. The proposed models were integrated into a web-based dashboard that enables users to monitor hydroponic conditions, perform image-based disease detection, and obtain prediction results with confidence scores. The results demonstrate that combining machine learning-based environmental monitoring with deep learning-based plant health assessment can provide an effective intelligent decision-support system for hydroponic farming. The proposed framework contributes toward improving crop management and enabling early detection of unfavorable growing conditions in controlled agricultural environments.

Keywords: Hydroponic farming, smart agriculture, machine learning, deep learning, water quality monitoring, plant disease detection, MobileNetV2, precision agriculture

A. Agnihotri et al. (eds.), *Proceedings of the Conference on Bridging Engineering Disciplines with AI and Machine Learning (BEDAIML 2026)*, Advances in Intelligent Systems Research 209,
https://doi.org/10.2991/978-94-6239-697-5_9

1 Introduction

Hydroponics is a soil-less agricultural technique in which plants are cultivated using nutrient enriched water rather than conventional soil. The idea of soil-less plant growth traces in history back to the several centuries but, it is in the twentieth century, that contemporary hydroponic systems received significant attention as a response to the rising land degradation, urbanization and water scarcity [1]. The hydroponic farm system can precisely determine how much nutrient a plant can consumed from the environment and more control over plant environment which results in a faster growth rate and increased productivity than the other conventional agriculture [2].

The most popular hydroponic system is Ebb & Flow where roots of plants are periodically immersed in the nutrient-rich solution and drained into reservoir. The cyclic flooding and draining cycle improve nutrient uptake yet maintains suitable root zone aeration. The hydroponic system consists of a nutrient reservoir, distribution tubes, a tray with growing plants and media, and timer activated with the pump as depicted in Fig. 1. Through upwards flow plants receive dissolved nutrients and the downward flow ensures there is no water standing and it also provides aeration. Such structured nutrient circulation improves growth efficiency and provides a stable environment for controlled-environment agriculture.

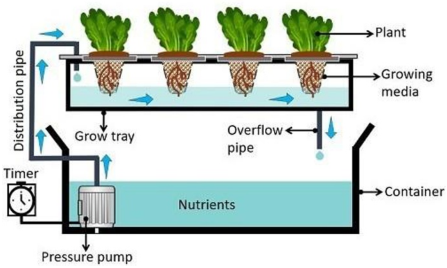

Fig.1: Schematic representation of an Ebb and Flow hydroponic system illustrating periodic flooding and draining of nutrient solution between the reservoir and grow tray [3].

In hydroponic cultivation, maintaining optimal water quality parameters is critical for plant health and system stability. Key parameters include pH, electrical conductivity (EC), and temperature. The pH level determines nutrient solubility and availability, while EC indicates the concentration of dissolved salts within the nutrient solution. Temperature influences root metabolism and dissolved oxygen levels. Deviations from recommended ranges can lead to nutrient deficiencies, toxicity, and reduced yield production [4]. Therefore, intelligent monitoring and automated assessment of these parameters are essential in modern hydroponic systems.

Although nutrient solution monitoring is critical, disease detection in plants represents an ongoing challenge. Even though hydroponics minimizes soil-borne pathogens, crops continue to be susceptible to fungal diseases, bacterial diseases and nutritional disorders. Timely detection is critical to stop disease spread, and maintain full yields. Advances in Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) during recent years have dramatically transformed the agricultural monitoring scenario. Many machine learning algorithms, including the random forest algorithm, and gradient boosting performed well on structured environmental data [5][6]. Deep learning models, and particularly Convolutional Neural Network (CNN), have achieved high accuracies in plant disease classification [7][8].

Although the research has shown many advancements, existing work focuses primarily on either water quality monitoring or plant disease detection as separate topics. There is a need for integrated frameworks where sensing of environmental parameters and plant disease detection based on images can be integrated into a single smart hydroponic monitoring system. However, a lot of them focus on hardware-centric Internet of Things (IoT) implementations with minimal emphasis on comparing models to one another, the reproducibility of the work and overall system performance metrics.

In response to such restrictions, we recommend here an intelligent hydroponic smart monitoring system based on AI and machine learning with the capacity of water quality classification as well as deep learning for disease prediction in plants. The proposed system analyzes multiple machine-learning approaches for nutrient solution assessment, and uses a convolutional neural network for computer vision-based plant disease detection. Through integration of all components into an integrated web-based interface, the system is intended to provide an intelligent decision-support system to improve modern hydroponic farming.

The scope of the work is restricted to data-driven modeling; algorithm comparison and system level software integration. The developed framework advances the state-of-the-art to develop scalable AI-driven solutions for smart hydroponic agriculture.

2 Related Work

Smart agriculture is becoming increasingly integrated with machine learning and deep learning for precision agriculture and smart decision-making. Several investigations have been carried out in hydroponic farming on using supervised learning for the enhancement of controlled environmental parameters and for maximizing plant growth through the applications of supervised learning algorithms.

Rahman *et al.* [9] reported an AIoT-based hydroponic monitoring system for the environmental monitoring through which the temperature, humidity and pH parameters are also detected by the sensor network of connected nature. The system of them depends on cloud computation to support the automation of adjustment and optimize the operational efficiency. Likewise, Mehare and Gaikwad [10] reported in their paper that supervised classification models of Logistic Regression, Random Forest, and Support Vector Machine were explored for agricultural monitoring by smart farming with respect to classification for agricultural monitoring.

Gokul *et al.* [11] proposed an IoT-based hydroponic automated system which utilizes ML based real-time controller for automatic environment adjustments. Further, Singh *et al.* [12] developed an extended research work by incorporating mobile-based remote monitoring in order to improve end-user accessibility and true-farm management in real-time. This work demonstrates the increasing trend of intelligent systems in hydroponic settings.

At the same time, deep learning has greatly improved plant disease detection through image-based classification. Sladojevic *et al.* [7], showed the application on plant diseases with leaf images CNN. Mohanty *et al.* [8], extending the validation for the potential power of the CNN and training the CNN models on some large-scale plant image datasets have resulted in very good classification accuracy for multiple crop species. Since transfer learning architectures like the widely adopted MobileNetV2 are computationally lightweight and real-time friendly for agricultural applications have seen widespread adoption.

While progress has been made in research, such as water quality monitoring or plant disease detection, most researches are focused on these two topics separately. Limited work has proposed a complete solution to jointly tackle the environment condition categorization and image-based disease prediction in a unified hydroponic monitoring system.

The applications of machine learning in environmental monitoring in smart farming are very popular, while deep learning methods have achieved great success in the plant disease detection for image-based analysis. Thus, there is a need for a data-driven and integrated hydroponic monitoring framework that evaluates different machine learning models for water quality classification while integrating DL-based plant health assessment into an intelligent decision-making system. The proposed work aims to address the gap by developing and evaluating a unified smart hydroponic monitoring platform.

3 Methodology

The section describes the overall methodology adopted in the study for developing an intelligent hydroponic monitoring system. The proposed system integrates machine learning for water quality classification and deep learning for plant disease detection. The methodology consists of several stages including system architecture design, dataset preparation, data preprocessing, feature engineering, model training, and system integration.

3.1 System Architecture

A novel approach has been presented with combined water quality sensing and plant disease detection for hydroponic farmers in the form of a system, as illustrated in Fig. 2.

The main indicators for health of nutrient solution are pH, electrical conductivity (EC) and temperature Water quality parameters. These parameters are fed to machine learning models to determine if the environment is safe or not for plant growth of hydroponic. In addition to water monitoring, a deep learning model is applied in processing plant leaf image data for disease detection. We also integrate the output of the above-mentioned two model through online dashboard, with which user can monitor in real-time the condition of the plant and nutrients.

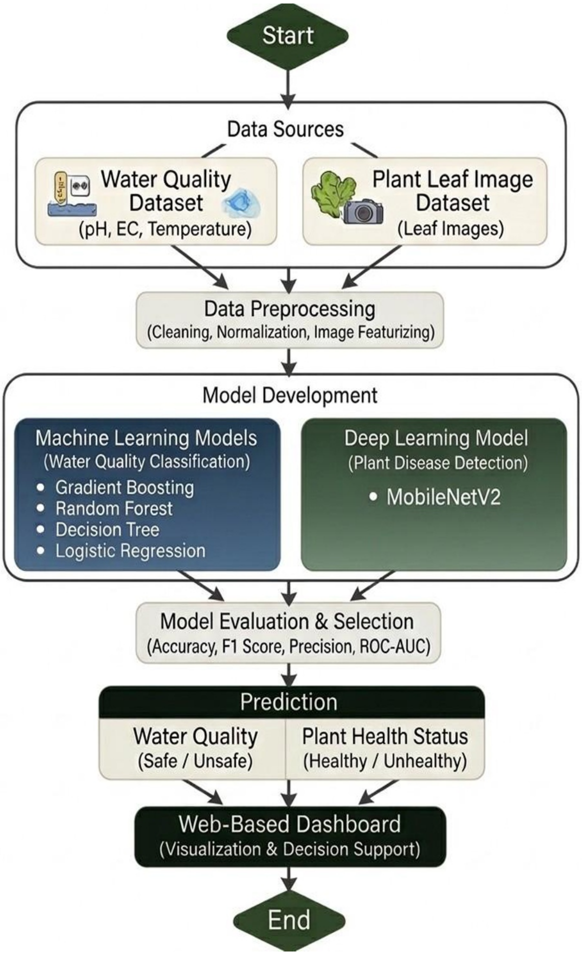

Fig.2: Architecture of the proposed smart hydroponic monitoring system integrating water quality classification and plant disease detection.

3.2 Dataset Description

Two types of datasets were used in the study: a hydroponic water quality dataset for machine learning classification and a plant disease image dataset for deep learning–based disease detection.

Hydroponic Water Quality Dataset Water quality data were obtained from two publicly available datasets related to hydroponic and aquaponic farming systems.

The first dataset, *HydroGrowNet of Batavia*, contains water quality measurements collected from hydroponic lettuce cultivation experiments, including pH, electrical conductivity, and temperature readings [13].

The second dataset contains water quality parameters collected from an IoT-based aquaponic lettuce farming system [14]. These datasets provide continuous environmental measurements that reflect real-world hydroponic operating conditions.

Both datasets were merged to create a comprehensive dataset for machine learning training. After cleaning and preprocessing, the final dataset contained 70,626 samples with three primary features: pH, electrical conductivity (EC), and temperature.

Table 1 presents sample records from the water quality dataset used for classification.

Table 1: Sample records from the hydroponic water quality dataset

pH	EC ($\mu S/cm$)	Temperature ($^\circ C$)	Label
6.1	1450	22.4	Safe
5.2	720	27.1	Unsafe
6.3	1650	21.8	Safe
5.0	500	29.2	Unsafe
6.0	1300	23.1	Safe

Plant Disease Image Dataset A lettuce leaf disease dataset fetched from Kaggle served as the basis for the research [15]. The dataset features multiple images of lettuce leaves with various plant diseases alongside images of healthy leaves. The images contain visible expressions of disease symptoms including leaf discoloration and mildew growth together with leaf spots and other widespread pathological issues which affect lettuce crops. The dataset functioned as a research tool for agricultural disease detection while supporting machine learning model development for crop disease recognition.

The raw dataset is a wide format with several disease types namely bacterial diseases like downy mildew in lettuce, powdery mildew in lettuce, Septoria blight in lettuce viral diseases, wilt and leaf blight in lettuce and pictures of shepherd purse weeds often found in the vicinity of lettuce cultivated areas. The categories denote different plant health issues which could reduce the lettuce production and its quality.

In the Study, the transformed data sets into a binary classification problem for the disease detection problem. Disease-related text classes were consolidated into a single “unhealthy” category for a unified healthy as “healthy,” all while the “healthy” class was kept on the form as is in the given dataset. The model is trained to distinguish a healthy plant from one that is diseased rather than to distinguish a diseased plant from several types of diseases. The dataset used in all the experiments included 621 images with 522 images of healthy lettuce leaves & 99 images of sick (all diseases taken as one class) plant leaves.

Before deep learning model training, all images were resized to 224×224 pixels to match the input specifications of MobileNetV2 architecture. On top of that, some of the Data Augmentations, which includes image rotations, Horizontal Flip, zooming and some brightness changes are applied to make dataset more diverse and also for better generalization ability on the model.

3.3 Data Preprocessing

Some preprocessing was performed on the data before training machine learning models to ensure high-quality data and reliable results. The dataset started by filtering out any invalid readings which included readings of very high or very low pH or temperature values. Duplicate records were also removed because they could introduce bias during model training. The dataset received labels based on optimal hydroponic nutrient ranges which experts commonly advise for lettuce cultivation. The safe growing conditions for lettuce included samples that had pH values between 5.5 and 6.5 along with EC readings between 800 and 2000 $\mu S/cm$ and temperatures between 18 and 26 $^\circ C$. We denoted the samples satisfying all three criteria as *safe* and those beyond these limits as *unsafe*. We used under sampling to balance the safe vs unsafe classes of the imbalanced dataset.

3.4 Machine Learning Models

Water quality classification effectiveness led researchers to evaluate multiple machine learning algorithms that work well with structured tabular data sets. The models considered in the study include Logistic Regression, Decision Tree, Random Forest and Gradient Boosting Classifier. The selection was based on their ability to represent linear and ensemble-based learning methods that have shown high performance in environmental monitoring and classification.

The datasets were partitioned randomly into training set to test data subsets using 80:20 split to guarantee valid model evaluation. We also standardized the numerical features before training with standardization to guarantee all numerical features contributing equally to model training. At train time, we performed five-fold cross-validation as well to test model stability and alleviate the danger of overfitting. The final model was selected according to the evaluation metrics of classification performance, i.e. accuracy, F1-score and ROC-AUC.

3.5 Deep Learning Model for Disease Detection

For plant disease detection, a CNN based on the MobileNetV2 architecture was employed. MobileNetV2 was selected as the base network because of its lightweight structure along with its superiority in image classification.

The pretrained ImageNet weights served as a basis for transfer learning. The network was fine-tuned by adding new classification layers which were trained on the lettuce disease dataset.

The input images were resized to 224 × 224 and normalized before training. Rotations and flip are some data augmentation techniques which are used for improving the performance of generalization. Such a model predicts how healthy or unhealthy a plant leaf would look visually from an image.

4 Results and Discussion

The section presents the experimental results for machine learning models used to classify hydroponic water quality and the deep learning model created to detect plant diseases. The findings show the system to be an effective method for monitoring hydroponic nutrient levels while detecting plant health problems. The model’s performance was assessed through several classification metrics including accuracy, precision, recall, F1-score and ROC–AUC.

4.1 Water Quality Classification Results

To identify the best model for classifying hydroponic water quality conditions, various machine learning algorithms were tested. In the research, the following models were tested using Logistic Regression and Decision Tree and Random Forest and Gradient Boosting classifiers. The models were trained on a processed dataset that included pH values and electrical conductivity readings and temperature values and engineered interaction features.

There were 70,626 samples in the dataset that was used to train the model after preprocessing and balancing. The data was split into training and test sets at 80:20. Five-fold cross-validation was also used to make sure that the models generalized well and to minimize the risk of overfitting. Table 2 presents the performance comparison of the evaluated machine learning models.

Table 2: Performance comparison of machine learning models for water quality classification

Model	Accuracy	F1 Score	ROC-AUC
Logistic Regression	92.83%	0.9290	0.9777
Decision Tree	95.16%	0.9518	0.9867
Random Forest	95.52%	0.9554	0.9909
Gradient Boosting	**95.58%**	**0.9561**	**0.9909**

The results also indicate that for the hydroponic water quality, ensemble models are better than linear models. Among the investigated algorithms, The Gradient Boosting Classifier had the best accuracy and F1 score indicating strong predictive performance in predicting both safe and unsafe nutrient conditions. ROC-AUC score is high, indicating good discrimination between the two classes.

To further analyze the performance of the selected model, a confusion matrix was generated for the Gradient Boosting classifier. The confusion matrix provides insight into the classification errors made by the model. The confusion matrix depicted in Fig. 3 demonstrates that the model successfully identifies the majority of samples across both categories. The model shows strong capability for detecting nutrient solution conditions harmful to plant growth.

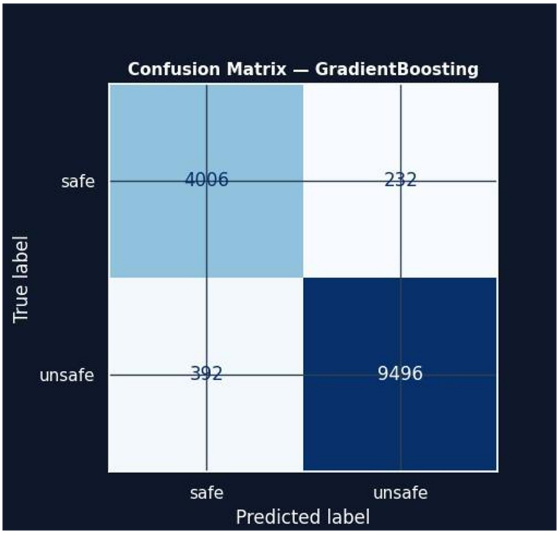

Fig.3: Confusion matrix for the Gradient Boosting water quality classification model

Receiver Operating Characteristic (ROC) analysis was performed to assess the classification performance of the models. The ROC curve plots the relationship between the true positive rate and the false positive rate at various classification thresholds. Fig. 4 explains that Gradient Boosting and Random Forest models achieved the highest ROC-AUC scores. The curves indicate that the models demonstrate excellent discriminative power for differentiating between safe and unsafe hydroponic water conditions.

4.2 Plant Disease Detection Results

The deep learning model based on the MobileNetV2 architecture was trained to detect plant diseases from lettuce leaf images. It is used to transfer learning with ImageNet pre-trained weights and fine-tuned the network with the lettuce disease dataset. Data augmentation operations improved the stability of model. The

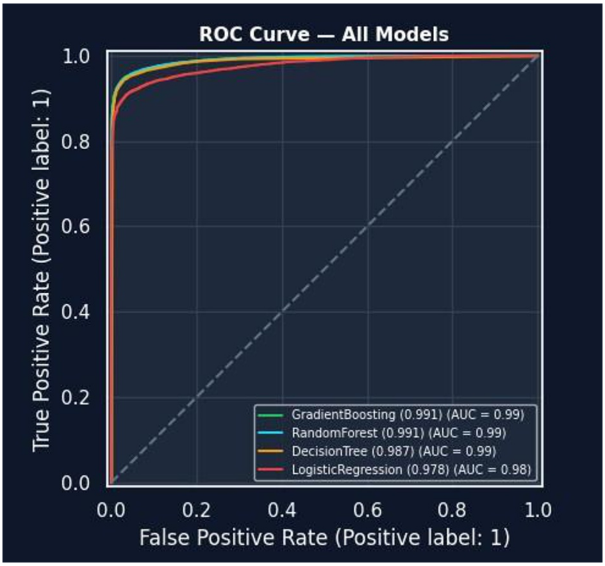

Fig.4: ROC curves of the evaluated machine learning models

validation accuracy was 98.40% when the validation accuracy stopped increasing, which confirmed that the model performed well in distinguishing between healthy and infected plants. Table 3 shows validation dataset's evaluation metrics.

Table 3: Performance of the MobileNetV2 disease detection model

Evaluation Metric	Score
Accuracy	98.40%
Precision	95.00%
Recall	95.00%
ROC-AUC	0.9976

The classification report indicates that the model performs particularly well in identifying healthy leaves with an accuracy of 99.05%, while also maintaining strong performance in detecting unhealthy plants with an accuracy of 95.00%. These results demonstrate that the deep learning model can effectively identify visual disease symptoms in lettuce leaves.

System performance was assessed with a disease detection system for the lettuce crop with predictive model validation outputs. The mobile application receives the lettuce leaf images as input and uses MobileNetV2 CNN for image processing to achieve classification results with corresponding confidence levels. The output prediction also contains a probability distribution which represents the likelihood of the leaf sample being classified to a specific class.

The model predicts that the input lettuce leaf sample belongs to an unhealthy class with a confidence of 93.3% as shown in Fig. 5. The probability distribution clearly shows that the unhealthy class has a much higher probability than the healthy class which proves the model's capability of identifying disease symptoms through visible indicators such as leaf spots and tissue damage. The illustration shows how such a system can be useful to farmers and researchers in the early diagnosis of crop health issues.

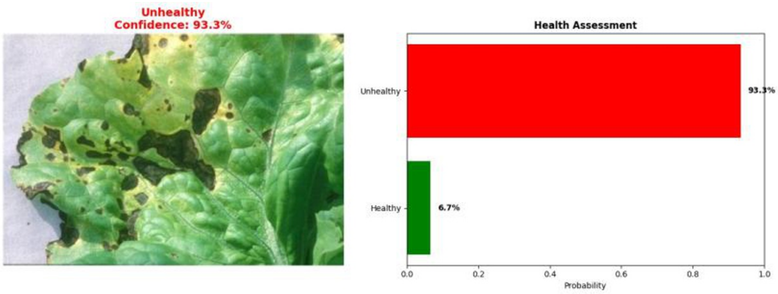

Fig.5: Example output of the proposed disease detection system showing a lettuce leaf classified as unhealthy with a confidence score of 93.3%.

4.3 System Implementation and Dashboard

Trained models were integrated into an interactive real time hydroponic system monitoring dashboard that offers live updates on the hydroponic system environment. The application enables the users to observe water quality factors and disease identification capabilities from plant images and visualize model predictions.

The dashboard has components for water quality observation, plant disease identification and alerting. Users can submit plant images for disease prediction services and evaluate the safety of their nutrient solution using sensor readings. It allows for convenient oversight and quick response to malfunctions in hydroponic farming systems through predictions generated by machine learning that can be controlled through an easy-to-use interface. Figure 6 shows the web-based dashboard interface developed for monitoring hydroponic system parameters and performing plant disease detection.

The experimental results show that the proposed system is able to effectively determine the hydroponic water quality state and identify plant diseases by means of image analysis. The combination of these models in a dashboard interface also increases the system's usability with farmers, researchers being able to use an intelligent decision-support system for hydroponic crop management.

Fig.6: Web-based dashboard for hydroponic system monitoring and disease detection

5 Conclusion and Future Work

In the research, a smart hydroponic monitoring system that utilized machine learning and deep learning approaches for water quality evaluation and plant disease detection was investigated. Hydroponic farming depends on optimal nutrient solution because pH, EC and temperature parameters directly affect plant development and nutrient absorption. The work developed a data-driven methodology to classify hydroponic water quality status and to identify plant health problems through image analysis.

Evaluation of multiple machine learning algorithms for classification of water quality using merged hydroponic, and IoT-based aquaponic systems data sets. Various models including (Logistic Regression, Decision Trees, Random Forest, Gradient boosting) were trained and compared after Preprocessing and feature engineering and several of them. The Gradient Boosting model produced the best results at around 95.58 % of accuracy and an almost perfect ROC-AUC score and also the superior capability to discriminate the safety from the in-safety conditions in HPNs water and nutrient composition.

The study includes water quality monitoring and develops a MobileNetV2based deep learning model for detecting lettuce diseases. The model reached a validation accuracy of 98.40% by utilizing the publicly available lettuce leaf disease dataset during its training process. The model shows that it can effectively detect plant health problems through visible leaf symptom analysis. The trained models were integrated into a web-based dashboard which allows users to track hydroponic conditions while uploading plant images for disease detection and receiving prediction results along with recommended actions.

The proposed system demonstrates that deep learning, machine learning and web-based visualization can be combined to support intelligent hydroponic farming. The proposed system enables detection of unfavorable nutrient conditions/plant diseases at the earlier stages, thus helping farmers to improve crop management and avoid potential loss of yield.

Future work is targeted to enhance the system capability to monitor extra environmental parameters including humidity, dissolved oxygen, nutrient concentration for full hydroponic management. Additionally, time series forecasting models could be integrated to predict forthcoming water quality and support predictive/preemptive system maintenance. Finally, in future studies a multiple disease classification will be evaluated for disease identification of affected plants and use real-time IoT sensor data in real fully automated hydroponics for the farmers.

References

1. Howard M. Resh. *Hydroponics: The Soilless Way to Garden*. Prentice-Hall, 1963.
2. Merle H. Jensen. Hydroponics worldwide. *Hort Technology*, 7(4):318–321, 1997.
3. Biology Reader. Types of hydroponic systems, 2020. Accessed: 2026-03-05.
4. L.I. Trejo-Tellez and F.C. Gomez-Merino. Nutrient solution management in hydroponic systems. *In: Nutrient Solutions for Hydroponics*, 2012.
5. Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
6. Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.

7. Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*, 2016:1–11, 2016.
8. Sharada P. Mohanty, David P. Hughes, and Marcel Salath´e. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7:1419, 2016.
9. M. Rahman et al. Aiot-based smart hydroponic monitoring system. *Journal of Smart Agriculture Systems*, 2024.
10. S. Mehare and P. Gaikwad. Classification models for smart farming applications. *International Journal of Agricultural Informatics*, 2022.
11. K. Gokul et al. Automation in hydroponic systems using iot and machine learning. *IEEE Access*, 2023.
12. R. Singh et al. Iot-based smart hydroponic monitoring with mobile integration. *International Journal of Smart Agriculture*, 2025.
13. HydroGrowNet Research Group. Hydrogrownet of batavia dataset, 2023. Mendeley Data.
14. Mendeley Data. Water quality parameters for lettuce aquaponic system (iot), 2023.
15. A. J. Star. Lettuce diseases dataset, 2022. Kaggle Dataset.