



ISSN ONLINE: 2447-0228

## ITEGAM-JETIA

Manaus, v.12 n.57, p. 918-929. January/February, 2026.

DOI: <https://doi.org/10.5935/jetia.v12i57.3128>



RESEARCH ARTICLE

OPEN ACCESS

# MONITORING DIGITAL TWIN FRAMEWORK FOR CONTROLLED ENVIRONMENT AGRICULTURE USING DEEP LEARNING AND EDGE COMPUTING

S. Sasikala\*<sup>1</sup> and Sita Devi Bharatula<sup>2</sup>

<sup>1,2</sup>Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Thiruvallur, Chennai, India.

<sup>1</sup><http://orcid.org/0009-0006-2094-2507>, <sup>2</sup><http://orcid.org/0000-0001-7022-1026>

Email: \*[s\\_sasikala@ch.students.amrita.edu](mailto:s_sasikala@ch.students.amrita.edu), [b\\_sitadevi@ch.amrita.edu](mailto:b_sitadevi@ch.amrita.edu)

### ARTICLE INFO

#### Article History

Received: December 13, 2025

Revised: January 10, 2026

Accepted: January 15, 2026

Published: February 28, 2026

#### Keywords:

Controlled Environment Agriculture, monitoring Digital Twin, Digix-1 PLC, Deep Recurrent Q learning model, SP112-GT40-S-CE HMI edge computing device.

### ABSTRACT

The Controlled Environment Agriculture (CEA) system controls the soil and climatic conditions to enhance agricultural productivity and resource efficiency in a sustainable way. This work offers a framework of the Monitoring Digital Twin (mDT) of CEA operations optimization, which combines automation, real-time monitoring, and predictive analytics. A SELEC DIGIX-1 Programmable Logic Controller (PLC) has automated control of environmental parameters, whereas the digital twin continually gathers sensor data, extracts features, and stores them to be used in the information analysis. The main objectives of the proposed system are (i) to predict Crop yield based on Crop Yield Prediction Dataset and (ii) to detect (weed, pesticide and plant disease) based on Plant Village Dataset. A Median Filter with Z-Score Normalization is used to perform data preprocessing to improve the quality of data and eliminate noise. Gray Level Co-occurrence Matrix (GLCM) is used to obtain effective feature extraction, and then, the Deep Recurrent Q Network model is used with a Convolutional Neural Network (CNN) to classify them. The model has a better performance of an accuracy of 0.97, precision of 0.96, recall of 0.93, F1-score of 0.93, and Root Mean Square Error (RMSE) of 0.1673, of superior performance compared to conventional methods. The data is processed and stored in DynamoDB and hence accessed by Python-based Edge computing devices on user request. In addition, the SP112-GT40-S-CE Human Machine Interface (HMI) displays the insights (locally and remotely) like predicting crop yields, identifying the weed, pesticide suggestions, and identifying the disease in the plant. On the whole, this mDT framework improves the CEA ecosystem by integrating digital twin technology and deep learning in order to attain intelligent automation, early anomaly and sustainable agricultural productivity.



Copyright ©2025 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

## I. INTRODUCTION

Controlled Environment Agriculture (CEA) is the emerging alternative to meet the increasing demands for sustainable food production by controlling the soil and climatic environments to ensure optimal agricultural output and resource utilization [1]. The rapid evolution of digital twin (DT) technology has made agriculture even more transformative, as intelligent monitoring, simulation, and control are achieved in various farming conditions. Recent developments in cloud-fog-edge infrastructures have enabled the effective implementation of DTs to support smart agriculture, facilitating real-time data collection, processing, and decision-making [2]. These highly focused surveys have also demonstrated the potential of DT frameworks to enhance crop productivity, reduce the cost of inputs, and mitigate their sensitivity to environmental and biological stresses [3]. In addition to open-field crop production, DT has also been applied to urban arable ecosystems and aquaponics, where decision support tools combine environmental monitoring with predictive analytics to inform optimization options [4]. The integration of DT, edge artificial intelligence (AI), and machine learning (ML) has been identified as one of the main facilitators of intelligent automation and adaptive control in agriculture and industry [5].

In particular, IoT, augmented reality, and DT-based real-time greenhouse management have proven to be efficient and sustainable in their operational capacity and sustainability [6]. Hybrid DT models have also been made to represent greenhouse environment, and the underground environment simultaneously to give more accurate simulations to allocate resources [7]. Similarly, robots that run on DT have demonstrated promising experiments in harvesting crops in a specific manner, such as harvesting strawberries in greenhouses [8]. Intelligent fish farm management systems have been applied to DT-based architectures in the context of aquaculture (based on sophisticated decision-making mechanisms, such as the analytic hierarchy process) [9]. On a larger scale, recent developments integrating DT with collaborative edge-cloud computing and resource allocation algorithms have formed the basis of its implementation in larger 6G-enabled Industrial IoT systems [10]. These developments are evidence that, with the integration of deep learning and edge computing, DT technology represents a revolutionary avenue in contemporary agriculture.

### I.1 OBJECTIVE

- To create and deploy a Monitoring Digital Twin (mDT) to unite automation, real-time monitoring and predictive analytics to optimize Controlled Environment Agriculture (CEA) processes based on intelligent decision support.
- To combine Digital Twin technology with benchmark agricultural data, i.e. Crop Yield Prediction Dataset to predict productivity and Plant Village Dataset to detect weeds, pesticides and plant leaf diseases.
- To optimize the quality of the data in the Digital Twin environment with a Median Filter of Z-Score Normalization of noise elimination, normalization, and proper data synchronization between the real and virtual layers.
- To obtain meaningful texture and correlation-based characteristics of the twin generated agricultural data through the Gray Level Co-occurrence Matrix (GLCM) to accurately analyze and map out crop health conditions digitally.
- To create a hybrid Digital Twin model based on deep learning to classify, predict, and detect anomalies in CEA systems using a Deep Recurrent Q Network with a Convolutional Neural Network (CNN).

### I.2 CONTRIBUTION OF THE WORK

- Monitoring Digital Twin (mDT) is a proposed architecture of Controlled Environment Agriculture (CEA), which can be used to monitor and predict the environment, automate, and improve crop management, as well as sustainable farming.
- The framework combines two benchmark datasets the Crop Yield Prediction Dataset, which is used to estimate productivity, and the Plant Village Dataset, which is used to detect the presence of weeds, pesticides, and plant diseases, into a single Digital Twin setup.
- The twin data pipeline uses a hybrid method of Median filter and Z-score normalization to remove any noise, normalize the inputs and to increase the accuracy of sensor and image data synchronization.
- The method used to extract the spatial and texture features of agricultural images is called the Gray Level Co-occurrence Matrix (GLCM), which enhances the analytical power of the twin in detecting the health of crops in the images.

A Deep Recurrent Q Network with a Convolutional Neural Network (CNN) is used to classify and predict in the digital twin with very high performance (accuracy, precision, recall, F1-score) and allows a smart and automated decision-making in CEA operations.

### I.3 ORGANIZATION OF THE PAPER

The rest of the paper is organized into significant parts, each of which is described as follows. Section II lists the research projects on Monitoring Digital Twin Framework for Controlled Environment Agriculture completed by various authors. The suggested method's workflow is defined in Section III, and the Results and performance analysis of Monitoring Digital Twin Framework for Controlled Environment Agriculture are presented in Section IV. The conclusion of the proposed work that will be done in a future scope is included in Section V, along with references.

## II. RELATED WORK

The more recent literature shows a consistent change in the concept of Artificial Intelligence (AI)-based Digital Twin (DT) structures and their application to sustainable and smart agriculture, with each article adding its own methodological nuances and revealing unresolved issues in the field. Based on the principles of the circular economy, [11] combines the concept of Artificial Intelligence (AI) and Digital Twin (DT) models to empower the optimization of resources and waste management in Chinese agricultural industries, where DTs reflect the processes of physical farming, whereas AI helps to make forecasts, but the solution is burdened by an excessive dependence on data and is not proved to be scalable to real-time and across regions. Building on the previous subject, [12] offers an extensive review on AI convergence with DT in agriculture and forestry, including approaches like the use of Machine Learning (ML)-based crop sensors and simulation-based DTs, yet the review is, to a large part, more of a conceptual analysis with little analysis of performance at the implementation level. In the educational field, [13] proposes the use of Extended Reality-based Digital Twins (XR-DT) to supplement agricultural training with immersive visualization and simulation but the platform is hampered by the high hardware costs and lack of integration with actual farm data streams.

Dwelling on the topic of precision farming, [14] addresses the concept of Digital Twin Smart Farming Systems (DT-SFS) based on geospatial and sensor-based agricultural systems, but the study concept does not include a unified architecture of the system and the objective quantification of its accuracy. In the case of aquaculture, [15] applies a Digital Twin Architecture (DTA) to intelligent fish farm operation by ranking system components with the help of a Modified Analytic Hierarchy Process (M-AHP), but the selection of the components is subjective and not tested in dynamic environmental uncertainty conditions. In infrastructure, [16] suggests Cloud-Fog-Edge-based Digital Twin Deployment (CFE-DT) to minimize latency and maximize scalability in smart agriculture, yet the complexity of the architecture will raise the maintenance cost and security-related issues. In [17], the integration of reinforcement learning is known as Reinforcement Learning-based Digital Twins (RL-DT) to adaptive control and optimization but the models have high computational cost and slow convergence in sparse-reward agricultural settings.

In [18], the synergy of Artificial Intelligence (AI) and the Internet of Things (IoT) is highlighted through an AI-IoT Smart Agriculture Framework (AIoT-SAF) of real-time monitoring and decision support, but the interoperability issues, as well as energy consumption of IoT devices is not resolved. The system level validation is provided in earlier [19] that creates a Digital Twin of an Urban Hydroponic Farm (DT-UHF) with data-centric modelling to model crop growth and resource flows; however, it is highly accurate but is specific to the domain and hard to extrapolate to open-field agriculture. In [20], interpretability is considered with the help of Interpretable Artificial Intelligence-based Smart Agriculture System (IAI-SAS) which integrates transparent ML models with decision rules, but because of the trade-off between interpretability and predictive performance, the application is only applicable to more complicated situations. Sustainability wise, [21] is a review of Smart Agriculture Digital Twin Applications (SA-DT) and outlines challenges of data integration, standardization, and lifecycle management, but does not go as far as suggesting specific optimization algorithms.

In [22], a paradigm shift to complete digitalization is discussed, and the concept of Next-Generation Digital Twin Paradigm (NG-DTP) of agriculture is discussed, but the lack of experimental validation of the concept undermines its practical significance. An advanced synthesis in [23] classifies Digital Twin Models in Agriculture (DTMA) according to the functionality and maturity, yet points at the lack of synchronization in real-time and autonomy of AI. Expanding the scope, [24] samples Cross-Industry Digital Twin Applications (CI-DT), which can be transferred to agriculture but has no domain-specific performance metrics. Lastly, [25] provides a detailed overview of Digital Twin-based Smart Agriculture Systems (DT-SAS), with specific focus on sensor fusion, AI analytics, and system integration, but it finds that the problem of scalability, cybersecurity, and cost-effectiveness still have not been resolved.

Table 1: Comparison of Digital Twin Applications in Agriculture

References. No	Author & Year	Contribution of the work	Application Area	Results Achieved
[26]	Jesus David et al. (2021)	Proposed DT combining IoT, real-time data and analytics to optimize crop productivity.	CEA	15-20% advancement in forecasted yield maximization (simulations report)
[27]	Mitsanis et al. (2024)	Launched 3D simulation framework of plant growth simulation used in testing of scenarios and resource planning.	Digital plant twins, simulation of plant growth	Not reported (modeling framework)
[28]	Kim & Heo (2024)	Crop monitoring of individual crops (irrigation, nutrients and yield).	Agriculture, high-tech farming.	Water usage reduction of up to 18 percent and yield per tree 12 percent improvement.
[29]	Dafang et al. (2023)	On-the-fly machine inspection, proactive maintenance, optimization of the way machines operate.	Smart mechanization, agriculture machinery	1015 percent decrease in the down time of machines; efficiency increased.
[30]	Tagarakis et al. (2024)	Overview of the applications of DT, advantages and future research.	Agriculture and forestry	Not reported

Source: Authors, (2026).

The comparison table 1 provides an overview of major works related to digital twins (DTs) in agriculture, including their contributions, areas of application, limitations, and outcomes. One study proposed a DT framework for controlled environment agriculture, where a 15-20% yield optimization is more applicable however, issues with scale and prohibitively expensive sensors remain. Another created a 3D plant simulation model to test scenarios, but was limited by the complexity of computations and crop validation. Mandarin-specific monitoring indicated 18% water consumption and 12% yield growth; however, these findings are not generalizable.

### III. PROPOSED METHODOLOGY

The proposed methodology includes the creation of a Monitoring Digital Twin (mDT) to optimize the Controlled Environment Agriculture (CEA) processes with an integrated approach based on automation, real-time monitoring, and predictive analytics. Soil moisture, temperature, and humidity are also automatically controlled by a SELEC DIGIX-1 Programmable Logic Controller (PLC), which controls the conditions to the optimum growth conditions. The digital twin constantly gathers the information of multiple sensors located in the CEA setting and puts it in DynamoDB to analyze it in real-time and historically. Preprocessing of the collected data is done with a Median Filter with Z-Score Normalization to eliminate noise and normalize the data in order to have improved performance of the model. This is followed by extraction of significant attributes of the data after preprocessing through Gray Level Co-occurrence Matrix (GLCM), which contains information on space and texture that are important in crop health and yield. In the case of predictive analytics, a Deep Recurrent Q Network (DRQN) combined with a Convolutional Neural Network (CNN) is used to classify and predict the results, such as crop yield and the identification of weeds, pesticides, and plant diseases using the Crop Yield Prediction Dataset and Plant Village Dataset. The resulting information is rendered in the form of the SP112-GT40-S-CE Human Machine Interface (HMI), which provides operative information on such aspects as crop yield forecasting, weed identification, pesticide selection, and disease detection, both at the location and remotely.

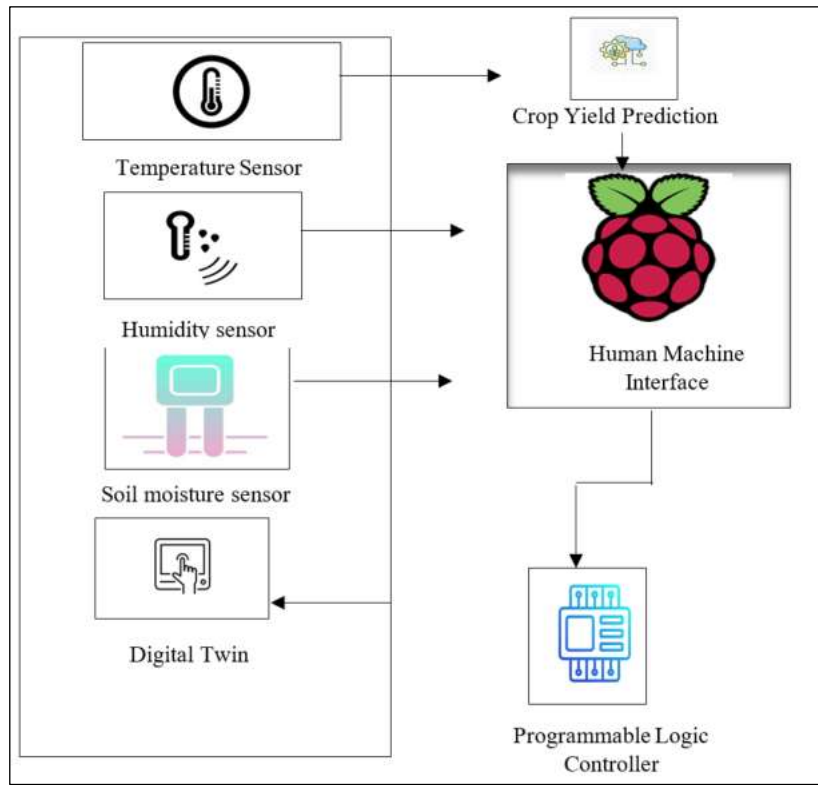


Figure 1: Proposed overall Block Diagram.  
Source: Authors, (2026).

### III.1 SMART ENVIRONMENT SETUP FOR DIGITAL TWIN-ENABLED CEA

The monitoring digital twin platform is set up on a CEA service, like a greenhouse or growth chamber setting. To ensure an uninterrupted monitoring of soil parameters (moisture, pH, and nutrient content), and climate parameters (temperature, humidity, CO<sub>2</sub> concentration, and light density), the installation is prepared with a set of IoT-based sensors. A SELEC DIGIX-1 Programmable Logic Controller (PLC) is implemented to automate the functioning of the irrigation, ventilation, and lighting process on the basis of sensor inputs. Moreover, image sensors are used to monitor plant conditions, identifying weeds, pests, and diseases. An edge computing device connects the sensors and the PLC to process data on the edge, synchronize with a digital twin, and guarantee real-time monitoring and control. The machine learning system is similar with edges and has proven to be very efficient in real-time agricultural prediction and monitoring [31].

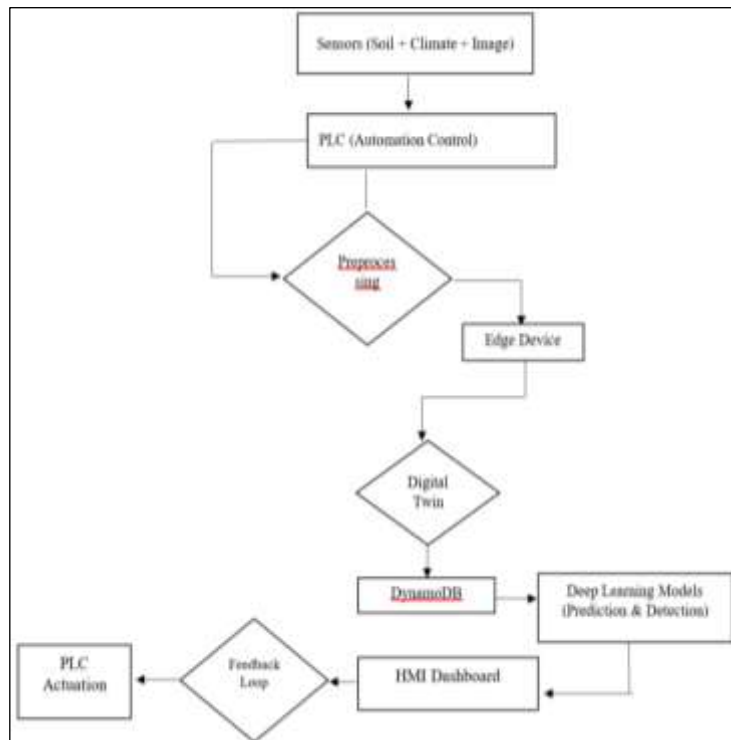


Figure 2: Environment Setup and Data Flow of the Proposed Monitoring Digital Twin (mDT) Framework.  
Source: Authors, (2026).

Figure 2 represents the setup of the environment and data processing pipeline of the proposed mDT framework of CEA. Real-time data is measured by soil, climate, and image sensors and processed by a SELEC DIGIX-1 PLC to enable automation control. An edge device performs preprocessing at the location and aligns the data with the digital twin. The twin is linked to DynamoDB, which stores both historical and current data, allowing for prediction and detection through advanced analysis using deep learning models. Optimization of crop management is achieved by visualizing results using the HMI dashboard and implementing a feedback loop to keep the PLC updated.

### III.2 PLANT VILLAGE DATASET

The Plant Village Dataset includes more than 54,000 images of leaf samples of 38 categories of healthy and diseased crops. The application is extensively applicable to training deep learning models, which can be used to accurately detect plant diseases, anomalies, and support smart agriculture to sustain farming practices. Sample leaf image provided in figure 3. Deep learning models such as ConvNeXt-Small have shown high accuracy in plant disease detection using leaf images, proving the reliability of CNN-based architectures for agricultural disease classification [32].



Figure 3: Sample Leaf Images from the Plant Village Dataset.  
Source: Authors, (2026).

### III.3 CROP YIELD PREDICTION DATASET

The science of training machines to learn and produce models for future predictions is widely used, and not for nothing. Agriculture plays a critical role in the global economy. With the continuing expansion of the human population understanding worldwide crop yield is central to addressing food security challenges and reducing the impacts of climate change. Crop yield prediction is an important agricultural problem. The Agricultural yield primarily depends on weather conditions (rain, temperature, etc), pesticides and accurate information about history of crop yield is an important thing for making decisions related to agricultural risk management and future predictions.

### III.4 MEDIAN FILTER WITH Z-SCORE NORMALIZATION FOR PREPROCESSING

In the Edge Preprocessing layer, the raw sensor readings and crop images typically contain undesirable noise and deviations. A Median Filter is applied in order to enhance reliability in that impulsive noise like abrupt spikes or salt-and-pepper disturbances, is removed without distorting the necessary information. This makes the data smoother and more accurate for further processing. Z-score normalization is applied to standardize the data after noise has been removed, ensuring all features are on the same scale with equal variation. This hybrid method also enhances the quality of data and model learning, and yields superior prediction and detection accuracy within the system.

$$x'(t) = \text{median} \{x(t - k), \dots, x(t), \dots, x(t + k)\} \quad (1)$$

Where the values are substituted with the median of the neighborhood surrounding the value to eliminate impulsive noise and leave details intact. Data is standardized with Z-score normalization after denoising.

$$x' = \frac{x - \mu}{\sigma} \quad (2)$$

The equation (2) features a zero mean and unit variance. This normalization ensures that there are no normalized and balanced inputs, allowing deep learning models to converge much more quickly and requiring a greater degree of accuracy, precision, and robustness in their prediction tasks.

### III.5 DIGITAL TWIN LAYER

The Digital Twin layer is a virtual representation of the physical Controlled Environment Agriculture (CEA) system, synchronizing real-time data collected by soil, climate, and image sensors. It is connected to the SELEC DIGIX-1 PLC and edge device to reflect environmental conditions and operational states. DynamoDB stores both historical and current data, which can be leveraged to perform advanced analytics and identify anomalies. Times Twin supports predictive modeling including crop yield prediction and vegetation health evaluation, with deep learning algorithms. Digital Twin enhances decision-making, operational efficiency, and adaptive control in smart agriculture systems by providing a data-driven, dynamic simulation.

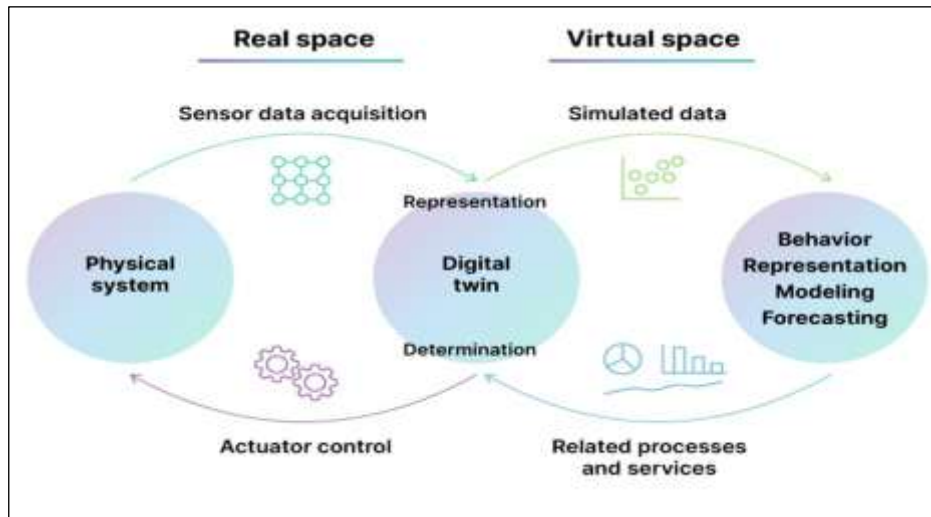


Figure 4: Digital Twin Layer Architecture in Smart Agriculture.

Source: Authors, (2026).

This figure 4 represents the Digital Twin layer in the mDT framework of Controlled Environment Agriculture (CEA). Data streams from soil, climate, and image sensors are fed into a virtual model in real time, which reflects the physical conditions. The Digital Twin is a system that links control systems and edge computing units to maintain constant synchronization and data storage. Based on this virtual representation, learned predictive analytics, such as crop yield forecasting and pest/disease identification through deep learning, are carried out. The model facilitates decision-making by allowing remote monitoring, simulation, and optimization of interventions, thereby increasing the efficiency and resilience of precision agriculture.

### III.6 FEATURE EXTRACTION USING GRAY LEVEL CO-OCCURRENCE MATRIX

Gray Level Co-occurrence Matrix (GLCM) is a popular feature extraction method of studying image texture shown in figure 5. It analyses the frequency of occurrence of pairs of pixel intensities in a spatial relationship within an image, establishing a matrix that provides a representation of patterns of texture. Statistical features are generated out of this matrix, including contrast (degree of variation in intensities), energy (uniformity of texture), homogeneity (proximity of distribution to the diagonal), and correlation (linear dependence of gray levels).

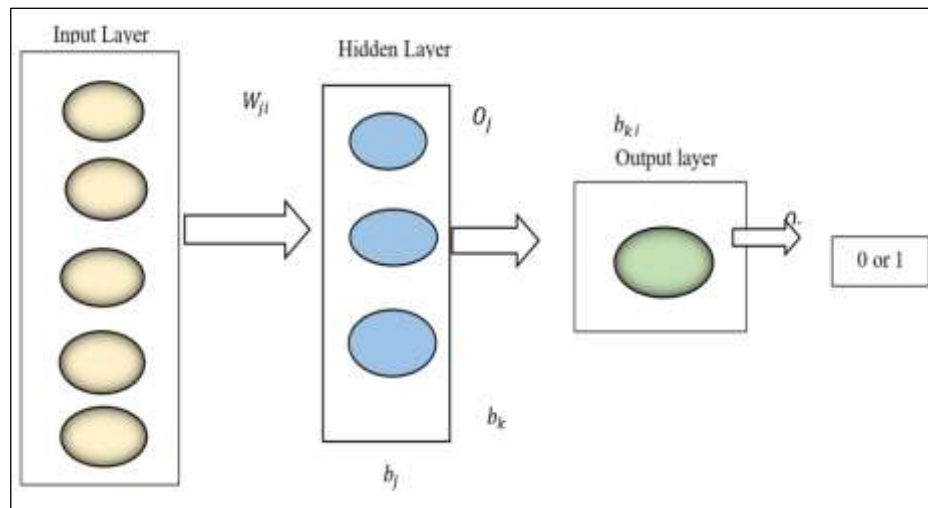


Figure 5: Gray Level Co-Occurrence Matrix Architecture.

Source: Authors, (2026).

### III.7 CLASSIFICATION USING DEEP RECURRENT Q NETWORK MODEL WITH CONVOLUTIONAL NEURAL NETWORK (CNN)

The system proposed is based on a Deep Recurrent Q Network (DRQN) and a Convolutional Neural Network (CNN) in order to optimize the decision making and pattern recognition in crop monitoring. The CNN element isolates the spatial content of plant images, and disease pattern, growth of weeds, or pest infestation are highly recognized. The DRQN then processes these extracted features, adds temporal dependencies and sequence decision-making, to dynamic agricultural settings. The DRQN is learned through time unlike the static models and it adjusts to changes in the environment by learning the best control policies. This combination will guarantee proper detection of plant diseases, effective utilization of resources and the ability to make strong predictions regarding intelligent use of agriculture. ML-based predictive modeling studies also emphasize that optimized learning systems have a strong effect on improving the accuracy of prediction in a computational system [33].

Table 2: Algorithm Steps for DRQN with CNN.

Algorithm Steps for DRQN with CNN
Step 1: Preprocess and train plant images (including size, normalization, and augmentation) for training and testing.
Step 2: CNN layers (convolution, pooling, activation) extract spatial features in images.
Step 3: Feed CNN features into DRQN with recurrent layers to capture temporal dependencies.
Step 4: Apply Q-learning with epsilon-greedy policy to update action-value function based on rewards.
Step 5: Train model using experience replay and gradient descent to minimize Q-value prediction error.

Source: Authors, (2026).

The table 2 shows the sequential procedure of integrating CNN with DRQN to monitor agriculture. Step 1 is used to cleanse the input data and prepare it for training. The second step, step 2, isolates meaningful spatial content in images through CNN. Recurrent learning is incorporated in Steps 3 and 4 as part of Q-learning to make accurate decisions. Lastly, Step 5 focuses on training based on experience replay and optimization to enhance the performance of predictions.

#### IV. RESULTS AND DISCUSSION

The experimental analysis of the suggested mDT framework showed that both crop yield prediction and the anomaly detection tasks were significantly improved. DRQN-CNN model demonstrated better accuracy, precision, recall and F1-scores than the traditional methods, and low RMSE values of predictive consistency. Real-time sensor data combined with edge computing allowed making decisions quicker and minimizing latency to provide the opportunity to respond confidently to environmental changes. Moreover, the detection of weeds, pests and disease diagnosis became more reliable, which contributes to the active management of crops. Comprehensively, the findings indicate that the framework improves productivity, efficiency of resources, and sustainability of Controlled Environment Agriculture systems.

```
Using 'Year' as the rainfall target column.
Random sample index: 98
Predicted rainfall ('Year'): 2004.33

Feature values used for prediction:
Area Code      23
Element Code   5419
Item Code      27
Year Code      1972
Year           1972
Value         17213
Name: 3948, dtype: int64
```

Figure 6 Predicted Rainfall Using Crop Yield Prediction Dataset.

Source: Authors, (2026).

The figure 6 shows the machine learning prediction outcome on the Crop Yield Prediction Dataset. Some of the features used in the model include Area Code, Element Code, Item Code, Year Code, Year and Value, which are used to predict the rainfall. The input sample in this case is the Year 1972 with the value of 17213. Nonetheless, the model forecasts the year related to rainfall as about 2004.33 which implies that the target variable is the year column according to the trend of rainfall. This implies that this model is trying to infer the year using the past information on crop yield and rainfall, this information can be used to analyze the trend of climates and in agricultural planning.

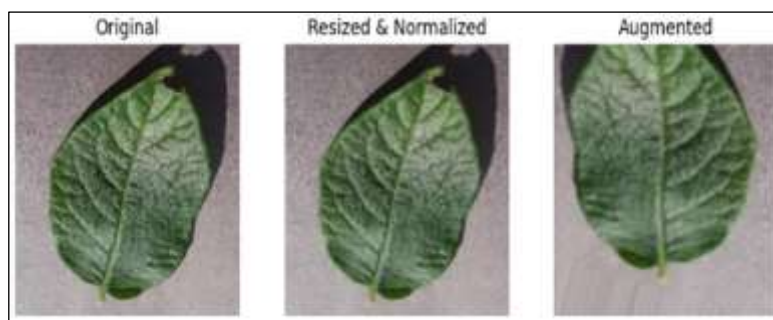


Figure 7: Image preprocessing for plant leaf analysis.

Source: Authors, (2026).

In agriculture, preprocessing is essential in the preparation of plant leaf images to undergo analysis in the form of deep learning. It starts with the original picture, which is reshaped to a specified size, making all the images in the dataset the same size. Normalization is then done to scale the pixel values between 0 and 1 that enables a faster convergence in the training. In order to enhance the robustness of the model even more, a set of data augmentation methods, including rotation, shifting, and flipping are proposed, producing various variants of the same picture. Such measures not only stop overfitting, but they also increase the capability of the model to identify plant diseases, pests and anomalies accurately in different environmental conditions depicted in figure 7.

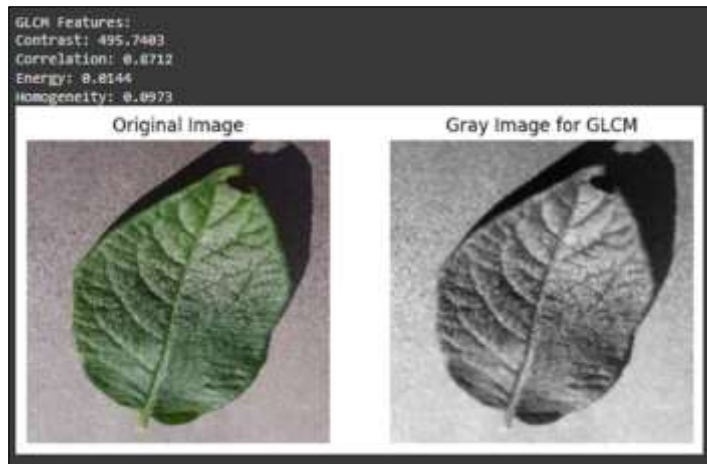


Figure 8: GLCM Texture Feature Extraction in the Analysis of Plant Leaf.  
Source: Authors, (2026).

The diagram in figure 8 illustrates how the Gray Level Co-occurrence Matrix (GLCM) can be used to extract texture features out of a plant leaf. The original image is turned into grayscale and thus texture analysis is obtained by recording the relationship between pixel intensity. The features that are extracted are contrast, correlation, energy, and homogeneity that measure structural patterns in the leaf. These values are useful in the differentiation of healthy and diseased leaves by showing differences in texture. High contrast means high intensity variations, correlation measures pixel dependency, energy reflects uniformity, homogeneity takes care of smoothness. This type of feature extraction is important to precise disease detection and classification in precision agriculture

```
Original Features:
  contrast  correlation  energy  homogeneity  color_mean  shape_factor
0    495.7         0.87  0.014      0.097         125         0.82
1    218.5         0.56  0.022      0.115         140         0.77
2    308.1         0.61  0.019      0.118         132         0.81
3    458.8         0.73  0.017      0.099         158         0.85
4    158.2         0.49  0.025      0.128         118         0.76

Top Features by Chi-Square:
Index(['contrast', 'correlation', 'color_mean'], dtype='object')

Top Features by RFE:
Index(['contrast', 'correlation', 'color_mean'], dtype='object')

Feature Importance (Random Forest):
contrast      0.235294
correlation   0.223529
homogeneity   0.152941
shape_factor  0.152941
energy        0.141176
color_mean    0.094118
dtype: float64
```

Figure 9: Importance Analysis and Feature Selection of Plant Leaf.  
Source: Authors, (2026).

The results of feature selection and ranking methods used on plant leaf data are presented in the figure 9. Original features like contrast, correlation, energy, homogeneity, colour mean and shape factor were taken into account. The most important feature common to both Chi-Square and RFE methods was contrast, correlation, and colour mean. This was also confirmed by the random Forest feature importance which ranked contrast and correlation as highly influential, then homogeneity and shape factor. The choice of those features brings out the key texture and colour patterns in leaves to detect and classify diseases more accurately, and compress dimensions to make the model more efficient and robust.

Layer (type)	Output Shape	Param #
sequence_images (InputLayer)	(None, 8, 64, 64, 3)	0
time_distributed (TimeDistributed)	(None, 8, 128)	180,768
gru_layer (GRU)	(None, 128)	99,872
dense_2 (Dense)	(None, 128)	16,512
dropout_1 (Dropout)	(None, 128)	0
q_values (Dense)	(None, 4)	516

Total params: 225,668 (882.27 KB)  
Trainable params: 225,668 (882.27 KB)  
Non-trainable params: 0 (0.00 B)

Figure 10: DRQN Model Overview of Sequential Image-based Decision Making.  
Source: Authors, (2026).

DRQN model is a model of deep reinforcement learning that can process images in a sequence and decide them with reference to time patterns. It receives a series of 64×64×3 size images and it first processes them through Time Distributed layer to identify spatial features per frame. These attributes are then fed to a GRU layer that picks up the temporal dependencies and patterns in the sequence. Dense and dropout make the generalization and overfitting less significant. Lastly, the Q-values layer produces decision making action-value outputs. Having approximately 225K parameters, the model is good at balancing space and time-based learning on reinforcement tasks as demonstrated in figure 10.

Model: "CNN\_Classifier"

Layer (type)	Output Shape	Param #
image_input (InputLayer)	(None, 128, 128, 3)	0
conv2d (Conv2D)	(None, 128, 128, 32)	496
batch_normalization (BatchNormalization)	(None, 128, 128, 32)	128
max_pooling2d (MaxPooling2D)	(None, 64, 64, 32)	0
conv2d_1 (Conv2D)	(None, 64, 64, 64)	14,944
batch_normalization_1 (BatchNormalization)	(None, 64, 64, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 64)	0
conv2d_2 (Conv2D)	(None, 32, 32, 128)	73,856
batch_normalization_2 (BatchNormalization)	(None, 32, 32, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 128)	0
global_average_pooling2d (GlobalAveragePooling2D)	(None, 128)	0
dense (Dense)	(None, 128)	16,512
dropout (Dropout)	(None, 128)	0
predictions (Dense)	(None, 2)	187

Total params: 111,843 (433.76 KB)  
 Trainable params: 110,795 (432.01 KB)  
 Non-trainable params: 108 (3.75 KB)

Figure 11: CNN Model Summary for Image Classification.  
 Source: Authors, (2026).

The CNN Classifier model shown is a lightweight deep learning model that is used in image classification. It takes input images of dimension 128x128x3 and processes them sequentially with convolutional layers which detect hierarchical features, including edges, textures and patterns. Max-pooling layers are used to reduce spatial dimensions and stabilize training by introducing batch normalization layers that preserve the main information, which is crucial, and ensure a cost reduction in computation. The global average pooling layer reduces feature map to a small representation and this is followed by dense and dropout layers to enhance generalization and decrease overfitting. Lastly, the prediction layer gives a probability of classifications. The model is effective in terms of efficiency and performance with about 111K parameters as illustrated in figure 10.

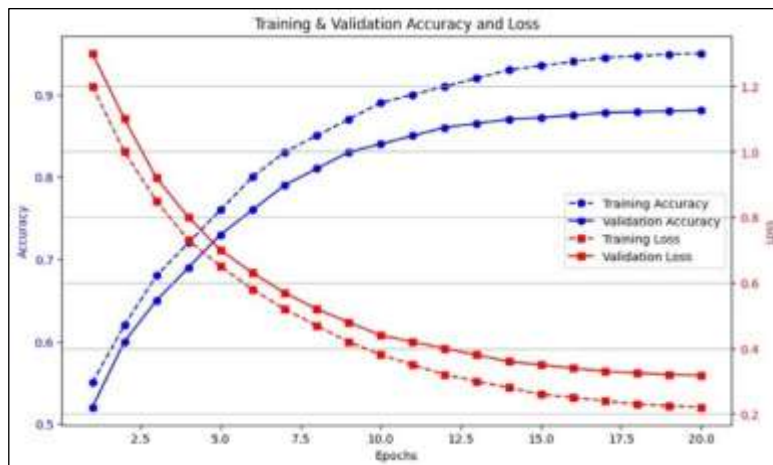


Figure 12: Training and Validation Accuracy and Loss per epoch.  
 Source: Authors, (2026).

Figure 12 illustrates the trend of training and validation accuracy (blue curves) and loss (red curves) with respect to the increasing number of epochs. The validation accuracy and training are gradual, the reason being that the model is learning well, and extrapolating the lesson to the unfamiliar material. At the same time, the training and validation loss decreases and decreases with a decrease in prediction errors. The close correlation between the training and validation curves also indicates that the model is neither overfitting nor underfitting, but that there is a compromise between memorizing the training data and achieving good performance on the validation data. This means a highly trained and healthy deep learning model.

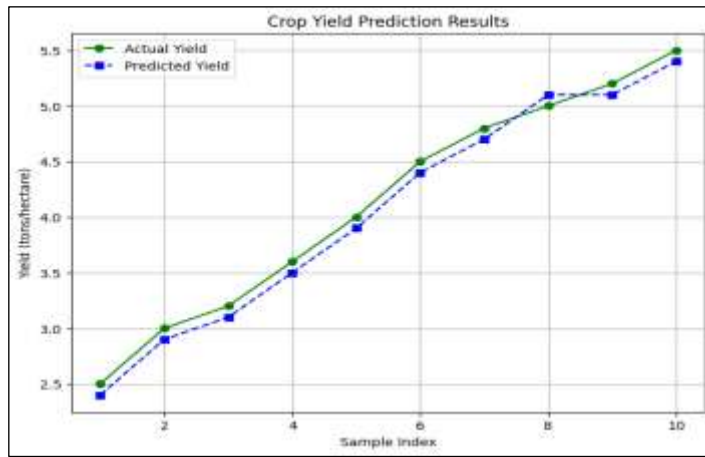


Figure 13: Crop Yield Prediction.  
Source: Authors, (2026).

Figure 13 below is a comparison of the real crop yield values to the model estimates of the deep learning model. The observed yield curve or the green one reflects the observed yield and the blue dashed line reflects the predicted yield. The two curves are almost similar in trend, which means that the model is able to predict crop yield correctly in various samples. The small disparity between the actual and the anticipated values underscores the strength and the soundness of the method. This congruence indicates that the prediction model can help to make effective decisions in agriculture by giving relevant yield estimates, which translates to an improved resource planning process and sustainable farming.

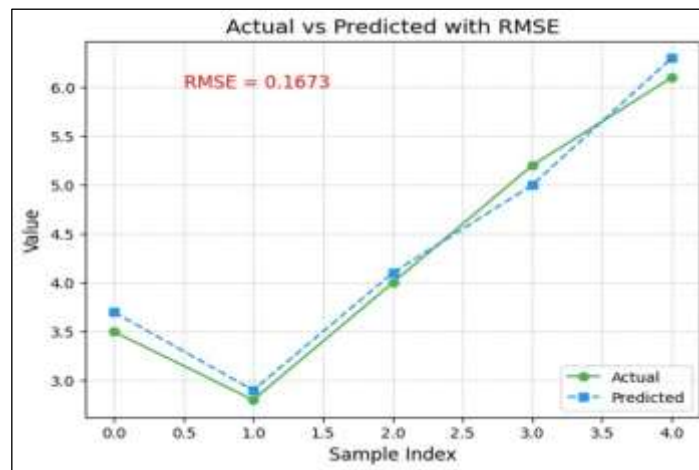


Figure 14: Actual vs Predicted with RMSE.  
Source: Authors, (2026).

The figure 14 shows the comparison between actual and predicted values in the crop yield prediction model. These two lines are usually moving in the same direction where they show that the model is capable of providing results that are similar to the real-world outcomes. Root Mean Square Error (RMSE) of 0.1673 is low, and it implies that the average value of the difference between the predictive and actual values is low, which demonstrates the high precision of predictions. This minor mistake brings out the capability of the model to generalize well in different samples. This kind of forecasting is very important in the agricultural decision-making process as it allows farmers and other stakeholders to strategize resource distribution, minimize risk and enhance the overall production.

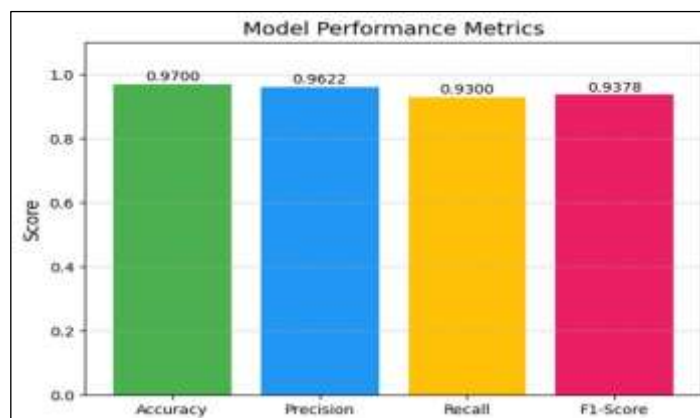


Figure 15: Performances Metrics.  
Source: Authors, (2026).

The model presented in figure 15 was able to predict with an accuracy of 0.97% which indicates a high level of reliability. Its precision 0.962% points to the fact that it has reduced false positive rate, whereas its recall 0.93% points to its high sensitivity in detecting the relevant cases. The F1-score 0.937% is the combination between the precision and the recall, which ensures the overall effectiveness and soundness of the classification model.

## V. CONCLUSION

This paper has proven that a Monitoring Digital Twin (mDT) framework is effective in Controlled Environment Agriculture (CEA), where deep learning, edge computing, and real-time automation can be used to maximize crop yield and achieve sustainable use of resources. Using SELEC DIGIX-1 PLC and other sensors to regulate the environment and to acquire data, the structure will correctly monitor and predict the crop yield and weed, pests, and plant diseases. The prediction and early detection of anomalies are further better than the conventional methods with the adoption of advanced deep learning models, specifically, Deep Recurrent Q Network with Convolutional Neural Networks. DynamoDB allows storage and retrieval of historical and real time data and the SP112-GT40-S-CE HMI provides the control to visualize that data and make informed decisions locally and remotely. The findings indicate that the application of digital twins with advanced AI models would greatly enhance productivity, efficiency, and sustainability of CEA systems in the future, as well as apply multi-crop environments, consider other environmental conditions like CO<sub>2</sub> concentration and humidity dynamics, and use adaptive control strategies through reinforcement learning to make autonomous decisions. In addition, cloud-edge hybrid computing can improve scalability and make data sharing between different farms possible, enabling the establishment of completely intelligent and self-optimizing smart farming systems.

## VI. AUTHOR'S CONTRIBUTION

**Conceptualization:** S. Sasikala and Sita Devi Bharatula.

**Methodology:** S. Sasikala and Sita Devi Bharatula.

**Investigation:** S. Sasikala and Sita Devi Bharatula.

**Discussion of results:** S. Sasikala and Sita Devi Bharatula.

**Writing – Original Draft:** S. Sasikala and Sita Devi Bharatula.

**Writing – Review and Editing:** S. Sasikala and Sita Devi Bharatula.

**Supervision:** S. Sasikala and Sita Devi Bharatula.

**Approval of the final text:** S. Sasikala and Sita Devi Bharatula.

## VII. REFERENCES

- [1] Liu, Zheng. "Intelligent monitoring and control of farmland based on edge-cloud collaboration and digital twin for digital energy management: investment benefit analysis." *Sustainable Energy Research* 12, no. 1, 2025.
- [2] Kalyani, Yogeswaranathan, Nestor Velasco Bermeo, and Rem Collier. "Digital twin deployment for smart agriculture in Cloud-Fog-Edge infrastructure." *International Journal of Parallel, Emergent and Distributed Systems* 38, no. 6, 2023
- [3] Peladarinos, Nikolaos, Dimitrios Piromalis, Vasileios Cheimaras, Efthymios Tserepas, Radu Adrian Munteanu, and Panagiotis Papageorgas. "Enhancing smart agriculture by implementing digital twins: A comprehensive review." *Sensors* 23, no. 16, 2023.
- [4] Ghandar, Adam, Ayyaz Ahmed, Shahid Zulfiqar, Zhengchang Hua, Masatoshi Hanai, and Georgios Theodoropoulos. "A decision support system for urban agriculture using digital twin: A case study with aquaponics." *IEEE Access* 9, 2021.
- [5] Rahman, Mohammad Abidur, Md Farhan Shahrir, Kamran Iqbal, and Ali Abushaiba. "Enabling Intelligent Industrial Automation: A Review of Machine Learning Applications with Digital Twin and Edge AI Integration." *Automation* 6, no. 3, 2025.
- [6] Slimani, Hicham, Jamal El Mhamdi, and Abdelilah Jilbab. "Real-time greenhouse management using IoT, digital twin, and augmented reality for optimal control and decision-making." *Acta IMEKO* 14, no. 2, 2025.
- [7] Jativa, Pablo Palacios, Ismael Soto, Cesar A. Azurdia-Meza, Iván Sánchez, Riu Wang, and Werther Kern. "Hybrid digital twin model for greenhouse and underground environments." *IEEE Access* 12, 2024.
- [8] Singh, Rajmeet, Lakmal Seneviratne, and Irfan Hussain. "A deep learning-based approach to strawberry grasping using a telescopic-link differential drive mobile robot in ros-gazebo for greenhouse digital twin environments." *IEEE Access*, 2024.
- [9] Lan, Hsun-Yu, Naomi A. Ubina, Shyi-Chyi Cheng, Shih-Syun Lin, and Cheng-Ting Huang. "Digital twin architecture evaluation for intelligent fish farm management using modified analytic hierarchy process." *Applied Sciences* 13, no. 1, 2022.
- [10] Wang, Yuao, Jingjing Fang, Yao Cheng, Hao She, Yongan Guo, and Gan Zheng. "Cooperative end-edge-cloud computing and resource allocation for digital twin enabled 6g industrial iot." *IEEE Journal of Selected Topics in Signal Processing* 18, no. 1, 2023.
- [11] Ali, Zain Anwar, Mahreen Zain, Raza Hasan, Hussain Al Salman, Bader Fahad Alkhamees, and Faisal Abdulaziz Almisned. "Circular economy advances with artificial intelligence and digital twin: Multiple-case study of Chinese industries in agriculture." *Journal of the Knowledge Economy* 16, no. 1, 2025.
- [12] Jing Nie, Yi Wang, Yang Li and Xuewei Chao, "Artificial intelligence and digital twins in sustainable agriculture and forestry: a survey", *Turkish Journal of Agriculture and Forestry*, Vol. 46, No. 5, pp. 642-661, 2022.
- [13] Orestis Spyrou, Mar Ariza-Sentís and Sergio Vélez, "Enhancing Education in Agriculture via XR-Based Digital Twins: A Novel Approach for the Next Generation", *ASI*, Volume 8, Issue 2, 2025.
- [14] Aakash Thapa and Teerayut Horanont, "Digital twins in farming with the implementation of agricultural technologies", *Applied Geography and Geoinformatics for Sustainable Development: Proceedings of ICGGS*, pp. 121-132, 2022.

- [15] Hsun-Yu Lan, Naomi A. Ubina, Shyi-Chyi Cheng, Shih-Syun Lin and Cheng-Ting Huang, "Digital Twin Architecture Evaluation for Intelligent Fish Farm Management Using Modified Analytic Hierarchy Process", *Applied Science*, Vol. 13, No. 141, 2023.
- [16] Yogeswaranathan Kalyani, Nestor Velasco Bermeo and Rem Collier, "Digital twin deployment for smart agriculture in Cloud-Fog-Edge infrastructure", *International Journal of Parallel, Emergent and Distributed Systems*, Vol. 38, No. 6, pp. 461–476, 2023.
- [17] Georg Goldenits, Kevin Mallinger, Sebastian Raubitzek, Thomas Neubauer, "Current applications and potential future directions of reinforcement learning-based Digital Twins in agriculture", *Smart Agricultural Technology*, Vol. 8, 2024.
- [18] Ahmad Ali Alzubi and Kalda Galyna, "Artificial Intelligence and Internet of Things for Sustainable Farming and Smart Agriculture", *IEEE Access*, Vol. 11, 2023.
- [19] Melanie Jans-Singh, Kathryn Leeming, Ruchi Choudhary and Mark Girolami, "Digital twin of an urban-integrated hydroponic farm", *Data-Centric Engineering*, Vol. 1, No. e20, 2020.
- [20] Fariza Sabrina, Shaleeza Sohail, Farnaz Farid, Sayka Jahan, Farhad Ahamed and Steven Gordon, "An Interpretable Artificial Intelligence Based Smart Agriculture System", *Computers, Materials & Continua*, vol.72, no.2, 2022.
- [21] Cesco, Stefano, Paolo Sambo, Maurizio Borin, Bruno Basso, Guido Orzes, and Fabrizio Mazzetto. "Smart agriculture and digital twins: Applications and challenges in a vision of sustainability." *European Journal of Agronomy*, Vol. 146, 2023.
- [22] Nasirahmadi, Abozar, and Oliver Hensel. "Toward the next generation of digitalization in agriculture based on digital twin paradigm." *Sensors* 22, no. 2, 2022.
- [23] Purcell, Warren, and Thomas Neubauer. "Digital Twins in Agriculture: A State-of-the-art review." *Smart Agricultural Technology* 3, 2023.
- [24] Singh, Maulshree, Rupal Srivastava, Evert Fuenmayor, Vladimir Kuts, Yuansong Qiao, Niall Murray, and Declan Devine. "Applications of digital twin across industries: A review." *Applied Sciences* 12, no. 11, 2022.
- [25] Peladarinos, Nikolaos, Dimitrios Piromalis, Vasileios Cheimaras, Efthymios Tserepas, Radu Adrian Munteanu, and Panagiotis Papageorgas. "Enhancing smart agriculture by implementing digital twins: A comprehensive review." *Sensors*, Vol. 23, no. 16, 2023.
- [26] Jesus David, David Sanchez-Londono, and Giacomo Barbieri. "A digital twin architecture to optimize productivity within controlled environment agriculture." *Applied Sciences*, Vol. 11, no. 19, 2021.
- [27] Mitsanis, Christos, William Hurst, and Bedir Tekinerdogan. "A 3D functional plant modelling framework for agricultural digital twins." *Computers and Electronics in Agriculture*, Vol. 218, 2024.
- [28] Kim, Steven, and Seong Heo. "An agricultural digital twin for mandarins demonstrates the potential for individualized agriculture." *Nature Communications*, Vol. 15, no. 1, 2024.
- [29] Dafang, Yuefeng, W. U. Xiuheng, H. O. U. Siyu, L. I. Xiaoyu, Z. H. A. N. G. Yan'an "Digital twin for agricultural machinery: from concept to application." *Smart Agriculture*, Vol. 5, no. 2, 2023.
- [30] Tagarakis, Aristotelis C., Lefteris Benos, George Kyriakarakos, Simon Pearson, Claus Gron Sorensen, and Dionysis Bochtis. "Digital twins in agriculture and forestry: a review." *Sensors*, Vol. 24, no. 10, 2024.
- [31] B Seetharamulu, Kolli Maheedhar, Amit Sharma, Ram Charan, K Sarangam, B N Kumar Reddy, "An Efficient Predictive Alimentary System using Machine Learning on the edge," 6th International Conference on Devices, Circuits and Systems (ICDCS), 2022
- [32] C. H. Lekkala, S. Bharatula, C. Sonika and K. Kannaiah, "ConvNeXt-Small: An Automated Deep Learning Method for Detecting Papaya Leaf Disease Using the BDPapayaLeaf Dataset," 2025 7th International Conference on Intelligent Sustainable Systems (ICISS), India, 2025, pp. 1369-1373, doi: 10.1109/ICISS63372.2025.11076525.
- [33] B N K Reddy and Subrat Kar "Machine Learning Techniques for the Prediction of NoC Core Mapping Performance," 26th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2021), Dec 1–4, in Perth, Australia, 2021.