International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 14, Issue 4, April-2026, Available online at: www.ijaresm.com

AgriTwin-GH: Agricultural Digital Twin for Smart Greenhouse Horticulture of Tomato Cultivation

Amala Margret. A[1], Dr. V. Govindasamy[2], Arjun Christopher[3], Vantapati Raja Rajeswari[4], Bhuvanalakshmi. J. P[5]

[1]Research Scholar, Department of Information Technology, Puducherry Technological University, Puducherry, India
[2]Professor, Department of Information Technology, Puducherry Technological University, Puducherry, India
[3,4,5]UG Student, Department of Information Technology, Puducherry Technological University, Puducherry, India

DOI: https://doi.org/10.56025/IJARESM.140426327

--*****************--

ABSTRACT

Growing tomatoes in a controlled greenhouse sounds straightforward, but in practice it means managing a constantly shifting web of temperature, humidity, CO_2, irrigation, and disease pressure — often all at once. Most greenhouses still handle this reactively, which means problems have already taken hold by the time anyone responds. This paper presents AgriTwin-GH, an AI-driven digital twin framework that flips that dynamic from reactive to predictive. The system fuses time-series environmental data with real-time plant image analysis to enable continuous health monitoring and early risk detection, driving a fully automated control strategy. For crop intelligence, EfficientNet-based deep learning models classify tomato growth stages and detect diseases, reaching 98.61 % and 98.23 % accuracy respectively. Weather forecasting combines a transformer-based Chronos-T5 model, XGBoost, and LSTM networks — capturing both long-range seasonal patterns and abrupt short-term shifts — with the ensemble outperforming any single model in terms of R^2 and MAE. Disease and growth progression are handled by LSTM and GRU sequence models that reliably predict symptom severity and stage transitions over 24- and 48-hour horizons. All of this feeds into a digital twin simulation layer coupled with Model Predictive Control, which continuously tests and refines actuator decisions inside a virtual copy of the greenhouse before anything is applied to the real environment. A web-based dashboard and a 3D Unity visualization surface these insights in real time. The net result is a greenhouse management system that senses earlier, predicts further ahead, and acts more efficiently — supporting higher yields, lower resource consumption, and more sustainable crop production.

Keywords: Smart Greenhouse, Growth Stage Classification, Disease Classification, Environmental Forecasting, Data-Driven Decision Support, Disease Progression Modeling, Growth Progression Modeling, Model Predictive Control, Virtual Greenhouse Simulation

INTRODUCTION

A Digital Twin is a live virtual copy of a real-world system — one that mirrors current conditions, predicts what comes next, and lets engineers safely test interventions before committing to them. In greenhouse farming, this matters enormously. Growing tomatoes well means keeping temperature, humidity, light, CO_2, irrigation, and nutrients in balance, often simultaneously, across an entire growing season.

Most greenhouses today still run on reactive logic — wait for something to drift out of range, then respond. Whether it's a human spotting a problem or a threshold alert firing, the intervention always comes after the fact. That delay means stressed crops, undetected disease, and wasted resources. The problem runs deeper too: greenhouse environments are tightly coupled, so correcting one variable can inadvertently push another the wrong way. Simple rule-based controllers struggle with this complexity.

AgriTwin-GH was built to close these gaps. It's a software-first digital twin framework for intelligent tomato greenhouse management that continuously draws together environmental time-series data, plant health signals, and AI-driven forecasts to catch risks early — before conditions deteriorate. Data flows through a real-time pipeline into a backend that makes the twin a genuine decision-support tool, not a passive display. A Unity-based 3D visualization represents crop growth stages, disease conditions, and live actuator states — vents, irrigation, lighting — all kept in sync through API communication. Control strategies can be tested against the simulation before touching the real greenhouse, reducing trial-and-error risk. Complex AI outputs are surfaced as simple visual cues — color codes, status badges, symbols — so growers can act confidently without needing to interpret raw data. The result is a genuine shift from reactive monitoring to proactive, data-driven greenhouse management.

LITERATURE SURVEY

Before any machine learning model can be reliably trained on agricultural data, the data itself needs serious attention. One body of work in this area emphasizes a multi-stage preprocessing routine: removing noisy and mislabelled records, rebalancing class distributions, randomly shuffling samples to prevent ordering bias, running correlation-based feature selection, and generating additional attributes using statistical measures such as minimum, maximum, and range [10]. After dimensionality reduction strips out redundant attributes, what remains is a leaner, more informative input space for downstream models. A separate line of work addresses plant disease diagnosis by stacking multiple CNN architectures — VGG16, VGG19, InceptionV3, and ResNet101V2 — in an ensemble learning framework [4]. Each model independently classifies leaf images from the PlantVillage dataset (54,000+ images), and their predictions are merged through voting. Augmentation — rotation, zoom, flipping — improves generalization, with ResNet101V2 reaching roughly 93.6 % accuracy.

A third line of work deals specifically with tomato growth-stage identification, combining StyleGAN3-generated synthetic images with real training data to expand dataset diversity, then using a Vision Transformer (ViT) that processes images as sequences of patches and captures global structure through self-attention [23]. Classification accuracy improves noticeably when synthetic data is included, which demonstrates that generative augmentation can compensate for limited annotated samples in specialized agricultural settings. On the forecasting side, one framework — DL2F — combines numerical weather prediction data with historical observations and feeds them into a GRU network to sharpen estimates of temperature, humidity, pressure, and solar irradiance [12]. The approach improves accuracy but stays in the realm of refinement: it doesn't generate decision-ready outputs or naturally accommodate multi-model integration. At the system level, an Industry 4.0 framework integrates IoT, AI, and cloud infrastructure to simulate and optimize greenhouse operations across climate, energy, and production scheduling [7]. It achieves meaningful efficiency gains but treats the greenhouse as an infrastructure optimization problem rather than a crop-level management challenge — disease intelligence and adaptive plant-health control are absent.

I. EXISTING SYSTEM

The baseline system treats the greenhouse primarily as a monitoring surface. It layers an AI advisory module on top of a real-time visualization interface, giving growers a window into environmental conditions as they evolve. The AI components serve as informational tools — a CNN-based model flags potential diseases from uploaded leaf images, and a data-driven recommendation engine suggests suitable crops based on soil and environmental inputs. These outputs provide genuinely useful context, but they don't influence anything. No control action is triggered, no resource is adjusted, no simulation runs in the background. The system tells you what it sees; it doesn't act on what it predicts.

Alongside the visualization layer, the system includes a CNN-based disease classifier and a crop recommendation module. The disease component takes a leaf image, runs it through standard preprocessing — resize, normalize — and outputs a predicted disease label with a confidence score. The crop recommendation module pulls from soil and environmental data to suggest a suitable crop. Both serve their purpose within the demo interface, but they operate as independent advisory tools rather than components of a connected control loop. There's no feedback between the disease detector and the actuators — a flagged pathogen doesn't trigger any corrective response. As a starting point, the baseline system does what it was designed to do: visualize greenhouse conditions, display actuator states, and show basic climate trends. That makes it a useful reference for comparison. What it isn't is a digital twin in any complete sense — there's no continuous model aligning simulated behavior with measured data, no simulation layer for testing control strategies safely, and no optimization that accounts for crop stage or disease risk. Those are exactly the gaps AgriTwin-GH addresses.

II. PROPOSED SYSTEM

The AgriTwin-GH system is built around three principles: data-driven, modular, and simulation-validated. It combines open-source environmental datasets, AI-based analysis, and a digital twin simulation layer to support intelligent greenhouse management without requiring a live sensor network. Each component is designed to work independently, integrate cleanly with the rest, and be extended or replaced without disrupting the system.

International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 14, Issue 4, April-2026, Available online at: www.ijaresm.com

A. Overall Architecture Diagram

AgriTwin-GH runs on a continuous data pipeline where environmental sensors and cameras feed temperature, humidity, soil moisture, light, and plant imagery into a unified preprocessing layer that cleans and structures everything before analysis begins. From there, image-based models handle growth stage classification and disease detection while time-series models forecast environmental conditions — giving the system two complementary views of greenhouse health. These predictions don't just describe the present; they project how growth, disease, and environment will evolve, feeding a decision layer that translates forecasts into concrete actuator adjustments for watering, ventilation, heating, and lighting. A digital twin ties it all together, running as a virtual greenhouse where control strategies are tested before reaching the real environment. The loop closes as predictions feed back into the models, forming a system that continuously adapts — with a web interface and 3D visualization keeping growers informed and in control at every step.

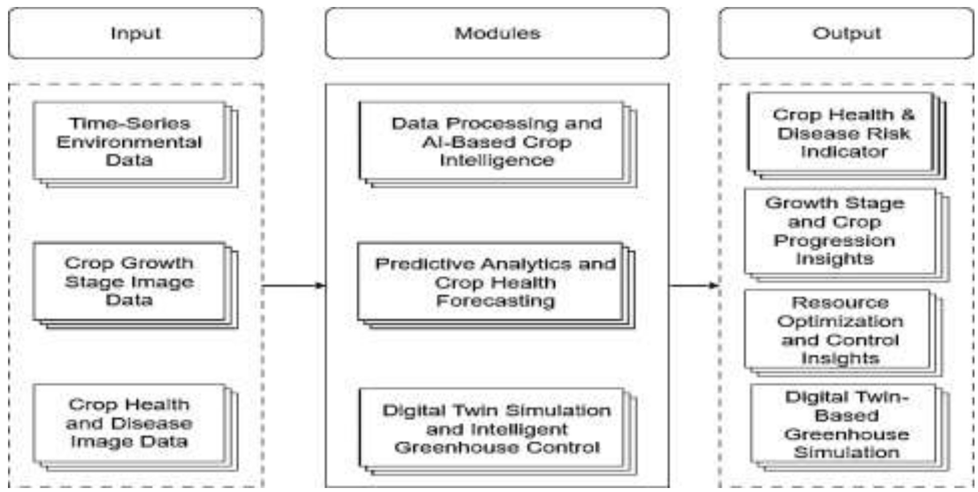

Fig. 1 High-Level Architecture Diagram of AgriTwin-GH

Fig. 1 illustrates the high-level architecture of the proposed AgriTwin-GH framework, depicting the flow of data from input acquisition to intelligent decision outputs. The architecture is structured into three primary layers: input data sources, core processing modules, and system outputs.

B. Data Processing and AI-Based Crop Intelligence

The Data Processing and AI-Based Crop Intelligence module turns raw greenhouse data into something the system can actually learn from. Environmental time-series feeds, external weather inputs, and plant images all flow through a shared pipeline — validated, cleaned, and standardized before anything else happens. Features are engineered to carry genuine agronomic meaning, and only then do the deep learning models for growth stage classification and disease detection take over. The result is a dependable stream of environmental and visual crop intelligence that the forecasting and control layers can build on confidently.

The pipeline begins with raw weather API data — messy, incomplete, and timestamp-inconsistent — which is integrated, cleaned, and aligned along a common time axis so that temperature, humidity, CO_2, and light readings can be analyzed together. From there, raw outdoor inputs are standardized and expanded into hourly records, then translated into meaningful indoor greenhouse estimates: temperature, humidity, and air velocity alongside agronomic indicators like dew point, Vapor Pressure Deficit (VPD) as a measure of plant transpiration demand, and a leaf wetness proxy that flags humidity conditions favorable to disease. The resulting structured dataset — covering indoor climate variables, CO_2, solar radiation, VPD, dew point, and leaf wetness — becomes the reliable foundation every AI module in the system builds on.

The system includes two image-based modules — one for growth stage classification, one for disease detection — both built on EfficientNet backbones with custom classification heads and softmax outputs that deliver confidence-weighted predictions rather than hard labels. The growth module uses EfficientNetB3 to classify tomato plants into six development stages — Seedling through Ripe Fruit — reading visual cues like leaf structure, flower presence, and fruit development. The disease module uses EfficientNetB0 to identify six leaf conditions, from Early Blight and Powdery Mildew to Healthy, catching visible symptoms early before they spread. Both follow the same two-stage transfer learning approach — training the classification head first, then fine-tuning upper backbone layers — and both use class-weighted loss functions to handle imbalanced data. The disease module applies heavier augmentation given the wider variation in lighting, orientation, and symptom severity it encounters in real greenhouse conditions. Together, they give the system a reliable visual understanding of both crop development and plant health without requiring manual inspection.

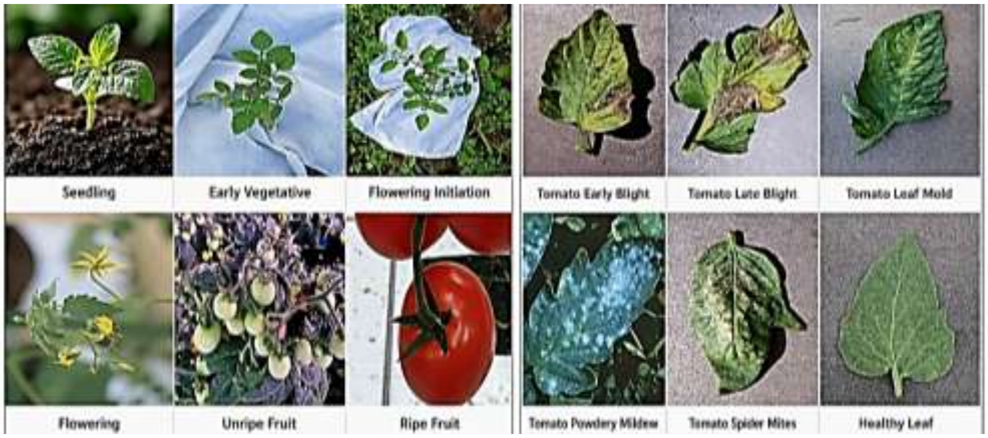

Fig. 2 Tomato Crop Growth Stages and Tomato Leaf Disease Dataset Samples

Fig. 2 illustrates sample input images representing different crop growth stages used for training and evaluating the growth stage classification model, and plant leaves showing various disease conditions used for training and evaluating the disease detection model.

C. Predictive Analytics and Crop Health Forecasting

The Predictive Analytics and Crop Health Forecasting module is built around three connected models: a weather forecaster, a growth progression model, and a disease progression model. The weather model looks ahead at how indoor temperature, humidity, and light will shift over the coming hours. The growth model builds on that by estimating how the plant itself will develop — which stage is next, and when the transition happens. The disease model rounds it out by projecting how active pathogen conditions are likely to evolve given recent environmental history. Together, these three perspectives — environment, growth, and disease — give the system a forward view that no individual sensor can provide on its own. Instead of waiting for something to go wrong, the module lets the system see it coming.

Weather Forecasting The module predicts external temperature, humidity, wind speed, and solar radiation over 24- and 48-hour windows — because actuator decisions should anticipate where conditions are heading, not just where they stand. It blends three complementary approaches: a Chronos-T5-small transformer for long-range temporal patterns, XGBoost for structured feature relationships, and stacked LSTMs for sequential trends. All three train on chronologically split historical data, enriched with engineered features — cyclical time encodings, lag features, rolling statistics, and interaction terms. Final predictions are produced by blending all three outputs using weights optimized per variable and horizon, with a Random Forest classifier mapping results to sky-condition categories for cleaner dashboard display.

Disease Progression Rather than simply flagging a pathogen's presence, the system tracks how disease develops, intensifies, or retreats — estimating current severity, projecting future states, and characterizing directional change. Each disease gets dedicated features and trend labels grounded in domain knowledge, all within a single unified framework. Data is cleaned, chronologically split, and fed first into gradient boosting models that establish a strong baseline for presence classification and severity regression. LSTM and GRU networks then add genuine temporal memory, with a shared encoder learning general disease rhythms and task-specific heads handling classification and regression independently. Final outputs blend both approaches using validation-optimized weights.

Growth Progression At every hourly step, the model outputs the plant's current growth stage, the next expected stage, estimated hours until transition, and transition probabilities at 24 and 48-hour horizons — letting growers adjust nutrients or plan harvests ahead of time rather than reacting after the fact. Training spans multiple independent growth cycles to ensure the model generalizes broadly. Random Forest and Gradient Boosting models serve as interpretable baselines, while the core system is a multi-output LSTM with a shared temporal encoder and separate heads for stage classification, transition likelihood, and time-to-transition regression. Post-processing rules ensure no predicted sequence contradicts known plant biology, and inference runs continuously on rolling sensor windows in deployment.

D. Digital Twin Simulation and Intelligent Greenhouse Control

This is where everything comes together. Environmental forecasts, crop growth data, disease signals, and live greenhouse state merge into a single control loop that acts continuously — not just monitors. At its core is a digital twin simulation where a virtual greenhouse evolves step by step, feeding its own outputs back into the next cycle. A Model Predictive Control (MPC) framework runs alongside it, calculating actuator settings — temperature, ventilation, irrigation, lighting — based on both current conditions and what's expected ahead. A FastAPI backend, web frontend, and 3D visualization surface the simulation state in real time, while monthly summaries support longer-term performance review.

The Model Predictive Control layer is the decision-making engine. At each time step, it takes the current greenhouse state — temperature, humidity, CO_2, soil moisture, light, VPD, leaf wetness — alongside disease risk, growth stage, and a 24-hour weather forecast, and computes a sequence of actuator actions that minimizes deviations from the crop's ideal conditions over a defined planning horizon. What makes this different from a simple rule-based controller is that the MPC reasons ahead: it considers how actions taken now will affect conditions later and respects operational constraints from the outset rather than treating them as violations to be corrected after the fact. State dynamics are modeled using an ARX (AutoRegressive with eXogenous inputs) framework, which predicts how each greenhouse variable responds to actuator commands and external weather disturbances. The general state transition equation is:

$$x_i[k+1] = \alpha_i x_i[k] + \Sigma\, \beta_{ij} u_j[k] + \Sigma\, \gamma_{im} d_m[k] + \varepsilon_i \qquad (1)$$

Where, $x_i[k]$ represents the i-th state variable at time step k, $u_j[k]$ represents the j-th control input, $d_m[k]$ represents external disturbances such as outdoor weather, α_i is the state persistence coefficient, β_{ij} represents actuator influence, γ_{im} represents disturbance influence, and ε_i captures modeling uncertainty.

$$T[k+1] = \alpha_T T[k] + \gamma_{ext}(T_{ext}[k] - T[k]) + \gamma_{sol} S[k] + \beta_{heat} u_{heater} + \beta_{fan} u_{fan} + \beta_{vent} u_{vent} \quad (2)$$

Where, T[k] is indoor temperature, Text[k] is outdoor temperature, S[k] is solar radiation, and the actuator terms represent heating and cooling contributions. Comparable ARX formulations govern humidity, CO_2, and light. Soil moisture follows a water balance equation. The MPC solves a finite-horizon cost minimization problem:

$$\min J(u) = \Sigma_{\{k=0\}}^{\{N-1\}} \ell\left(x_k, u_k, u_{\{k-1\}}\right) + V_{f(x_N)} \qquad (3)$$

Where, N is the prediction horizon, $\ell(\cdot)$ is the stage-wise running cost, and is the terminal cost. This formulation optimizes both immediate performance and longer-term system behavior simultaneously.

The Digital Twin simulation loop is the execution engine that keeps the virtual greenhouse alive step by step. At each cycle, it takes the current greenhouse state, applies the MPC's actuator commands and the prevailing weather, advances the physical model by one time step, and feeds the result back as the starting point for the next cycle — a true closed loop. The schedule is multi-rate: the simulation advances every 5 minutes, MPC re-optimizes every 15 minutes, and image-based observations refresh every 30 minutes. This balances computational cost against control responsiveness without breaking the continuous loop. The greenhouse transition itself is based on the same ARX model used by the MPC module. Among individual state equations, temperature is especially important because it directly drives crop stress, humidity coupling, and disease risk. The temperature update is:

$$T[k+1] = \alpha_T T[k] + \gamma_{ext}(T_{ext}[k] - T[k]) + \gamma_{sol} S[k] + \beta_{heat} u_{heater} + \beta_{fan} u_{fan} + \beta_{vent} u_{vent} \quad (4)$$

Where, T[k] is current indoor temperature, $T_{ext}[k]$ is external temperature, S[k] is solar radiation, u_{heater} is heater output, u_{fan} is fan speed, and u_{vent} is vent opening. α_T thermal persistence, γ_{ext} models heat exchange with the outside environment, γ_{sol} models solar gain, and the β terms represent actuator effects. Positive heater gain raises temperature, while fan and vent terms reduce temperature. Likewise for other factors computation is done. To keep simulations biologically realistic over long runs, the loop includes automatic growth-stage progression: when a stage's configured duration elapses, the loop advances to the next stage and resets the stage-hour counter automatically — allowing a full crop cycle from seedling to ripe to unfold without manual intervention. Emergency-triggered MPC re-solves add a reactive safety layer on top of the scheduled 15-minute cadence: if any step produces a state that crosses an emergency threshold, the optimizer is invoked immediately rather than waiting for the next scheduled interval.

The frontend, API, and monthly snapshot system are what users actually see and interact with. The frontend covers environmental monitoring, crop growth intelligence, disease risk, historical trends, and manual controls — all without requiring technical expertise. FastAPI handles communication between layers, delivering simulation updates to the UI, validating user inputs, and bridging the 3D visualization over a live connection. Monthly snapshots compress highfrequency telemetry into PostgreSQL summaries — energy and water usage, actuator behavior, disease incidents, growth progression — making month-over-month comparison practical for longer-term review. The Unity-based 3D greenhouse model gives the digital twin a physical presence. Vents, fans, irrigation systems, lighting rigs, and crop plants are all modeled as interactive objects that respond dynamically to live actuator commands. Running in WebGL with no installation required, the scene updates in real time as simulation state changes — crop plants shift appearance as growth stages advance, lighting reflects time of day, and color cues communicate system status at a glance, so growers can read complex conditions without parsing raw diagnostic data.

RESULT ANALYSIS

E. Results of Crop Growth Stage Classification Model and Plant Disease Detection Model

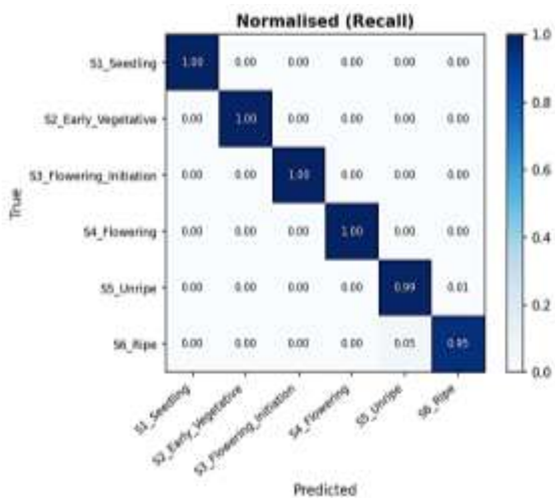

Fig. 3 Confusion Matrix of Tomato Growth Stage Classification Model

Fig. 3 shows the confusion matrix for the growth stage classification model. Classification performance is strong across most stages, with the model correctly distinguishing the visually distinct seedling and ripe phases. Minor confusion occurs between adjacent mid-growth stages, which is expected given how gradually leaf morphology transitions.

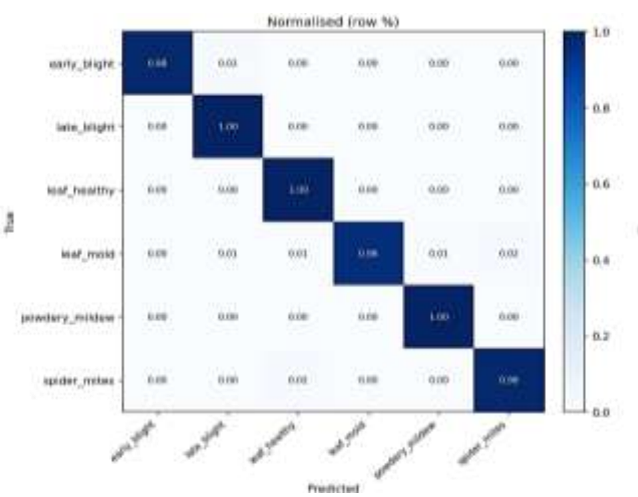

Fig. 4 Confusion Matrix of Tomato Disease Detection Model

Fig. 4 shows the confusion matrix for the plant disease detection model across six disease categories. The model correctly identifies the majority of disease classes, with healthy leaves and late blight achieving particularly clear separation.

F. Results of Weather Forecasting Model

Table I compares the proposed hybrid weather forecasting model to existing single-model approaches. The ensemble consistently outperforms individual models on both 24-hour and 48-hour horizons across evaluation metrics.

Table I. Performance Comparison of Weather Forecasting Models

Parameter	Proposed Work
Model Type	Chronos + GRU-based Deep Learning Model & Random Forest Classification Model
Forecast Horizon	2 Days
Variables Predicted	Temperature, Wind Speed, Humidity, Solar Radiation
R^2 (Solar Radiation)	0.886
R^2 (Temperature)	0.943
R^2 (Wind Speed)	0.885
R^2 (Humidity)	0.629

G. Results of Disease Progression Model and Crop Growth Progression Model

Table II presents regression performance for disease severity prediction across all disease outputs in the LSTM model. R^2 values near 1.0 and low RMSE and MAE figures indicate strong predictive accuracy across forecast horizons.

Table II. Performance Metrics for Disease Severity Prediction

Metric	Value
Mean Absolute Error (MAE)	3.297
Root Mean Square Error (RMSE)	7.143

Table III presents time-to-next-stage prediction performance. The regression metrics confirm that the model provides useful hour-level accuracy in predicting how long a plant will remain in its current growth stage.

Table III. Performance Metrics for Time-to-Next-Stage Prediction

Metric	Value
Mean Absolute Error (MAE)	6.58
Root Mean Square Error (RMSE)	8.45

H. System Interface and 3D Greenhouse Visualization

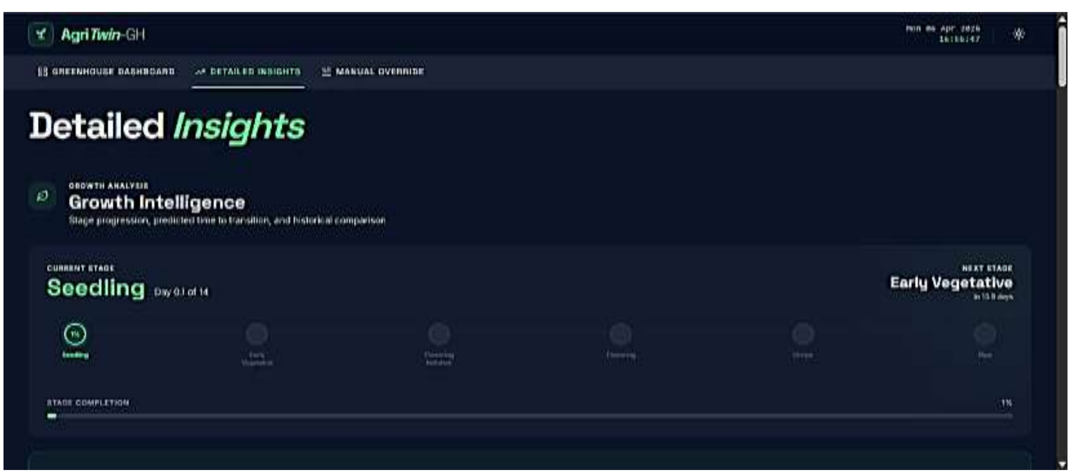

Fig. 5 Detailed Insights Interface

Fig. 5 shows detailed insights interface, which displays all model outputs in a compact, readable format.

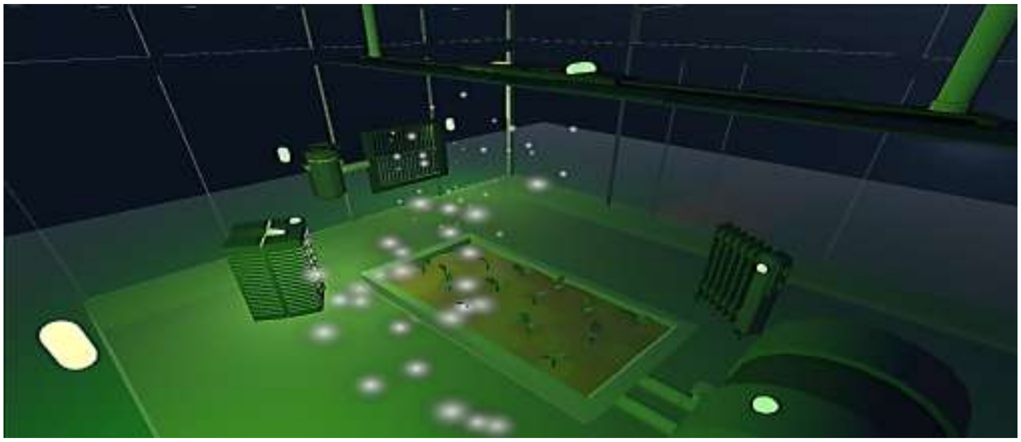

Fig. 6 Top-Left Perspective View of Greenhouse in Active State

Fig. 6 shows a top-left perspective view of the greenhouse interior in active state, providing a broader spatial overview of how actuators and crops are distributed across the structure.

I. Comparative Performance Evaluation

Table IV evaluates existing greenhouse management systems against AgriTwin-GH across dimensions including predictive capability, disease intelligence, digital twin integration, and sensor independence.

Table IV. Comparative Analysis of Existing Systems and Proposed AgriTwin-GH Framework

Criteria	IoT-Based Systems	Digital Twin Systems	AgriTwin-GH (Proposed)
Monitoring Capability	Real-time sensor-based monitoring	Real-time + virtual monitoring	Multi-source intelligent monitoring
Disease Detection	Not supported	Limited	Integrated AI-based detection
Control Strategy	Semi-automated	Predictive (MPC-based)	Predictive + MPC & Simulation-based control
AI/ML Integration	Minimal	Moderate	Fully integrated AI framework
Decision Support	Limited	High	Intelligent decision-support system

CONCLUSION

AgriTwin-GH is a software-based digital twin framework that moves tomato greenhouse management from reactive to proactive. It brings together environmental data processing, EfficientNet-based growth stage and disease classification, time-series forecasting for weather and crop dynamics, and a digital twin simulation coupled with Model Predictive Control. Rather than waiting for something to go wrong, the system continuously reads environmental and crop health signals, projects how conditions will evolve, and tests control strategies against a virtual greenhouse before applying them to a real one. A core contribution is the sensor-independent design — by treating processed datasets as virtual sensors, the framework operates without physical sensing infrastructure, making it practical in cost-sensitive settings. Across all evaluation tasks, AgriTwin-GH demonstrates strong predictive accuracy and stable simulation behavior, providing a scalable foundation for AI-driven greenhouse management.

Looking ahead, several directions stand out. Integrating physical IoT sensors would enable true synchronization between the virtual and real greenhouse. Expanding models to cover crops beyond tomatoes would broaden applicability, while reinforcement learning could replace manual reconfiguration with control strategies that improve through operational experience. Richer environmental inputs — regional forecasts, satellite indicators, local microclimate data — would sharpen the system's ability to anticipate external disturbances. Edge-based AI deployment would reduce latency in low-connectivity settings, and field evaluations measuring yield, resource efficiency, and carbon footprint would validate the system's contributions to sustainable precision agriculture at scale.

REFERENCES

1. A. Benyezza, A. Bouhedda, and S. Rebouh, "Digital twins in greenhouse horticulture: A review," Computers and Electronics in Agriculture, 2022.
2. A. F. Subahi and K. E. Bouazza, "An intelligent IoT-based system design for controlling and monitoring greenhouse temperature," IEEE Access, vol. 10, pp. 1–12, 2022.
3. A. Kaur and R. Kaur, "AI-enabled energy management for sustainable smart greenhouses: An integrated review," Journal of Cleaner Produelbasiction, vol. 382, pp. 1–15, 2024.
4. A. Oad, S. S. Abbas, A. Zafar, et al., "Plant leaf disease detection using ensemble learning and explainable AI," IEEE Access, vol. 12, pp. 156038–156052, 2024.
5. A. Soussi, E. Zero, C. D. C. Herrera, et al., "Integrating digital twins and model predictive control for sustainable greenhouse management in smart agriculture," IEEE Transactions on AgriFood Electronics, vol. 1, no. 1, pp. 1–17, 2025.
6. B. V. Nikitha, N. K. S. Keerthan, P. Praneeth, and T. Amrita, "Leaf disease detection and classification," Procedia Computer Science, vol. 218, pp. 291–300, 2023.
7. D. A. Howard, Z. Ma, C. Veje, A. Clausen, J. M. Aaslyng, and B. N. Jørgensen, "Greenhouse industry 4.0 – digital twin technology for commercial greenhouses," Energy Informatics, vol. 4, no. 1, pp. 1–20, 2021.
8. D. K. Lee, B. K. Kang, Y. J. Kim, T. G. An, and Y. T. Chae, "A digital twin-based framework for precision tomato cultivation: Integrating XGBoost-based prediction with model predictive control," SSRN Electronic Journal, 2024.
9. D. S. Joseph, P. M. Pawar, and K. Chakradeo, "Real-time plant disease dataset development and detection of plant disease using deep learning," IEEE Access, vol. 11, pp. 1–15, 2023.
10. E. Elbasi, N. Mostafa, C. Zaki, Z. AlArnaout, A. E. Topcu, and L. Saker, "Optimizing agricultural data analysis techniques through AI-powered decision-making processes," Applied Sciences, vol. 14, no. 8018, 2024.
11. J. Silva, P. Costa, and L. Gomes, "A predictive greenhouse digital twin for controlled environment agriculture," Proc. ICINCO, 2025.
12. L. Caroprese, M. Pierantozzi, C. Lops, and S. Montelpare, "DL2F: A deep learning model for the local forecasting of renewable sources," Computers & Industrial Engineering, vol. 187, 109785, 2024.
13. M. Ghiasi, Z. Wang, M. Mehrandezh, and A. Mohammadi, "Optimal strategy for energy-efficient management of greenhouse climate control," IEEE Open Access Journal of Power and Energy, vol. 12, pp. 1–14, 2025.
14. M. R. Islam, et al., "Towards digital twin-enabled greenhouse management: Enhancing climate resilience and sustainability through machine learning-driven controlled environment agriculture," Sustainable Energy Technologies and Assessments, 2026.
15. M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," Proc. ICML, 2019.
16. M. Umar, S. Altaf, S. Ahmad, H. Mahmoud, A. S. N. Mohamed, and R. Ayub, "Precision agriculture through deep learning: Tomato plant multiple diseases recognition with CNN and improved YOLOv7," IEEE Access, vol. 12, pp. 49167–49180, 2024.
17. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, 1997.
18. S. Kanakala and S. Ningappa, "Detection and classification of diseases in multi-crop leaves using LSTM and CNN models," Journal of Innovative Image Processing, vol. 7, no. 1, pp. 161–181, 2025.
19. S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, 2016.
20. S. R. Luna, M. P. García, and J. R. Torres, "Digital twin technology in greenhouse-based precision agriculture," Computers and Electronics in Agriculture, vol. 209, pp. 107–119, 2023.
21. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," Proc. ACM SIGKDD, 2016.
22. T. He, M. Li, and D. Jin, "Deep learning-based time series prediction for precision field crop protection," Frontiers in Plant Science, vol. 16, 1575796, 2025.
23. Y. Huo, Y. Liu, P. He, L. Hu, W. Gao, and L. Gu, "Identifying tomato growth stages in protected agriculture with StyleGAN3–synthetic images and Vision Transformer," Agriculture, vol. 15, no. 120, 2025.
24. Y. Zhang, X. Wang, et al., "Data-driven robust model predictive control for greenhouse temperature control and energy utilisation assessment," Applied Energy, 2023.
25. Z. Zhang, J. Li, and Y. Wang, "Digital twin-driven smart greenhouse management using deep learning and IoT integration," Computers and Electronics in Agriculture, 2024.