Harvesta: Sustainable Greenhouse Management through Smart Technology Integration

Kumaravithana D.B
Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21342608@my.sliit.lk

Padukka P.D.M.D
Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21277504@my.sliit.lk

Jayawardena R.R.A.N.N
Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21894510@my.sliit.lk

Sandeepani A.W.S
Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21173004@my.sliit.lk

Sanjeevi Chandrasiri
Dep. Of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
sanji.c@sliit.lk

Shalini Rupasinghe
Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
shalini.r@sliit.lk

***Abstract*— The cultivation process is highly dependent on several factors, such as ripeness, plant diseases, resource management, and pest infestations. Previously, evaluating these factors was highly dependent on visual observation, a process that has often been characterized by inefficiencies, susceptibility to mistakes, and overall ineffectiveness. Advances in machine learning and deep learning technologies open a new door to building and integrating data-based autonomous systems to improve accuracy in agriculture. This research focuses on building and integrating an intelligent system specifically for tomato cultivation with features for forecasting harvesting ripeness, assessing and predicting plant disease severity, and adapting fertilization and irrigation processes in real-time according to soil and environmental factors. Additionally, significant breakthroughs in autonomous pest detection and control have enabled pest species identification and execution of targeted interventions. All in all, all these breakthroughs are crucial to enhancing tomato quality, lowering environmental footprint, and enhancing economic sustainability. Practical applicability in real-world agricultural environments is guaranteed by portability and ease of use. Empirical findings show that incorporating machine learning in tomato cultivation results in increased yields, reduced wastage, and a reduced dependency on toxic chemical applications. This paper demonstrates how automation, through machine learning, can revolutionize conventional methods in tomato farming into efficient and sustainable agriculture.**

Keywords*— *Crop Management, Machine Learning, Precision Agriculture, Smart Farming, Tomato Farming.

I. INTRODUCTION

Tomatoes are one of the most widely cultivated fruits in the world and contribute greatly to food production and economic viability. However, traditional methods for cultivating tomatoes are highly dependent on subjective judgments in numerous activities, ranging from determining ripeness to detecting diseases, irrigation and fertilizer management, and pest control. These traditional methods not only require a large amount of resources but are also subjective to a certain extent, making them prone to mistakes. Such limitations in methodologies can lead to ill-informed harvesting decisions, substantial crop losses, resource inefficiency, and unnecessary use of chemical pesticides. Recent progress in computational modeling and imaging technologies has made it possible to create data-based strategies for enhancing cultivation techniques for tomatoes and enhancing decision-making and agricultural output. Ripeness is a critical requirement for deciding when to harvest tomatoes.

Conventional approaches using imaging technologies are often plagued with inaccuracies and inconsistencies, which in turn negatively influence fruit quality and market price. Image processing algorithms like YOLOv8 and Faster R-CNN present a non-destructive way to accurately measure tomatoes' ripeness, thus providing farmers with objectivity and reliability. Yields and overall quality of tomatoes are highly prone to numerous diseases. Manual disease detection in tomatoes is time-consuming and inapplicable [1]. Automated methods, specifically Convolutional Neural Networks (CNNs), have proven effective in detecting and quantifying diseases and enabling timely interventions and specific treatments. Efficient irrigation and fertilization strategies are imperative to improve agricultural output with reduced environmental pressures. Standard resource allocation techniques normally run according to pre-specified schedules and result in ineffective use of water and nutrients. The incorporation of environmental analytics in plant growth monitoring enables a system with the ability to adjust irrigation and fertilization levels dynamically and thus reduce wastage and increase agricultural output. Additionally, pest control in tomato farming is normally achieved through the use of chemical pesticides, which are environmental pollutants and contribute to pest resistance. Utilizing sophisticated pest detecting technologies in the form of image-based models like Xception enables earlier pest-related issues to be detected and thus pest control to be made more effective and sustainable. This research is aimed at developing a sophisticated system for managing tomato cultivation using image processing algorithms, algorithmic methods, and environmental sensing technologies to enable decision-making in ripeness determination, disease detection, and irrigation, fertilization, and pest control management. Using state-of-the-art analytical methods, this research seeks to advance precision agriculture processes in tomato cultivation to enhance sustainability, reduce wastage in resources, and enhance both agricultural quantity and quality.

II. LITERATURE REVIEW

Agriculture faces numerous challenges, including plant diseases, pest infestations, inefficient resource use, and improper harvest timing. Emerging technologies such as deep learning, predictive analytics, and IoT-based systems have been leveraged to enhance efficiency in disease detection, pest identification, irrigation optimization, and ripeness assessment. Despite advancements, challenges persist in improving model accuracy, integrating real-time monitoring, and ensuring adaptability in diverse agricultural settings.

Plant disease identification and severity prediction are critical for timely intervention and crop protection. Deep learning models, particularly Convolutional Neural Networks (CNNs), have shown success in classifying plant diseases and predicting their severity [2]. However, distinguishing between healthy, mild, moderate, and severe infection stages remains challenging. To address this, labeled datasets and data augmentation techniques have been employed to improve model accuracy and reduce misclassification errors. A system integrating disease detection with severity classification would significantly improve precision agriculture, allowing farmers to implement targeted control measures.

Pest infestations are another major concern, and deep learning models such as Xception have proven effective in pest identification and classification. Image-based pest detection offers real-time monitoring, improved accuracy, and reduced reliance on manual scouting. The integration of mobile-based pest detection systems with automated recommendation tools ensures sustainable pest control. However, distinguishing visually similar pest species and adapting models to different environments remain challenges that require further research.

Fertilization and irrigation optimization are crucial for resource-efficient agriculture. Conventional methods overuse water and fertilizers, leading to environmental degradation and inefficiency. IoT-enabled smart irrigation systems and automated nutrient monitoring improve water-use efficiency and prevent over-fertilization, ensuring crops receive the right amount of nutrients. However, sensor calibration, cost, and scalability remain challenges that need to be addressed.

Ripeness detection and harvest prediction involve colour-based classification, environmental monitoring, and growth rate analysis to determine the best harvesting time [3]. By analyzing ripeness percentage differences over time, predictive models help schedule optimal harvesting while ensuring high yield quality. An integrated system that combines vision-based classification, environmental condition monitoring, and predictive modelling would significantly enhance precision harvesting and reduce post-harvest losses [4].

The integration of advanced technologies such as IoT, predictive analytics, and automated monitoring systems in agriculture enhances efficiency, sustainability, and precision in farming practices. However, to achieve widespread adoption, further research is required to refine existing models, improve scalability, and address challenges related to environmental adaptability, ensuring these solutions can be effectively implemented across diverse agricultural settings.

III. METHODOLOGY

Through the use of machine learning models, this research provides a comprehensive approach to optimize agricultural practices for disease detection and pest identification, while also providing environmental monitoring through irrigation and fertilization. The research incorporates a range of deep learning models and datasets obtained from local farms, agricultural research institutions, and publicly available repositories to tackle specific challenges in precision agriculture. The classification of tomato diseases, ripeness phases, pest detection, and fertilization and irrigation methods are all accomplished using the models' capabilities. These models take into account various environmental factors and plant conditions. This research utilizes data preprocessing, transfer learning, model development, and evaluation to enhance agricultural productivity by reducing resource wastage and improving crop management.

A. Disease Detection and Severity Prediction with Management Suggestions

The present study employed a dataset of images of tomato plants, which were collected from local greenhouses and complemented with publicly available databases. The images were carefully analyzed and classified based on disease type and severity, including four specific categories: healthy, mild, moderate, and severe, with a pathologist's oversight to ensure accuracy and consistency. This classification system enabled the model to make accurate predictions and provide specific recommendations for the efficient control of the plants. Public databases, especially those related to common tomato diseases in Sri Lanka, contributed high-quality images that increased the accuracy and diversity of the dataset.

In order to promote uniformity in the input data and to improve its amenability to the model, all the images were resized to a standardized resolution of 128×128 pixels before being converted to tensor format for being incorporated in the PyTorch framework. Furthermore, in order to enhance the model's robustness and flexibility in handling challenges related to limited data availability, a range of data augmentation methods were used. These methods included random rotations, horizontal flipping, and changes in brightness levels. The overall aim of these augmentations was to simulate real-world environments, thus improving the model's capability of handling varied input conditions and preventing the possibility of overfitting [5].

The EfficientNet-B0 architecture was used as it is lightweight and performs exceptionally well in the execution of image classification operations. The model was pre-trained using the ImageNet dataset, followed by subsequent fine-tuning using a specifically designed dataset optimized for tomato disease [6].To meet the demands of multi-class classification, the output layer was adapted to correctly predict the specific disease type and severity level present. The Cross-Entropy Loss function was used during training as the function is ideal for the resolution of problems related to multi-class classification operations.

Learning efficiency and acceleration of convergence was improved through the use of the Adam optimizer. The dataset was divided into three sections: 70% for training, 20% for validation, and 10% for testing. The performance of the model during the training process was carefully monitored using the validation dataset to detect and counteract any possibility of overfitting and improve the generalization ability for new instances. The system has two major functions: the diagnosis of plant health and treatment recommendations based on disease severity. When analyzing an image, the model categorizes the plant's health into one of four categories: healthy, mild, moderate, or severe. Healthy plants do not require immediate action; however, they need consistent observation and optimal environmental conditions. Plants showing a mild infection require minimal interventions, including pruning and environmental adjustments. For plants showing a moderate infection, stricter measures must be implemented, including isolation and the use of insecticides. On the other hand, plants with severe infections require extreme measures, including complete removal or the application of pesticides to curb further spread of the virus. This methodical approach to accurate disease prediction and customized treatment protocols is designed to improve plant disease management and allow for timely remedial measures to reduce crop losses.

B. Automated Ripeness Sensing with Environmental and Harvest Prediction System

The data contains images captured in greenhouses under controlled lighting. The images are labeled with bounding boxes for easy detection and classification of objects. Pre-processing involved the elimination of background objects, contrast adjustment, and noise reduction for better classification accuracy. Data augmentation methods involving brightness and contrast adjustments and distortion were applied to render the model stronger and enable better tomato classification under various farm conditions.

The system employs two object-recognition models based on deep learning, specifically YOLOv8 and Faster R-CNN. YOLOv8 detects various stages of ripeness of tomatoes in real time and is therefore appropriate for greenhouse monitoring, whereas Faster R-CNN predicts the best time and conditions to harvest when tomatoes are expected to be ripe. The two models together enable accurate monitoring of ripeness stages, thereby enhancing the harvesting decision. YOLOv8's speed and integration ability enable quick and efficient classification, and Faster R-CNN improves the precision of bounding box prediction and post-harvest ripeness assessment. Integrating these two models creates a holistic system for tracking tomato growth and its readiness for harvest.

Data detection models were trained with supervised learning on a labeled dataset for the correct classification of fruit ripeness. It was trained for multiple epochs to improve feature detection. The model architecture was optimized using ResNet-50 to improve efficiency and accuracy of ripeness classification. Contrast stretching and data augmentation were applied to make the model more robust to environmental condition variations. These improvements enable the system to effectively classify the ripeness of tomatoes, thereby facilitating optimum harvesting decisions.

The crop-harvest prediction system has been developed to determine the optimum time for harvest based on analyzing growth and maturity rates. The system classifies crops into half-ripe, ripe, and unripe as the three maturity levels, thereby enhancing decision-making. These classifications enable farmers to refine harvesting strategies, ultimately leading to enhanced quality and yield outcomes. Moreover, the system constantly tracks the stages of ripening and identifies zones where environmental conditions need to be regulated in order to reach the optimum ripeness. With the tracking of the ripening distribution, it provides accurate recommendations for temperature and lighting control. Temperature control is adjusted dynamically based on the percentages of ripening, thereby providing a suitable environment for uniform maturity. Light intensity is also controlled based on the distribution of ripening, and farmers manually or automatically adjust it.

The system also predicts the growth rate of the crops and provides timely notification of when it is best to harvest. The system tracks various growth stages, allowing farmers to better estimate the ripeness of the crops. The system recommends that farmers begin harvesting when 70% of the crops are ripe, reducing losses due to over-ripening. The system improves farming practice through the use of environmental recommendations, predictions, and ripeness assessments. This results in less monitoring labor and post-harvest losses, and enhanced crop quality and revenues.

C. Pest Identification and Management System

The pest identification model was created using a range of images of agricultural pests, with special focus on tomato pests. Images were sourced from local greenhouses, publicly available databases, and online collections of pest specimens to provide representation across various species in various lighting and environmental conditions. Pre-processing steps included resizing the images to sizes of 224 × 224 pixels, scaling the pixel values, and the application of data augmentation techniques including rotation, scaling, and perspective changes to facilitate generalization and reduce the overfitting risk. These techniques greatly improved the accuracy and reliability of the model for pest detection [7].

In pest identification, the Xception model was used as a feature extractor with a pre-trained variant built on the ImageNet dataset. The terminal layer was removed, and the features extracted were further processed through a custom-designed classifier, comprising fully connected layers, batch normalization, and dropout regularization to alleviate the issues of overfitting [8]. The final SoftMax layer produced probability distributions over different pest classes. For the training process, the Adam optimizer was used, leveraging its adaptive learning rate, while categorical cross-entropy was used as the loss function, given the nature of the multi-class classification problem. The training process was carried out over 50 epochs, using mini batches of 8 samples for training and 4 samples for validation. Techniques like early stopping and checkpointing were employed to save the model that gave the best performance [9].

The system enables instant detection of pests and provides management suggestions. Using a smart phone, agricultural farmers can take pictures of infested crops, which are then pre-processed and processed by a trained model. The SoftMax function is used for the classification of pests, mapping it to a pre-defined pest dictionary. After successful detection, pest management techniques are retrieved from a structured dataset stored in an Excel sheet and processed using Pandas. The system then provides instant control suggestions that include organic, chemical, and preventive measures.

D. Irrigation and Fertilization System

The data set utilized in this research relates to influential variables affecting environmental conditions and soil health, which are of vital importance for the application of precision irrigation and fertilization methods within the scope of precision agriculture. The data was gathered using advanced agricultural sensors and real-time monitoring systems that measure key parameters. The plant maturity and soil moisture levels were central factors in determining irrigation settings, while the levels of nitrogen, phosphorus, and potassium in the soil, combined with the fertilizer type used and the development stage of the crops, determined the fertilization methods. To ensure the validity of the data, a number of preprocessing operations were performed, including handling missing values (using mean or median for continuous and mode for categorical variables), removing duplicate records, and the application of anomaly detection algorithms like Isolation Forest and One-Class SVM. Categorical variables were represented using either one-hot encoding or label encoding methods, while numerical data were normalized using Min-Max scaling or Z-score normalization. The data set was then split into training and testing subsets [10].

To improve the accuracy of predictions, an ensemble learning methodology was employed. This was facilitated by the RandomForestRegressor, which enabled the training of an ensemble of decision trees, thereby reducing overfitting and improving the reliability of our irrigation models. Additionally, gradient boosting helped improve predictive accuracy by targeting residual differences with each subsequent iteration. To ensure reliability at the lowest computational cost, a VotingRegressor was utilized to combine the predictions of both the RandomForestRegressor and the XGBRegressor. In fertilization, the LGBMRegressor proved to be better performing compared to other models in dealing with high-dimensional features, while the CatBoostRegressor guaranteed precision through its efficient handling of categorical features with lower preprocessing demands [11]. The performance of the models was evaluated using a variety of metrics. The Mean Absolute Error (MAE) and Mean Squared Error (MSE) were quantitative measures of overall predictive accuracy, with MSE weighing more heavily on more substantial differences. The square root of the MSE was used to measure the model's ability in dealing with large variations. R-squared (R^2) was an indicator of the proportion of variance in the dataset explained by the model [12] [13].

IV. RESULTS & DISCUSSION

A comparative analysis was conducted on the performance of tomato ripeness classification models, i.e., YOLOv8 and Faster R-CNN, for assessing their performance in detecting and classifying tomatoes at different levels of ripeness. Their performance was assessed on key parameters like accuracy, recall, mean average precision (mAP), and Intersection over Union (IoU). The findings reveal that YOLOv8 outperforms Faster R-CNN when it comes to real-time detection, making it more appropriate to be integrated into autonomous agricultural systems. On the other hand, Faster R-CNN showed better localization accuracy and precision owing to its enhanced bounding box detection approach, but at the cost of higher computational time.

Visual observation was carried out in order to contrast two models' bounding box predictions. YOLOv8 demonstrated a clear difference between unripe, half-ripe, and ripe tomatoes with bounding boxes labeled accordingly. Low accuracy of detection as a result of occlusion and overlapping tomatoes was observed in certain cases, however. Faster R-CNN, although more precise at localization, demonstrated tighter bounding boxes and better aligned object boundary. This is supported by the higher IoU scores that were noted for Faster R-CNN.

Figure 1 presents the bounding box outputs for tomato ripeness classification, including confidence scores assigned by each model. These confidence scores reflect the reliability of each classification, helping to distinguish between correctly and incorrectly classified tomatoes. YOLOv8 exhibited high confidence in well-lit conditions but struggled in areas with shadow or occlusion. Conversely, Faster R-CNN maintained stable confidence levels but occasionally misclassified tomatoes at transitional ripeness stages.

Fig. 1. Bounding Box Outputs with Confidence Scores for Tomato Ripeness Classification.

The system also examined the percentage distribution of tomatoes at various stages of ripeness. Based on predefined thresholds, it dynamically regulated environmental factors like temperature and light intensity to enhance the ripening process without allowing overripening. When the percentage of ripe tomatoes crossed a critical point, the system determined the ideal harvesting time, and the crop was harvested when it was most ripe.

The result shows that as the number of ripe tomatoes increases, the system adjusted its recommendations to give the best time to harvest. While it worked fine, there were some problems noted. The change in lighting affected the ability of YOLOv8 to classify tomatoes accurately, specifically in identifying half-ripe and unripe tomatoes. Moreover, when the tomatoes were closely packed, it made detection less precise. Faster R-CNN performs better with obstructed objects, but it is more time-consuming, thus less appropriate for real-time applications. Some future enhancements can be the utilization of multi-spectral imaging for enhanced classification of ripeness and the implementation of hardware acceleration techniques to enable Faster R-CNN to perform faster.

The irrigation system was improved with ensemble learning techniques applied using RandomForestRegressor, XGBRegressor, and VotingRegressor. The combination of these three models led to a decrease in error predictions, thus ensuring a high degree of accuracy in irrigation suggestion. Overfitting was tackled by RandomForestRegressor, and XGBRegressor helped in enhancing accuracy by using gradient boosting techniques, and VotingRegressor ensured general stability. A performance comparison using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R^2 Score proved to be effective, as illustrated in Table I.

A similar ensemble learning strategy was utilized in constructing a fertilization recommendation system using LGBMRegressor, CatBoostRegressor, and Voting Regressor. The LGBMRegressor was very effective in dealing with huge datasets with low computational complexity and high processing rates. The CatBoostRegressor was best suited to deal with categorical inputs with low preprocessing needs. The VotingRegressor integration helped to increase the reliability of predictions, thus ensuring effective fertilization recommendation. The performance metrics such as MSE, MAE, and R^2 Score confirmed the effectiveness of this approach, and the comparison is also presented in Table I, highlighting the efficiency of both systems.

TABLE I

Performance Metrics for Irrigation and Fertilization Prediction

Metric	Irrigation Prediction	Fertilization Prediction
MSE	0.001233	14.91821
MAE	0.0268	3.1703
R^2	0.87642	0.88868

The proposed pest identification system using deep learning methods has a high classification rate for different categories. The precision, recall, and F1-score tested are in favor of the performance of the model in separating different classes from one another. The overall performance results are tabulated into the classification report in Table 1. The model had a total accuracy of 97%, thus validated in the actual-world way of identification. The macro average precision and recall were 0.97 and 0.96, respectively, and the weighted averages for both were similarly 0.97. From these data, it is possible to make the reasonable conclusion that the model shall perform equally in its different classes. The higher precision values signal near-perfect classification in certain sections and the lower ones point towards some likely misclassifications in certain classes. The recall values vociferously announce good detection capability, actively suppressing false negatives for all the classes. The performance of the proposed model over existing detection models for similar conditions is either equal or superior. High recall and precision values are indicative of classifying the varying instances into their respective classes with minimal misclassifications.

TABLE II

Pest Classification performance summary

Metric	Score
Accuracy	0.97
Macro Avg precision	0.97
Macro Avg Recall	0.96
Macro Avg F1-Score	0.96
Weighted Avg Precision	0.97
Weighted Avg Recall	0.97
Weighted Avg F1-Score	0.97

Results have shown that the designed system can proficiently categorize the different kinds of categories and recommend the control for them. With extreme accuracy, the model might significantly contribute to practical applications by reducing the overdependence on human detection, thereby attaining control measures in a timely manner and with great accuracy. Nevertheless, with some more refinement to the model, and with further training using different buckets of data, classification efficiency could still be improved, when it comes to examples carrying a value of precision that is less.

The Disease Identification Model performed at an accuracy of 81.94%, whereas the Severity Level Prediction Model performed slightly better at 83.2%. The results thus demonstrate that severity prediction of plant disease is a comparatively simpler task than the identification of the actual disease. The Severity Level Prediction Model predicted the disease severity as Mild, Moderate, and Severe with high accuracy, while the Disease Identification Model struggled due to the complexity of differentiating between various types of diseases. When we examined the performance measures, we could observe that both models performed well, yet it was hard to distinguish between Moderate and Severe cases of the disease. This is because the disease progression is fine-grained in nature and that visual features for these levels of severity might overlap, thus perplexing the model. The performance of the two models is consolidated in the table below, which also highlights key metrics like Accuracy, Precision, Recall, and F1 Score. Severity Level Prediction Model reported higher values for both Precision and Recall than the Disease Identification Model, and this suggests its better ability to manage the classification process efficiently regarding disease severity.

TABLE III

Disease Detection and Severity Prediction performance summary

Metric	Disease Identification Model	Severity Level Prediction Model
Accuracy	81.94%	83.20%
Precision	0.85	0.86
Recall	0.85	0.86
F1 Score	0.87	0.86

Figure 1 shows the confusion matrix of the Severity Level Prediction Model, indicating that the model correctly classified most cases in the Mild and Severe classes, but confused some cases between the Moderate and Severe classes. The confusion matrix provides an overview of the capability of the model in differentiating among the different severity levels.

This study demonstrates the transformative potential of computational models in agriculture, optimizing crop monitoring, resource allocation, and disease management. The YOLOv8 model proves suitable for real-time classification, while Faster R-CNN excels in post-harvest quality assessment. The ensemble learning models for irrigation and fertilization successfully improve predictive accuracy, promoting efficient resource utilization. The pest and disease identification models exhibit high accuracy, although further improvements in data diversity and feature extraction are necessary. Future research should focus on integrating multi-spectral imaging, hardware acceleration, and advanced computational techniques to enhance the precision and effectiveness of agricultural monitoring systems.

V. Conclusion

Consequently, contemporary studies emphasize intelligent automation's capability to redefine tomato cultivation by solving major problems such as determination of maturity, detection of diseases, monitoring pests, and managing resources. The application of decision-making through data in these processes has the capability to improve crop quality considerably, minimize wastage, and ensure sustainability in future agriculture.

The results indicate that automation in detection during harvesting mature fruits has the potential to reduce losses during and after harvesting, subject to being performed in a timely and high-quality manner. It should be noted that traditional visual assessments have high variation; however, using systematic classification can improve reliability and accuracy in decision-making. Additionally, early detection of diseases can reduce damage to crops to a great extent and allows for effective use of pesticides, while better classification of severity allows for targeted action to maintain plant health and conserve nature's ecosystem.

A considerable problem in tomato cultivation comes from pests, causing heavy losses in crops and a near double application of pesticides. A new computer-aided detection system has now been developed to improve the accuracy in identifying pests and provide timely interventions. Early detection of pests would allow farmers to cut back on chemical pesticides, thus ensuring environmental sustainability. The efficient use of targeted irrigation and fertilization schemes in such environments is paramount. Traditional management techniques have led to high water and fertilizer withdrawal, with implications for soil health and negatively impacting the environment. This study explains how monitoring and automation technologies can be used to improve regulation of real-time water and fertilizer application in accordance with real-time plant and soil conditions, ensuring optimum growth with minimal wastage. Though these suggested strategies will be advantageous to farmers, they will face a number of challenges. Geographic variability plays a major role in contributing to inaccuracies in evaluations; therefore, it is critical to ensure high-resolution adaptations in predictive modeling. Another challenge that has to be overcome to enable effective application in a variety of agriculture environments includes enhancing real-time responsiveness and scalability.

This study demonstrates great potential in sophisticated agriculture methods that have the potential to radically change practices with tomato farming. By combining automation and value-based analyses, farmers can improve operational efficiencies, increase production, and implement environmentally friendly methods, thus creating a basis for a successful and ecologically friendly future in agriculture.

VI. References

[1] C. Maraveas, "Incorporating Artificial Intelligence Technology in Smart Greenhouses: Current State of the Art," *Applied Sciences,* vol. 13, p. 14, 2022.

[2] P. B. Savitha, A. Rai, N. Singh, C. C. Keshav and V. G. Neelambike, "Smart Greenhouse and Warehouse Monitoring with Disease," *IOP Conference Series: Materials Science and Engineering,* p. 12, 2023.

[3] W. A. Matondang, "Utilising machine learning and image processing, this computer vision system detects defects and improves tomato quality.," *Journal of Integrated Artificial Intelligence Science and Engineering,* vol. 1, 2025.

[4] Z. Lin, C. Zhang, Z. Liang, X. Zou and X. Li, "A method for nighttime tomato fruit detection and occlusion judgment based on deep learning and image processing.," *International Conference on Optical and Photonic Engineering,* vol. 13509, pp. 213-218, 2025.

[5] S. P. Mohanty, D. P. Hughes and M. Salathé, " Using deep learning for image-based plant disease detection.," *Frontiers in Plant Science,* p. 10, 2016.

[6] J. Liu, M. Wang and X. Li, "EfficientNet based recognition of maize diseases by leaf image classification," *Journal of Physics: Conference Series,* vol. 1693, 2020.

[7] S. Shahinfar, K. Kelman and L. Kahn, "Prediction of sheep carcass traits from early-life records using machine learning.," *Computers and Electronics in Agriculture,* vol. 156, pp. 159-177, 2019.

[8] J. Ma, S. Yang, D. Wang, K. Tang, X. Feng and X. Z. Feng, "Genetic mapping of a light-dependent lesion mimic mutant reveals the function of coproporphyrinogen III oxidase homolog in soybean.," *Frontiers in Plant Science,* vol. 11, p. 557, 2020.

[9] A. Fuentes, S. Yoon, S. C. Kim and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition.," *Sensors,* vol. 17, 2017.

[10] D. S. Parihar, B. Dogra, M. K. Dogra, S. K. Singh and R. Khurana, "Development and evaluation of notched concave disc seed drill for direct seeding of wheat in paddy stubble field.," *Journal of Agriculture and Food Research,* vol. 10, 2022.

[11] J. C. Moso, S. Cormier, C. de Runz, H. Fouchal and J. M. Wandeto, "Anomaly detection on data streams for smart agriculture," *Agriculture,* vol. 11, p. 1083 , 2021.

[12] R. Szczepanek, " Daily streamflow forecasting in mountainous catchment using XGBoost, LightGBM and CatBoost," *Hydrology,* vol. 9, p. 226 , 2022.

[13] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research,* vol. 30, pp. 79-82, 2005.