

REAL-TIME IMAGE-BASED MONITORING OF CHILLI PLANT MATURATION USING OBJECT DETECTION TECHNIQUES

***Zahida Yaseen**

Center for Research and Development (CRD), Minhaj University Lahore. Forman Christian College (A Chartered University Lahore), Pakistan

Zarka Saeed

School of Software Engineering, Minhaj University Lahore, Pakistan

Usama Asif

School of Software Engineering, Minhaj University Lahore, Pakistan

Khalid Masood

School of Computer Science, Minhaj University Lahore, Pakistan

Muhammad Sajjad

School of Computer Science, Minhaj University Lahore, Pakistan

RECEIVED

02 Feb2025

ACCEPTED

25 Feb2025

PUBLISHED

22March2025

ABSTRACT

Chilli crop monitoring plays a vital role in maximizing harvest, ensuring quality, and determining market willingness. Traditional manual approaches to judging crop maturity are labor-intensive and prone to error because of human involvement. Advancements in AI-driven technologies have allowed for efficient and accurate crop monitoring. This study proposed a Deep learning image-based monitoring framework for using advanced object detection techniques (YOLOV11 model) due to its superior performance in speed and accuracy in real-time object detection. The proposed system is designed to help farmers make informed yield decisions, thereby minimizing post-harvest losses and enhancing profitability, achieving a validation accuracy of 98.4%. Results show the robustness and practicality of the model for precision agriculture applications.

Key words: Chilli crops monitoring, Object detection, Deep learning, Yolo, Real-time object detection

Introduction

Agriculture is an economic pillar for people, particularly for people where a large portion of the economy depends on agriculture [1]. There are various crops cultivated among them; chilli has a significant position due to

high demand for domestic use, along with the markets. Chilli is not only valuable for cooking but also important for economic development [2]. But the optimal time for harvesting chilli is the persistent challenge.

The final stage (maturation) of chilli is very significant because it directly affects the quality, market price, flavor, and nutritional value [3]. If someone harvests it at an inappropriate time, it can lead to extensive losses due to inferior product quality. Traditionally, rely on visual inspection to determine the maturity stage of chilli crops [4]. Traditional methods are very simple, and no tools are needed. Traditional methods may lead to human error. Factors such as weather

variations, light conditions, and individual perception of color and size can affect the accuracy of traditional methods [5]. Furthermore, large farms make it impossible to watch every plant manually. These challenges limited the traditional methods and necessitated the development of automated solutions that may perform excellently and efficiently [6]. Figure 1 depicts the growth cycle of the chilli crop.

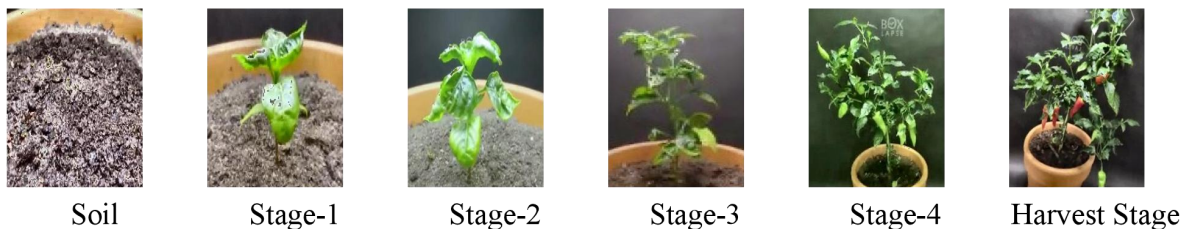


Figure 1: Chilli growth stages

The invention of digital technology, particularly in the field of artificial intelligence (AI), Machine Learning (ML), and Deep learning (DL), has paved the way for pivotal variations in technologies to analyze, monitor, and manage crops at a granular level [7]. Smart farming also involves the use of technology to monitor, analyze, and manage crops at a granular level [8]. Image-based monitoring is one of the best ways of smart farming [9]. Capturing images

and checking the crop at different levels may increase the production and prevent the chilli crops from different diseases. In the context of chilli crop maturation, image-based monitoring delivers many advantages. Firstly, it enables continuous monitoring of crops for large farms. Secondly, it ensures object evaluation and consistency, reducing subjectivity in decision-making. Thirdly, it makes timely decisions when integrated with real-time systems, which is important for

optimizing harvest schedules.

This study proposed a real-time, image-based monitoring system for monitoring and detecting the maturation stage of chilli by using object detection techniques. The main part of the system is a DL model YOLOv11, which is trained for accurate results and fast object detections. It has several enhancements over its predecessors, including best accuracy, overlapping objects, and improved real-time performance.

The study contributions are summarized as follows:

1. A labeled dataset of chilli crops was developed, containing various images of maturation stages.
2. Evaluation and implementation of the model in real-time conditions.
3. To get real-time feedback for precision harvesting from an integrated model with the system

To gain the objective, a labelled dataset for chilli was downloaded, representing various maturity stages. The dataset contained images of different light conditions and multiple images of different angles to improve the performance of the model. The YOLOv11

model was trained on that dataset and evaluated by Precision, recall, F-1 score, and mean Average Precision (mAP). The performance metrics values signify that the proposed model performed well as compared to the primary established models. A flow chart of the system is given below in Figure.

Related work:

Historically, manual measurements, visual inspections, and expert intuition have been used to monitor the crops' growth and determine when the crops are on time. These methods have many drawbacks, such as scalability, subjectivity, and inconsistency [10]. Some traditional tools, for example, rulers, color charts, and maturity indices, have been used to enhance the accuracy, but due to human involvement, these methods are still limited in their effectiveness [11]. Furthermore, these approaches fail to give real-time monitoring, which is very important for managing the large farms efficiently.

Deep learning has demonstrated remarkable potential in addressing complex real-world problems across diverse domains. In healthcare, CNN-based models have been applied to mobile applications for skin disease classification with emphasis on user privacy

[12], fungal disease diagnosis [13], and Urdu handwritten alphabet recognition for linguistic preservation [14]. Security and surveillance systems have been enhanced through intelligent ammunition detection and classification using CNNs [15]. In urban infrastructure, IoT-enhanced autonomous parking solutions utilizing transfer learning [16] and multi-sensor data fusion with deep learning techniques for parking lot detection [17] have been developed to optimize transportation efficiency. Waste management in smart cities has also benefited from CNN-powered intelligent agents for automatic image-based waste segregation [18]. Moreover, CNN architectures have been analyzed in depth for general image classification applications [19], reflecting their versatility in computer vision. These works collectively highlight how deep learning continues to enable innovative, accurate, and scalable solutions across healthcare, urban development, and intelligent security systems.

The application of ML in agriculture has limited the traditional methods [20]. ML techniques such as Random Forest (RF), Support Vector Machines (SVM), and K-

Nearest Neighbors (KNN) have been used for classifications, including disease detection, yield prediction, and crop type detection [21]. These models have failed because they required structural data and handcrafted feature engineering extracted from images. These models lack generalizability across different environmental conditions.

Deep learning models, especially Convolutional Neural Networks (CNNs), are designed to automatically extract features. CNNs have shown the best performance for detecting objects such as fruit detection, weed identification, and disease detection [22]. CNNs are also used in the classification of chilli maturity stages based on shape and color cues [23]. However, early models were not optimized for object detection in adverse environmental conditions.

Object detection models such as RNN [24], Faster-RNN [25], SSD [26], and the YOLO series have set an example by enabling real-time object detection. These models classified the objects as well as told about the location of the objects at once, making the models ideal for crop monitoring in adverse conditions. The Faster-RNN gives high accuracy but has limited speed compared to

YOLO [27]. YOLO provides a balance between real-time feasibility and performance. The YOLO series has undergone various improvements since its inception. The most recent model (YOLOv11) includes improved loss functions, enhanced object detection capabilities, and hybrid attention mechanisms, making the model suitable for monitoring chilli crops.

Methodology:

The leading objective of this study is to invent a real-time and robust monitoring system that can detect the maturation stages of chilli crops using deep learning-based object detection. The methodology tells how the data is collected, preprocessed, annotated, model selected, trained, evaluated, and deployed.

Dataset acquisition:

A high-resolution dataset was downloaded from Robo-Flow, a widely used platform for managing computer vision datasets. The dataset includes images of chilli crops captured under adverse environmental conditions and growth stages. All the images are labeled based on the three maturity categories: Green chilli, Orange /Yellowish

chilli, and Red Chilli.

Data Annotation:

Robo-flow has built-in annotated tools that are used to draw boundaries around each visible chilli pod in the images. Every image is tagged with its maturity level. This stage ensures that the model not only classifies the chilli but also locates the chilli in the fields.

Data Preprocessing:

Before model training, the dataset underwent several stages, including image resizing, data augmentation, and normalization.

Image resizing:

All the images were resized to a standard size of 416×416 or 640×640, suitable for the proposed model.

Data Augmentation:

In this stage, the images are converted into different angles, such as flipping, rotating, adjusting brightness, and cropping techniques were applied to increase the robustness of the model and reduce the overfitting.

Normalization:

The pixel values were set between 0 and 1 to enhance the training scalability of the model.

The data sets were split as follows:

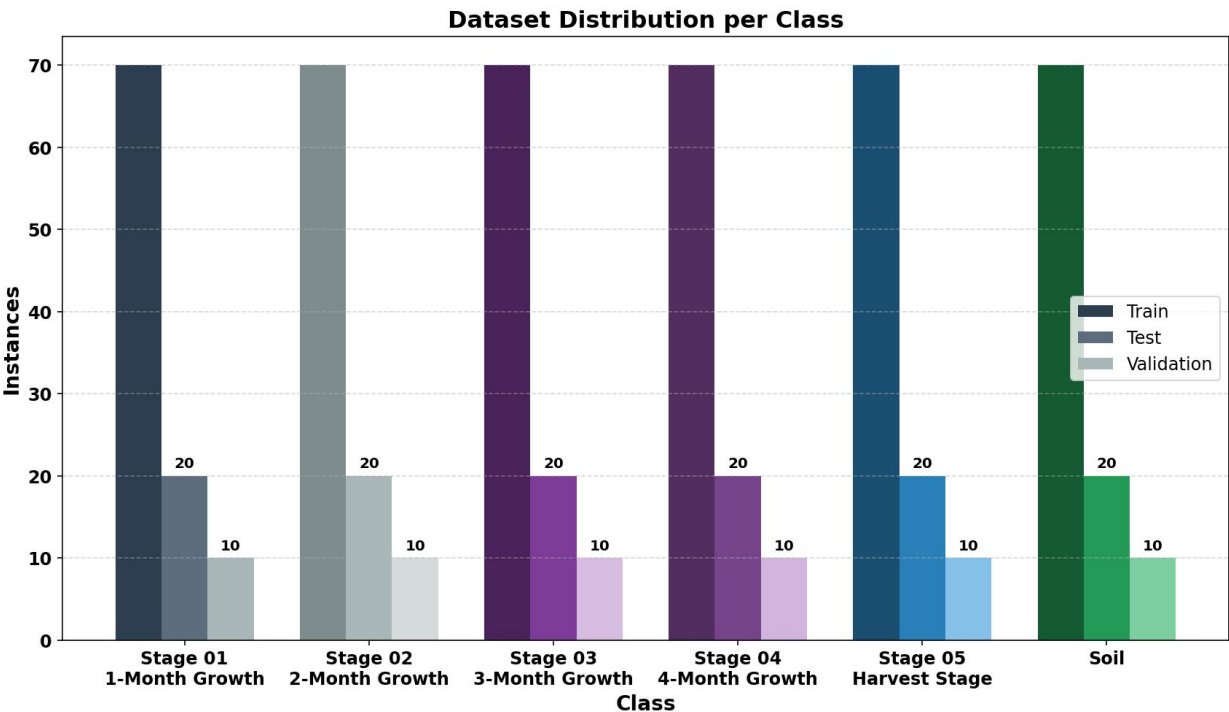


Figure 2: Data distribution

The graph in Figure 2 depicts that all the classes have balanced distributions, enabling the model to perform effectively in real-time monitoring of chilli crops across all growth stages.

Proposed Model Workflow:

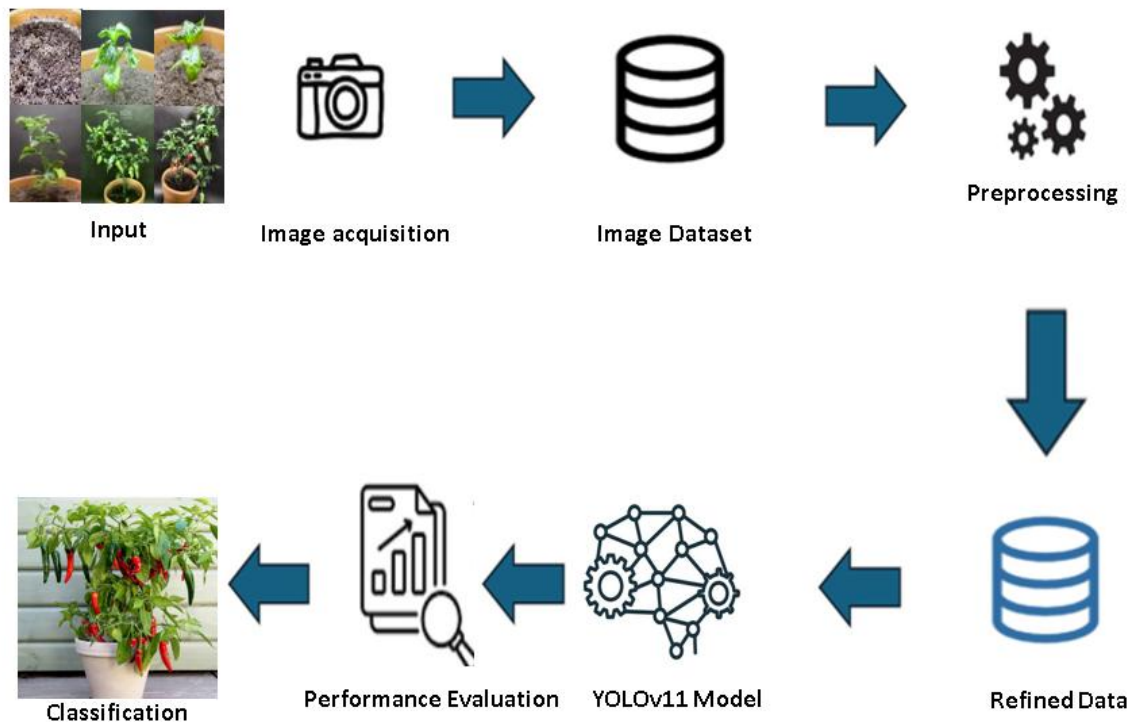


Figure 3: Proposed Model

Figure 3 depicts the complete workflow of the proposed chilli crop maturation monitoring system. The process starts with the input images of the chilli plant having different stages. The IOT is used to capture the images. Then images are compiled into a structured image dataset. The images pass through different preprocessing steps before training the model, such as image resizing, augmentation, and normalization. The dataset is split into training, validation, and testing sets to confirm the proper learning of the

model. The preprocessed and labeled dataset was then fed to the proposed model for training, the best object detection algorithm for its accuracy and speed. After the training phase, the system is evaluated by the validation phase and the testing phase. The performance evaluation metrics assist the model in validating its effectiveness in identifying and classifying the chilli fruits under adverse conditions. The final output is the real-time classification of the system. The proposed model individually identifies and

labels the maturity level of each plant, enabling automated tasks. The figure depicts the complete flow from data collection to intelligent classification, representing a practical application of DL in smart

agriculture.

Model Training:

The model was trained by annotated images with the following setup:

Table 1: Model Training Parameters

Framework	Pytorch
Input size	640×640
Batch size	16
Epochs	100
Optimized	Adam
Learning rate	0.001
Loss function: Classification loss, regression loss, and abjectness loss.	

The training was performed in a GPU-enabled environment.

Evaluation Metrics:

following metrics were used:

To calculate the model's performance, the

Table 2: Performance metrics

Metric	Formula
Accuracy	$\frac{Tp + Tn}{Tp + Tn + Fp + Fn}$
Precision	$\frac{Tp}{Tn + Fp}$
Recall	$\frac{Tp}{Tp + Fn}$
F-1 score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$
Specificity	$\frac{Tn}{Tn + Fp}$

Confusion Matrix:

performance for classifications.

The Confusion metric in Table 3 shows strong

Table 3: Confusion Metric

		Predicted Values					
Actual Values	N=120	1-Month	2 Months	3 Months	4 Months	Harvest	Soil
	1 Month	20	0	0	0	0	0
	2 Month	0	20	0	0	0	0
	3 Months	0	0	20	1	0	0
	4 Months	0	0	0	19	0	0
	Harvest	0	0	0	0	20	0
	Soil	0	0	0	0	0	20

Evaluation Metrics:

The bar chart represents the performance of the model across six classes including crop growth stages and soil. The model performed accurately 100% in stage-1, stage-2, stage-5, and soil with 0% miss rate. For the stage-4

model, perform effectively with a miss rate of 1.5%. However, a notable drop is seen in stage 3 with a miss rate of 9.09%. Overall model exhibits reliable and accurate classification abilities, with minimal errors in one class, as shown in Figure 5.

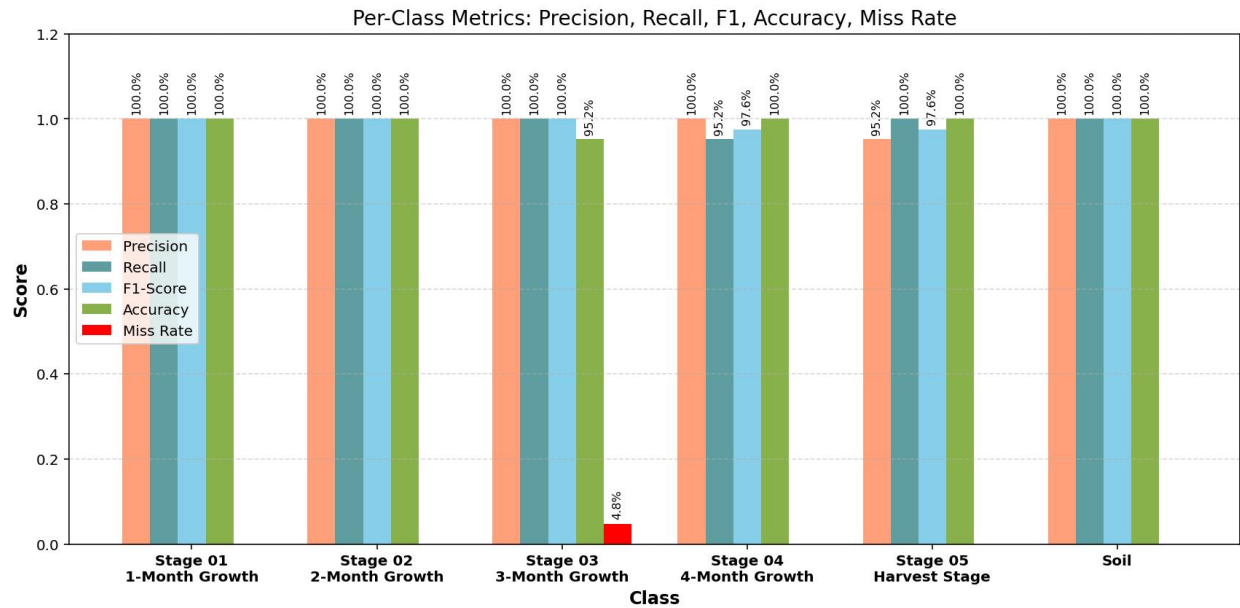


Figure 5: Performance Metrics

Precision and Recall:

The graph in Figure 6 depicts the performance

of the model over 150 training epochs. Initially, precision and recall values are

relatively low. After 20 epochs, both metrics started improving rapidly and became stable at epoch 40, gaining 95 % values. After 60 to 150 epochs, they gain a value of 99.98%,

indicating that the system is performing efficiently. Both metrics show the robustness and generalization of the model.

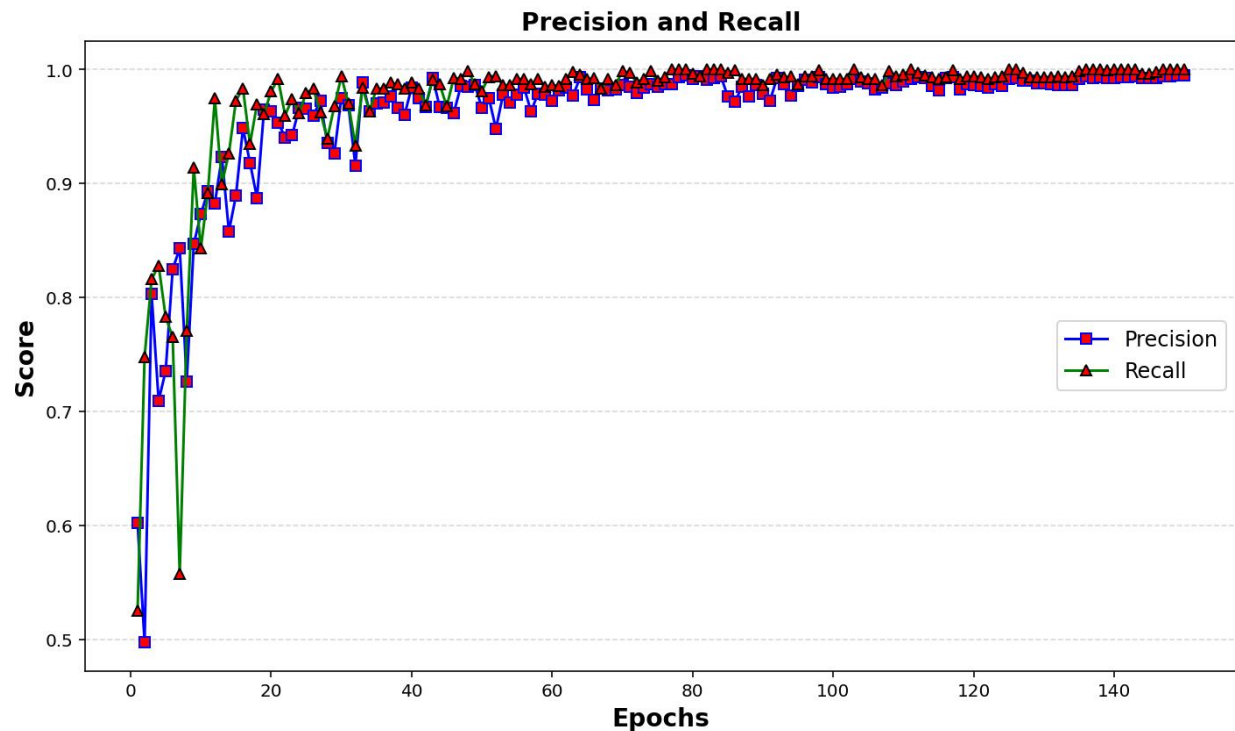


Figure 6: Precision-Recall Curve

Training vs Validation Loss:

The graph shown in Figure 7 indicates that both training and validation losses were high, but after 20 epochs, both losses decrease quickly, indicating a high learning rate of the model. The relative alignment of both losses

indicates a positive sign for robustness and generalization of the model. The graph shows that the model training is effective and stable. The consistently low values of both confirm that the model will perform accurately in real-world data.

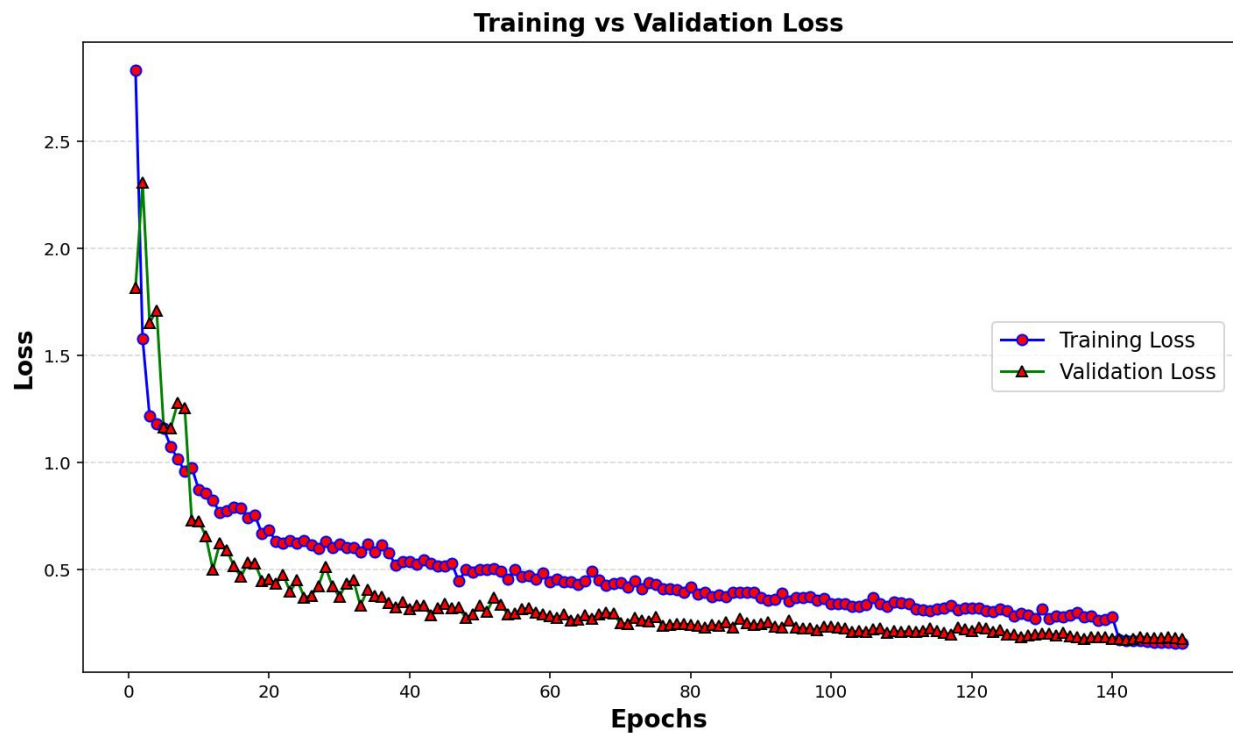
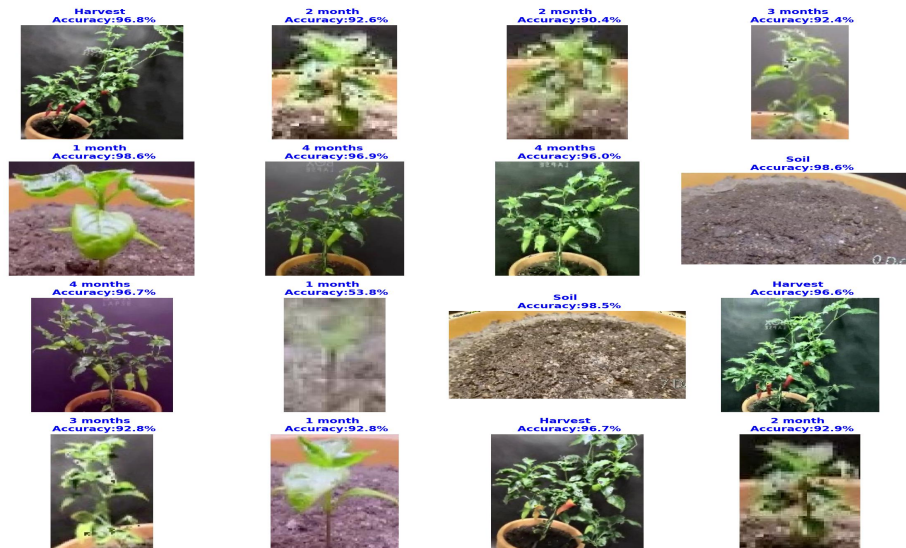


Figure 7: Training vs Validation loss

Model Performance Results:

The model efficiently classifies the growth stages and soil under adverse conditions with high accuracy, particularly for soil. Harvest months 4, 2, and 3 indicate moderate accuracy,

but the month one stage predicts less reliably, with one consistent drop. Overall model performance is good, with some uncertainty in early-stage classifications as shown in Figure 8 below.



Comparison Result of ML and DL Models:

Table 4 represents the comparison results of different ML, DL models, and our proposed model.

Table 4: Comparison Result

Model	Type	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
SVM (RBF Kernel) [28]	ML	81.4	78.9	80.1	79.5
Random Forest [29]	ML	84.2	80.5	82.3	81.7
k-NN (k=5) [30]	ML	79.6	77.3	78.4	78.0
CNN (Custom, 5 layers) [31]	DL	91.5	90.2	90.8	90.4
Faster R-CNN [31]	DL	93.2	92.8	93.0	92.5
YOLOv8n [32]	DL	95.0	94.7	94.8	94.5
Proposed Model	DL	99.20	99.20	99.20	99.2

Conclusion:

This study shows real-time image-based object detection techniques for monitoring chilli crop maturation, highlighting the efficiency of the models. YOLOv11 model demonstrates superior performance as compared to predecessors and other convolutional models in terms of speed, accuracy, and flexibility in variable agricultural conditions. The model handled the real-world challenges effectively. Its real-time capability, lightweight, and compatibility make it highly adoptable for field use.

In addition to enhancing technical performance, YOLOv11 supports automated crop monitoring, enabling timely decision-making for crop harvesting and resource optimization. The proposed system reduces human involvement, enhances the reliability, and these capabilities make the model accurate and efficient for deployment in the real world. The proposed model aligns with the broader goal of smart agriculture and sustainable production. Additionally, the YOLOv11 provides a reliable and efficient solution for real-time chilli crop maturity monitoring, providing a vital contribution to smart farming and AI-driven agriculture

systems.

References:

- G. Ahmad, A. Khan, M. Sajjad, A. Saddiqa, M. M. Naveed, and Z. Riaz, "TRANSFER LEARNING-BASED SMART CROP PROTECTION SYSTEM FOR ANIMAL DETECTION AND DETERRENCE," *Spectr. Eng. Sci.*, vol. 3, no. 7, pp. 810–899, 2025.
- J. Mariyono, "The economic aspects of chilli production in Central Java," *Econ. J. Emerg. Mark.*, vol. 8, pp. 85–97, Oct. 2016, doi: 10.20885/ejem.vol8.iss2.art1.
- S. Bal *et al.*, "Biomedical and Antioxidant Potentialities in Chilli: Perspectives and Way Forward.," *Molecules*, vol. 27, no. 19, Sep. 2022, doi: 10.3390/molecules27196380.
- F. Schneider, J. Swiatek, and M. Jelali, "Detection of Growth Stages of Chilli Plants in a Hydroponic Grower Using Machine Vision and YOLOv8 Deep Learning Algorithms," 2024. doi: 10.3390/su16156420.
- M. Półrolniczak and L. Kolendowicz, "The influence of weather and level of observer expertise on suburban landscape perception," *Build. Environ.*, vol. 202, p. 108016, 2021, doi: <https://doi.org/10.1016/j.buildenv.2021.108016>.
- N. N. Thilakarathne, M. S. Abu Bakar, P.

- E. Abas, and H. Yassin, "Internet of things enabled smart agriculture: Current status, latest advancements, challenges and countermeasures," *Heliyon*, vol. 11, no. 3, p. e42136, 2025, doi: <https://doi.org/10.1016/j.heliyon.2025.e42136>.
- [7] T. Anukiruthika and D. S. Jayas, "AI-driven grain storage solutions: Exploring current technologies, applications, and future trends," *J. Stored Prod. Res.*, vol. 111, p. 102588, 2025, doi: <https://doi.org/10.1016/j.jspr.2025.102588>.
- [8] S. John and A. L. R. P J, "Smart Farming and Precision Agriculture and Its Need in Today's World," 2024, pp. 19–44. doi: 10.1007/978-3-031-51195-0_2.
- [9] J. F. Restrepo-Arias, J. W. Branch-Bedoya, and G. Awad, "Image classification on smart agriculture platforms: Systematic literature review," *Artif. Intell. Agric.*, vol. 13, pp. 1–17, 2024, doi: <https://doi.org/10.1016/j.aiia.2024.06.002>.
- [10] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine Vision Systems in Precision Agriculture for Crop Farming.," *J. imaging*, vol. 5, no. 12, Dec. 2019, doi: [10.3390/jimaging5120089](https://doi.org/10.3390/jimaging5120089).
- Anjali *et al.*, "State-of-the-art non-destructive approaches for maturity index determination in fruits and vegetables: principles, applications, and future directions," *Food Prod. Process. Nutr.*, vol. 6, no. 1, p. 56, 2024, doi: [10.1186/s43014-023-00205-5](https://doi.org/10.1186/s43014-023-00205-5).
- Gulzar Ahmad *et al.*, "Mobile Application for Skin Disease Classification Using CNN with User Privacy," *J. Comput. Biomed. Informatics*, no. SE-Articles, Apr. 2024, [Online]. Available: <https://jcabi.org/index.php/Main/article/view/432>
- T. Alyas, K. Alissa, A. S. Mohammad, S. Asif, T. Faiz, and G. Ahmed, "Innovative Fungal Disease Diagnosis System Using Convolutional Neural Network," 2022. doi: [10.32604/cmc.2022.031376](https://doi.org/10.32604/cmc.2022.031376).
- G. Ahmed *et al.*, "Recognition of Urdu Handwritten Alphabet Using Convolutional Neural Network (CNN)," 2022. doi: [10.32604/cmc.2022.029314](https://doi.org/10.32604/cmc.2022.029314).
- G. Ahmad *et al.*, "Intelligent Ammunition Detection and Classification System Using Convolutional Neural Network,"

- Comput. Mater. Contin.*, vol. 67, no. 2, pp. 2585–2600, 2021, doi: <https://doi.org/10.32604/cmc.2021.015080>.
- [16] Q. Abbas, G. Ahmad, T. Alyas, T. Alghamdi, Y. Alsaawy, and A. Alzahrani, “Revolutionizing Urban Mobility: IoT-Enhanced Autonomous Parking Solutions with Transfer Learning for Smart Cities,” 2023. doi: 10.3390/s23218753.
- [17] kashif iqbal Shah, S. Abbas, M. Khan, A. Athar, M. Saleem Khan, and G. Ahmad, “Autonomous Parking-Lots Detection with Multi-Sensor Data Fusion Using Machine Deep Learning Techniques,” *C. -Tech Sci. Press.*, vol. 66, pp. 1595–1612, Nov. 2020, doi: 10.32604/cmc.2020.013231.
- [18] J. Vargas, G. Ahmad, B. Sheikh, Hameed, R. Bokhari, and T. Tajammal, “Smart Cities’ Automatic Image-Based Waste Segregation through an Intelligent Agent Using CNN,” Mar. 2025.
- [19] X. Wang, Y. Lu, Y. Wang, and W. B. Chen, “Diabetic retinopathy stage classification using convolutional neural networks,” *Proc. - 2018 IEEE 19th Int. Conf. Inf. Reuse Integr. Data Sci. IRI 2018*, no. October 2019, pp. 465–471, 2018, doi: 10.1109/IRI.2018.00074.
- M. Waqas, A. Naseem, U. W. Humphries, P. T. Hlaing, P. Dechpichai, and A. Wangwongchai, “Applications of machine learning and deep learning in agriculture: A comprehensive review,” *Green Technol. Sustain.*, vol. 3, no. 3, p. 100199, 2025, doi: <https://doi.org/10.1016/j.grets.2025.100199>.
- S. Shitole, “Analysis of Crop disease detection with SVM, KNN and Random forest classification,” *Inf. Technol. Ind.*, vol. 9, pp. 364–372, Mar. 2021, doi: 10.17762/itii.v9i1.140.
- Z. Wang *et al.*, “An improved Faster R-CNN model for multi-object tomato maturity detection in complex scenarios,” *Ecol. Inform.*, vol. 72, p. 101886, 2022, doi: <https://doi.org/10.1016/j.ecoinf.2022.101886>.
- Y. Hendrawan *et al.*, “Classification of large green chilli maturity using deep learning,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 924, p. 12009, Nov. 2021, doi: 10.1088/1755-1315/924/1/012009.
- Y. Wang *et al.*, “A multi-scale feature selection module based architecture for the diagnosis of Alzheimer’s disease on [18F]FDG PET,” *Int. J. Med. Inform.*, vol. 191, p. 105551, 2024, doi: <https://doi.org/10.1016/j.ijmedinf.2024.105551>.

- 05551.
- [25] B. Karbouj, G. A. Topalian-Rivas, and J. Krüger, “Comparative Performance Evaluation of One-Stage and Two-Stage Object Detectors for Screw Head Detection and Classification in Disassembly Processes,” *Procedia CIRP*, vol. 122, pp. 527–532, 2024, doi: <https://doi.org/10.1016/j.procir.2024.01.077>.
- [26] W. Liu *et al.*, “Ssd: Single shot multibox detector,” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, Springer, 2016, pp. 21–37.
- [27] J. Redmon, “Yolov3: An incremental improvement,” *arXiv Prepr. arXiv1804.02767*, 2018.
- [28] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018, doi: <https://doi.org/10.1016/j.compag.2018.02.016>.
- [29] L. V Tan, T. Tran, and H. H. Loc, “Soil and Water Quality Indicators of Diversified Farming Systems in a Saline Region of the Mekong Delta, Vietnam,” 2020. doi: 10.3390/agriculture10020038.
- [30] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine Learning in Agriculture: A Review,” 2018. doi: 10.3390/s18082674.
- E. Gomes, L. Bertini, W. R. Campos, A. P. Sobral, I. Mocaiber, and A. Copetti, “Machine Learning Algorithms for Activity-Intensity Recognition Using Accelerometer Data,” 2021. doi: 10.3390/s21041214.
- Y. Wang *et al.*, “YOLO-ALW: An Enhanced High-Precision Model for Chili Maturity Detection,” 2025. doi: 10.3390/s25051405.