Precise Potato Leaf Disease Detection with You Only Look Once Models: A Comparative Analysis

Touhid Alam 1, a), Abir Bokhtiar 1, b), Md Saef Ullah Miah 1, c), Md. Jakir Hossen 2, d)

1Department of Computer Science, American international University-Bangladesh, 408/1 (Old KA 66/1), Kuratoli, Khilkhet, Dhaka 1229, Bangladesh

2 Center for Advanced Analytics (CAA),COE for Artificial Intelligence, Faculty of Engineering and Technology (FET), Multimedia University, Jalan Ayer Keroh Lama 75450, Bukit Beruang, Melaka

*d) Corresponding author: jakir.hossen@mmu.edu.my*

*a) 22-4330-1@student.aiub.edu  
b) 22-47038@student.aiub.edu*

*c) saef@aiub.edu*

**Abstract.** Potato leaf diseases are significant challenges in agriculture, which causes substantial yield losses, economic difficulties for farmers, and potential food security threats. Early blight, caused by fungi, thrives in warmer temperatures and is characterized by brown-black oval spots on leaves. Late blight, caused by oomycetes, prefers cooler, moist environments and presents dark brown spots with whitish fungal growth. Early detection of these diseases is necessary to reduce crop loss and ensure sustainable agricultural practices. This study systematically implemented You Only Look Once (YOLO) models in detecting potato leaf diseases, focusing on three different classes. The dataset used for this research was sourced from Kaggle containing a total of 2,152 images. Preprocessing techniques such as data resizing and dummy annotations were applied. Among the YOLO models, YOLOv11 demonstrated the highest performance, achieving a mAP50 of 99.5% and a mAP50−95 of 98.4%, precision of 98.1% and a recall of 97.1%. These results outperformed the previous versions of YOLO models and other transformer models. These results show the advancements of architectural components used in YOLOv11 such as C3K2 and C2PSA, improving feature extraction and attention mechanism techniques. Furthermore, YOLOv11 requires only 2.6 million parameters, significantly less than its previous versions and other transformers.

# INTRODUCTION

In Over 140 countries, potato is the 4th most cultivated food following maize, wheat and rice [1]. This shows the importance of potatoes in world food crops. Nevertheless, potatoes face numerous diseases that threaten the yield and quality of potatoes [2]. Both early blight and late blight can cause serious problems under favorable conditions, with potential loss of 50% cross in various scenarios [3]. Among them, late blight is a more significant problem for potatoes, causing an annual loss of around $10 billion annually [4]. It has been shown that integrating disease management strategies can reduce disease incidence by 30-70% [5]. Therefore, it is important to intervene and prevent these diseases timely.

One of the key solutions of addressing this challenge is automated detection system using computer vision. While traditional Convolutional Neural Networks (CNNs) have shown success in image classification, object detection models can be more effective for leaf disease detection in specific instances of disease on leaves. YOLO models are one of the key advancements used for detection related implementations. However, while existing YOLO implementations face limitations regarding their accuracy and generalizability with proper evolution metrics [6]. Furthermore, current studies focus on broader applications without diving into the unique complexities of specific leaf diseases [6]. This gap undermines the development of effective, integrated disease detection techniques.

As a result, this research focuses on systematic implementation of recent YOLO models for detecting potato leaf disease accurately. The objective of the study was to identify the model offering highest accuracy and efficiency so that agricultural productivity and sustainability can be enhanced. Our study highlights the superior performance of YOLOv11 compared to previous versions of the model, achieving detection accuracy of 99.5% while using significantly less parameters (2.6 million). This advancement not only shows the potential of implementing the model for timely intervention for potato leaf diseases but also represents a significant step forward in sustainable agriculture.

# LITERATURE REVIEW

This section of the paper gives an overview of recent studies involving YOLO models and potato leaf disease detection.

J. Wang and H. Zhao proposed YOLOv8-MSS algorithm for surface object detection [7]. The model achieved 5% increase in mAP50 (Mean Average Precision with 50% confidence) compared to YOLOv8. Their model implemented 4x down sampling detection head to the YOLOv8n model and replaced the C2f module in the backbone with C2fMLCA to the noise from complex water surface environments. Furthermore, the use of SloU loss function instead of CloU was used to improve the bounding box accuracy. X. Wang et al. introduced YOLOv8n-vegetable model which was able to outperform existing object detection models in terms of precision, recall, mAP, parameters, model size, and processing speed for vegetable disease detection [8]. R. He et al. developed a model named YOLOv9-LSBN model that had better precision and recall compared to other YOLO models for cotton pest detection and identification [9]. They replaced the RepNCSPELAN4 module in YOLOv9 with advanced RepLanLsk module and employed a series of depth-wise convolution kernels to extract more precise information. D. Lu and Y. Wang demonstrated a model named MAR-YOLOv9 superior performance in object detection and counting tasks on four plant datasets, achieving significant improvements in metrics like precision, recall, and mAP [10]. S. Guan et al. integrated BiFPN module with YOLOv10 that achieved mAP50 of 95.10% in wheat spike detection, outperforming YOLOv10 model [11]. J. Ou et al. proposed GAS\_YOLO model based on YOLOv10 for road defect detection, especially for small object detection with an improvement of 10.8% in parameter count and with a mAP50 of 86.5% [12]. A. T. Khan and S. M. Jensen showed that YOLOv11 had a mAP50 of over 0.85 and 0.7 respectively for bounding box tasks with spring crops and winter crops [13]. Using primary data, they used two training approaches: a combined dataset across all crops and individual datasets for each crop. L. He et al. devised YOLOv11-Seg model achieving high accuracy in object detection and segmentation on construction sites, with over 80% of test samples having confidence scores above 90% in constructor sites object segmentation [14]. These advancements using YOLO models are potentially transferable to detect various agricultural diseases.

H. Ghosh et al. recommended transfer learning model for potato leaf disease including VGG19, DenseNet121, and ResNet50 where VGG19 had an accuracy of 98.77% and loss of 0.028 on the training set [15]. Although DenseNet121 outperformed VGG19 and ResNet50 in terms of validation loss and accuracy. J. Pasalka et al. presented CNN-based approach achieved 97.4% accuracy in classifying healthy potato leaves and infected leaves [16]. J. Rashid et al. developed a custom model called PDDCNN that was able to detect potato leaf disease with an accuracy of 98.75% with data augmentation [17]. M. G. Lanjewar et al. used four transfer learning models. Among them, DenseNet achieved an accuracy of 99% for potato leaf disease detection [18]. Furthermore, their modified VGG19 and NasNet model had an accuracy of 98%.

# METHODOLOGY

This section of the paper gives a detailed description of the proposed methodology and the YOLO models. Figure 1 shows the methodology of the study.

|  |
| --- |
|  |

**FIGURE 1.** Proposed methodology of the study

## Data Collection

The Potato Leaf Disease dataset is a well-structured collection of images obtained from Kaggle [19]. This dataset comprises 2,152 high-quality images distributed into three categories: 152 images of healthy leaves and two categories (Early Blight and Late Blight) evenly distributed into 1,000 images each. This balanced dataset simplifies the training and evaluation process for machine learning models. Figure 2 shows examples for three classes of the dataset. The images were resized to a uniform dimension of 128 × 128 for optimal performance across the models and device compatibility. For YOLO- compatible data preparation, annotation files were generated in the required format, where each image was associated with a corresponding .txt file containing the coordinates. For YOLO compatible data preprocessing, dummy annotations (0.5 0.5 0.5 0.5) were initially used as placeholders to structure the label files. These annotations ensure proper alignment with the training and validation datasets for YOLO model training. The bounding boxes were manually annotated to accurately demarcate the regions of interest using

|  |
| --- |
|  |

**FIGURE 2.** Example of the classes of the dataset

## YOLO Model Architecture

In 2015, YOLO was introduced [20]. Since then, significant advancements have been made, enhancing both detection accuracy and computational efficiency across successive YOLO versions. Each version introduced innovative techniques [21] such as YOLOv2 introduced multi-scale training and dimension clustering. YOLOv3 incorporated the SPP block and the Darknet-53 backbone. YOLOv4 featured the Mish activation function and the CSPDarknet-53 backbone. YOLOv5 introduced anchor-free detection and SWISH activation. YOLOv7 introduced architectural advancements such as Extended-ELAN (E-ELAN), model scaling techniques for different computational budgets [22]. Ultralytics improved the architecture further with YOLOv8, introducing five scaled versions for objectness, classification, and regression tasks. It utilizes a modified CSPDarknet53 backbone with a C2f module, an SPPF layer for feature pooling, batch normalization, and SiLU activation [23]. YOLOv9 introduced improvements such better gradient flow and information transfer [24]. YOLOv11, released in 2024 [25], introduces The C3k2 Block enhances feature extraction with an efficient convolutional structure. SPPF (Spatial Pyramid Pooling - Fast) effectively captures multi-scale features. Attention mechanism was introduced by C2PSA (Cross Stage Partial with Spatial Attention). Moreover, YOLOv11 has 22% less parameters compared to YOLOv8m. These enhancements enable YOLOv11 to achieve higher accuracy, faster processing speeds, and adaptability across diverse computer vision tasks, outperforming its predecessors in efficiency and precision.

## Experimental Setup

The experimental setup involved 50 epochs with a batch size of 16 to ensure consistency across evaluations. All experiments were conducted on a system containing an Intel Core i3-13th generation processor, utilizing its integrated GPU.

## Evaluation Metrics

Various evolution metrics such as precision, recall, mAP50, mAP50-95, F1 score, confusion matrix were used to ensure constant performance and efficiency of the models in detecting the diseases accurately.

# RESULT AND DISCUSSION

Table 1 presents the performance metrics of the evaluated models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TABLE 1.** Performance of the YOLO models for potato leaf disease detection | | | | |
| **Model** | **Precision** | **Recall** | **mAP50** | **mAP50-95** |
| YOLOv8 | 88.7% | 91.8% | 98.3% | 94.1% |
| YOLOv9 | 97.6% | 97.8% | **99.5%** | 91.9% |
| YOLOv10 | 83.2% | 91.2% | 96.1% | 93.9% |
| YOLOv11 | **98.1%** | **97.1%** | **99.5%** | **99.0%** |

Based on evolution metrics, YOLOv11 performed the best across all the evolution metrics. YOLOv11 achieved the highest precision at 98.1%, highest recall at 97.1%, highest mAP50-95 at 99.0% and highest mAP50 at 99.5%. Both YOLOv11 and YOLOv8 achieved the highest mAP50 at 99.5%. The plots depict various metrics for both the training and validation datasets. The performance of the YOLOv11 model throughout the epochs can be seen on Figure 3. The consistent and generally improving trends observed in these plots indicate that the YOLOv11 model achieves a robust and generalizable performance in potato leaf disease. Figure 4 depicts the F1-confidence curve. The model achieved an F1 score of 0.96 with a confidence of 79.8%, while maintaining a precision-recall value of 99.5%. This shows the excellent trade off ability of the YOLOv11 model for false positive rate and true positive rate at high confidence level for potato leaf disease.

|  |
| --- |
|  |
| **FIGURE 3.** Evolution metrics of the YOLOv11 model across epochs |

Figure 5 shows the precision-recall tradeoff for the YOLOv11 model. YOLOv11 had excellent performance in the trade-off for precision recall at all three classes at 99.5%. This shows the great generalization ability of the YOLOv11 model for potato leaf disease detection. Figure 6 presents the confusion matrix of the YOLOv11 model, highlighting its strong classification performance across all categories. For the validation set, the YOLOv11 showed great performance across all the classes. Although, in the total, the model only detected 21 images as background out of 430 images in the validation set, while they were classes of the dataset. Figure 7 showcases sample detections made by the YOLOv11 model across various classes, demonstrating its effectiveness in real time object detection. Finally, Table 2 compares accuracy of YOLOv11 with previous research of potato leaf disease detection with transfer learning models. YOLOv11 has also showed better performance than other models.

|  |  |
| --- | --- |
|  |  |
| **FIGURE 4.** F1-confidence curve of the YOLOv11 model. | **FIGURE 5.** Precision-recall tradeoff of the YOLOv11 model. |

|  |  |
| --- | --- |
|  |  |
| **FIGURE 6.** Confusion matrix of the YOLOv11 model. | **FIGURE 7.** Detection examples of the YOLOv111 model on the validation set. |

|  |  |  |
| --- | --- | --- |
| **TABLE 2.** Comparison of YOLOv11 model with other studies. | | |
| **Authors** | **Models** | **Accuracy (%)** |
| H. Ghosh et al. [15] | VGG19 | 98.77 |
| J. Pasalka et al. [16] | VGG16 | 97.40 |
| J. Rashid et al. [17] | PDDCNN | 98.75 |
| M. G. Lanjewar et al. [18] | DenseNet | 99.00 |
| This Study | **YOLOv11** | **99.50** |

# CONCLUSION

This study demonstrated the performance of state-of-the-art YOLO models in detecting potato leaf diseases. The experimental results unequivocally demonstrate that YOLOv11 outperforms its predecessors across key evaluation metrics, achieving exceptional performance. These superior performance metrics underscore the enhanced capabilities of YOLOv11 in accurately and reliably detecting potato leaf diseases. While these results are promising, they are based on a specific dataset and focus on a comparative application rather than developing novel modifications in architecture. Future research should focus on evaluating YOLOv11's performance on more diverse datasets, encompassing images collected from various geographical locations, under different lighting conditions, and across multiple potato varieties, to robustly assess real-world generalizability and applicability. Nevertheless, the findings underscore the importance of utilizing cutting-edge machine learning techniques to address critical challenges in agriculture.

# References

1. Z.-J. Wang, H. Liu, F.-k. Zeng, Y.-c. Yang, D. Xu, Y.-C. Zhao, X.-f. Liu, L. Kaur, G. Liu, and J. Singh, “Potato processing industry in china: Current scenario, future trends and global impact,” Potato research 66, 543–562 (2023).
2. 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 (2024).
3. J. Luck, M. Asaduzzaman, S. Banerjee, I. Bhattacharya, K. Coughlan, G. Debnath, D. De Boer, S. Dutta, G. Forbes, W. Griffiths, et al., “The effects of climate change on pests and diseases of major food crops in the asia pacific region,” Final Report for APN (Asia-Pacific Network for Global Change Research) Project 73 (2012).
4. M. BERHAN, “Review on epidemiology, sampling techniques, management strategies of late blight (phytophthora infestans) of potato and its yield loss,” Asian Journal of Advances in Research 4, 199–207 (2021).
5. S. Khanal, K. Karimi, S. Majumdar, V. Kumar, R. Verma, S. K. Bhatia, K. Kuca, J. Esteban, and D. Kumar, “Sustainable utilization and valorization of potato waste: State of the art, challenges, and perspectives,” Biomass Conversion and biorefinery 14, 23335–23360 (2024).
6. M.Y. Xin, L.W. Ang, and S. Palaniappan, “A Data Augmented Method for Plant Disease Leaf Image Recognition based on Enhanced GAN Model Network,” Journal of Informatics and Web Engineering 2(1), 1–12 (2023).
7. J. Wang and H. Zhao, “Improved yolov8 algorithm for water surface object detection,” Sensors 24, 5059 (2024).
8. X. Wang and J. Liu, “Vegetable disease detection using an improved yolov8 algorithm in the greenhouse plant environment,” Scientific Reports 14, 4261 (2024).
9. R. He, P. Li, J. Zhu, F. Zhang, Y. Wang, T. Zhang, D. Yang, and B. Zhou, “Yolov9-lsbn: An improved yolov9 model for cotton pest and disease identification method,” (2024).
10. D. Lu and Y. Wang, “Mar-yolov9: A multi-dataset object detection method for agricultural fields based on yolov9,” Plos one 19, e0307643 (2024).
11. S. Guan, Y. Lin, G. Lin, P. Su, S. Huang, X. Meng, P. Liu, and J. Yan, “Real-time detection and counting of wheat spikes based on improved yolov10,” Agronomy 14, 1936 (2024).
12. J. OU, J. Zhang, H. Li, and B. Duan, “An improved yolov10-based lightweight multi-scale feature fusion model for road defect detection and its applications,” Available at SSRN 4970753.
13. A. T. Khan and S. M. Jensen, “Leaf-net: A unified framework for leaf extraction and analysis in multi-crop phenotyping using yolov11,” Agriculture 15, 196 (2025).
14. L. He, Y. Zhou, L. Liu, and J. Ma, “Research and application of yolov11-based object segmentation in intelligent recognition at construction sites,” Buildings 14, 3777 (2024).
15. H. Ghosh, I. S. Rahat, K. Shaik, S. Khasim, and M. Yesubabu, “Potato leaf disease recognition and prediction using convolutional neural networks,” EAI Endorsed Transactions on Scalable Information Systems 10 (2023).
16. J. Pasalkar, G. Gorde, C. More, S. Memane, and V. Gaikwad, “Potato leaf disease detection using machine learning,” Current Agriculture Research Journal 11 (2023).
17. J. Rashid, I. Khan, G. Ali, S. H. Almotiri, M. A. AlGhamdi, and K. Masood, “Multi-level deep learning model for potato leaf disease recognition,” Electronics 10, 2064 (2021).
18. M. G. Lanjewar, P. Morajkar, and P. P, “Modified transfer learning frameworks to identify potato leaf diseases,” Multimedia Tools and Applications 83, 50401–50423 (2024).
19. M. A. Putra, “Potato leaf disease dataset,” (2021), accessed: [20-2-25].
20. X. Han, J. Chang, and K. Wang, “You only look once: unified, real-time object detection,” Procedia Computer Science 183, 61–72 (2021).
21. R. Khanam and M. Hussain, “Yolov11: An overview of the key architectural enhancements,” arXiv preprint arXiv:2410.17725 (2024).
22. C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, “Trainable bag-of-freebies sets new state-of-the-art for real-time object detectorsv.org/abs/2207.02696,” arXiv preprint arXiv: 2207.02696 (2022)
23. Ultralytics, “Ultralytics yolov8,” (2024), accessed: [20-2-25].
24. C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, “Yolov9: Advancing the yolo legacy,” LearnOpenCV (2024), accessed: [20-2-25].
25. Ultralytics, “Ultralytics yolov11,” (2024), accessed: [20-2-25].