INSECT DAMAGE DETECTION IN CORN LEAF USING YOLOv10 MODELS: A TRANSFER LEARNING APPROACH

Suresh Timilsina[1,2], Sandhya Sharma[1], Biplove Pokhrel[3], Pradip Aryal[2], Kazuhiko Sato[1], Yoshifumi Okada[1], Shinya Watanabe[1], Satoshi Kondo[1]*

[1] Muroran Institute of Technology, 27-1 Mizumoto-cho, Muroran, Hokkaido 050-8585, Japan

[2] Tribhuvan University, IOE, Pashchimanchal Campus, Pokhara, Gandaki Province, Nepal

[3]Pokhara University, Gandaki College of Engineering and Sciences, Gandaki Province, Nepal

Abstract

Insect pests that affect corn leaves are significant contributors to agricultural losses. Despite corn being a globally consumed food crop, there is lack of research on the automated identification of these pest's damage, which limits farmers` ability to implement effective control measures. This research aims to fill these gaps by implementing various YOLOv10 architectures (10n, 10s, 10m, 10l, 10x) for the automated identification and localization of insect pests on corn leaves. A total of 1,395 images of highly affected corn leaves from Pokhara, Nepal, were collected, showing damage from grasshoppers and fall armyworms. The collected images underwent5-fold cross validation and split dataset into Train, Valid, Test in a ratio of 3:1:1 being fed into various architectures. Among the pest detection models tested, YOLOv10x outperformed the others in terms of mean average precision (mAP50 = 0.799) at an IoU threshold of 50%. Regarding the Precision-Recall (PR) relationship YOLOv10m has balancing nature of highest precision (Precision= 0.808) with high recall value (Recall= 0.807), this trend is not followed by YOLOv10l and YOLOv10x. In terms of latency, YOLOv10n required the least average time to detect each image during testing, making it the most lightweight model among those tested. The central tendency of YOLOv10m is highest among all YOLOv10 variants with decent spread box over inter-quartile range. Overall, with sufficient resource availability YOLOv10x is beneficial while with resource constraints condition YOLOv10m is efficient choice for insect damage detection in corn leaves. This research is valuable for farmers in detecting insect pests on corn leaves, enabling timely implementation of measures to reduce agricultural damage caused by pests. We also recommend optimizing the architectures to enhance the precision of pest detection.

Keywords: *corn, insect damage, object detection, YOLOv10*

1. INTRODUCTION

Corn (*Zea mays*) ranks at the top among cereal crops, contributing 20% to global production. Its versatility spans food, fuel, textiles, and medicinal applications [1,2,3]. However, pathogens and pests significantly reduce agricultural yield and quality, with global yield loss for corn estimated at 22.5% [4]. The U.S. Department of Agriculture reports that declines in corn production during 2019 and 2021 are primarily attributed to disease and pest pressures [5,6]. In 2022, corn yields experienced a notable reduction due to pest infestations, including grasshopper damage, which resulted in the loss of 44,341.4 thousand bushels, and Fall armyworm infestation, which caused a loss of 938.7 thousand bushels [7]. The Fall Armyworm damages maize by creating holes in the leaves and burrowing into the cob, disrupting nutrient flow and photosynthetic capacity, ultimately reducing yields. Similarly, grasshoppers cause substantial damage by feeding on leaves, stems, and flowers, leading to rapid defoliation and stunted growth.

Detecting pest-induced damage in crops is traditionally a slow and error-prone process, requiring expert manual inspections that are impractical for large-scale fields. However, advancements in deep learning offer promising solutions for rapid and accurate detection of pest-induced damage to leaves. Early detection of pest damage prevents widespread crop losses, minimizes pesticide use, and enhances overall yields. Object detection models are particularly useful in assessing damage severity and identifying overlapping damage caused by multiple pest species.

For example, a study implementing single-shot object detection models in cruciferous crops identifies YOLOv5l as the best-performing variant in terms of accuracy, precision, recall, and F1-score [8]. Similarly, a machine learning and deep learning approach for identifying insect-damaged and undamaged corn plants achieves the highest accuracy using a Random Forest model for early detection [9]. Another study demonstrates a deep learning model's high predictive performance, achieving an average AUROC of 0.917 in determining pest risk labels in crops such as strawberries, peppers, grapes, tomatoes, and paprika [10]. Research on vineyard pest damage detection using YOLOv8x, YOLOv7, DETR, and RTMDet shows that YOLOv8x achieves the best metrics, with an mAP50 score of 0.926 [11]. Recent studies in cereal crops primarily focus on identifying pest types [12–15]. However, pests are not always present on leaves during data collection, as some are nocturnal. Additionally, assessing the severity of insect damage remains a critical challenge. Object detection techniques address this by identifying pest-damaged portions of leaves, enabling farmers to estimate damage severity and implement timely remedies.

In this research, we develop a model to detect insect damage in corn leaves. This model leverages YOLOv10 variants, and performance metrics are used to identify the best-performing model. The optimized model is further tested on an unseen dataset, demonstrating robust performance and practical utility for real-world applications.

2. MATERIALS AND METHODS

2.1 Data Collection and Pre-processing

A total of 1395 images of insect damaged corn leaves were collected from Pokhara, Nepal. Two main damage types were identified in that region i.e. Grasshopper and Fall Armyworm. Canon80d camera was used for data collection purpose, which is non-invasive approach. The images were categorised with the help of agriculture expert. Table 1 below provides detailed information on datasets along with types of datasets and data collection sources. Images with multiple leaves in a single image were cropped before categorisation, but images with any background are accepted to generalise its implementation. Images of Grasshopper, Fall Armyworm and collective sample of both damages are presented in Fig. 1.

Table 1. Detailed information on the corn leaf insect damage datasets, types of datasets and data collection sources.

Dataset	No. of Images	Source	Types of Datasets
Train	837	Field Collected	Grasshopper damaged images.
Validation	279	Field Collected	Fall Armyworm damaged images.
Test	279	Field Collected	Leaves damaged by both types of insects.

Grasshopper Damaged Corn Leaf | Grasshopper and Fall Armyworm Damaged Corn Leaf | Fall Armyworm Damaged Corn Leaf

Fig. 1. Image samples of corn leaf damaged by insect attack

2.2 Object Detection Module

Object detection plays a crucial role in deep learning techniques particularly in plant diseases and pest damage detection. It is not always necessary to classify the images, but sometimes the severity of disease or pest damage present in leaves, and sometime the exact location with multiple objects present, also plays crucial role for farmers to apply appropriate remedial measures. There are many object detection techniques out of which we selected the YOLOv10 model due to its consistent dual assignments that eliminate the need for Non-Maximum Suppression (NMS) [16]. Fig. 2 below provides an architectural block diagram of the YOLOv10 model used in our insect damage detection mechanism. Previous versions of YOLO encountered issues with redundant or overlapping bounding box detections. However, YOLOv10 addresses this challenge through the implementation of consistent dual assignments, which significantly reduces the reliance on Non-Maximum Suppression (NMS). The input image is first passed to YOLOv10's backbone, responsible for feature extraction. In YOLOv10, the backbone is based on the Cross-Stage Partial Network (CSPNet), which divides the feature maps into two parts. One part is processed through the network as usual, while the other is handled through a separate path. After passing through their respective processing stages, these two parts are merged, which helps lower computational costs and enhance gradient flow [17].

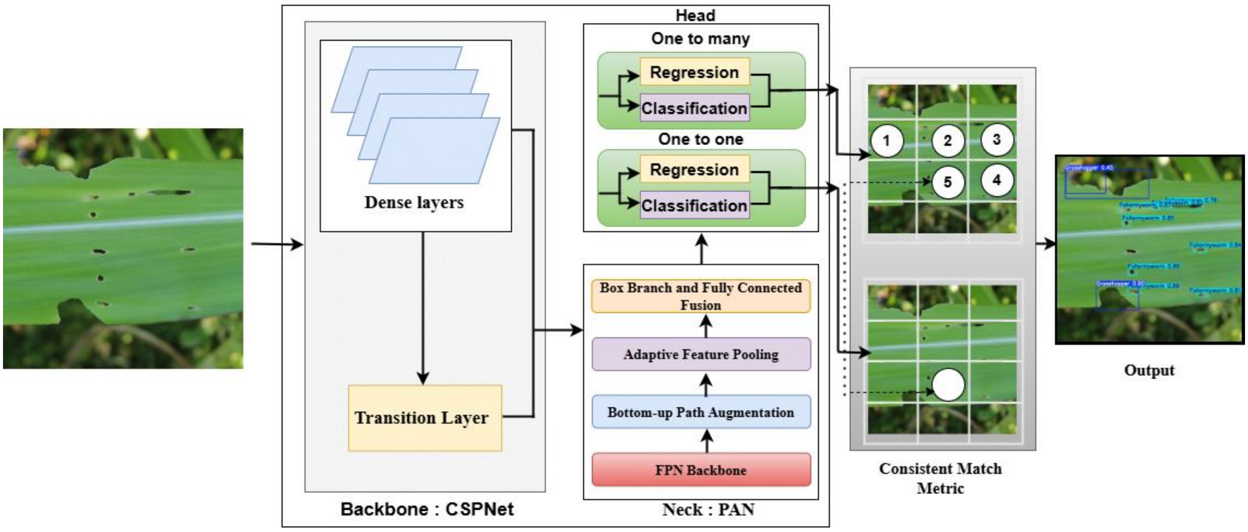

Fig. 2. Architectural block diagram of implemented YOLOv10 model.

These features are passed to the PAN (Path Aggregation Network) in the neck which improves feature aggregation for better object detection. The FPN (Feature Pyramid Network) generates multi-scale features for detecting objects of various sizes. Bottom-up Path Augmentation (BPA) enhances the flow of low-level features to deeper layers whereas Adaptive Feature Pooling (AFP) optimizes pooling by focusing on relevant regions, while the Box Branch predicts bounding box coordinates [18]. Finally, fully connected fusion merge features from multiple layers to improve detection accuracy. The head generates the final predictions from the feature maps by predicting bounding box coordinates, class labels, and confidence scores for each detected object. In YOLOv10, the one-to-one head predicts a single bounding box and class per grid cell, suitable for sparse objects. The one-to-many head uses multiple anchor boxes per grid cell to predict multiple bounding boxes, enabling better detection in complex or crowded scenes. For this research we used all variants of YOLOv10 i.e. YOLOv10n, YOLOv10s, YOLOv10m, YOLOv10l, YOLOv10x and compared their performance metrics for determining best model.

2.3 Hyper-parameters Tuning

Hyper-parameter play an important role in deep learning for object detection analysis [19]. The training hyper-parameters used in this evaluation are provided in Table 2. The AdamW optimiser was used with the batch size 4 for the purpose of optimising models that adopted a default learning rate. This learning rate was used for the training of the models over a fixed number of epochs, which was set at 100. All other Hyper-parameters and image augmentation were used default from the YOLOv10 model. The experiment was run on a 13th generation Intel Core i7-13700F processor, an NVIDIA RTX 3090 GPU (24,576 MiB of memory), and 96 GB of RAM hardware. The software environment included operating system of Ubuntu 20.04.6 LTS, Python 3.11.5 and the PyTorch framework (version: 2.2.0+cu121). The optimal trained model provides an optimal weight and using this best weight model detection for unseen images of insect damage leaves was performed.

Table 2. List of hyper-parameter values for model

Hyper-parameters	Description
Batch size	4
Optimiser	AdamW
Epoch	100
Learning rate	0.001

2.4 Performance Metrics

We conducted various performance evaluations to assess the effectiveness of our model. The selection of evaluation metrics is determined by the specific task and objectives of the machine learning or deep learning application. It is crucial to choose metrics that are suitable for the task and the data under analysis [20-21]. For object detection, the key performance metrics commonly used are:

a) Mean Average Precision (mAP-50; mAP50-95): The mean average precision (at IoU threshold 0.50) is a specific evaluation metric that measures the mean of average precision values. The mean average precision (at IoU threshold 0.50-95) measures the mean of average precision scores across multiple IoU thresholds, from 0.50 to 0.95.

$$mAP50 = \frac{\sum_{i=1}^{N} APi}{N} \quad (1)$$

where, N is the total number of classes.

APi is the average precision for class i.

b) Precision: For object detection model, particularly in terms of the accuracy of predicted bounding boxes, precision quantifies the proportion of correctly predicted bounding boxes (true positives) among all predicted bounding boxes (true positives and false positives).

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

c) Recall: For object detection recall is a metric used to assess a model's capability to accurately identify and detect all relevant objects within an image. It specifically quantifies the ratio of correctly identified objects (true positives) to the total number of actual objects (ground truth) present in the image.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

3. RESULTS AND DISCUSSION

In deep learning, an evaluation metric is a measure used to evaluate a model's performance on a given task. These metrics help quantify how well a model is performing and provide insight into its strengths and weaknesses. The choice of evaluation metrics depends on the specific task and objectives. In this research our objective is to detect the damaged portion of corn leaves by insects. Fig. 3 shows the dual axes curve for the mean average precision with respect to the YOLOv10 model variants. The green line represents mAP50-95 while the blue line resembles mAP50, both following the increasing trend with model strength due to their better feature extraction capabilities. Detecting insect damage in leaf considering evaluation metrics as mAP50 only could be sufficient if the farmer needs just leaf is affected or not. For exact boundaries of the disease the averaging precision across IoU thresholds from 50% to 95% reflects the model's robustness and performance of model's localisation precision. Since there is not much variance in the trend so, the YOLOv10 model shows potential in detecting both types of insect damage present in corn leaves within these datasets.

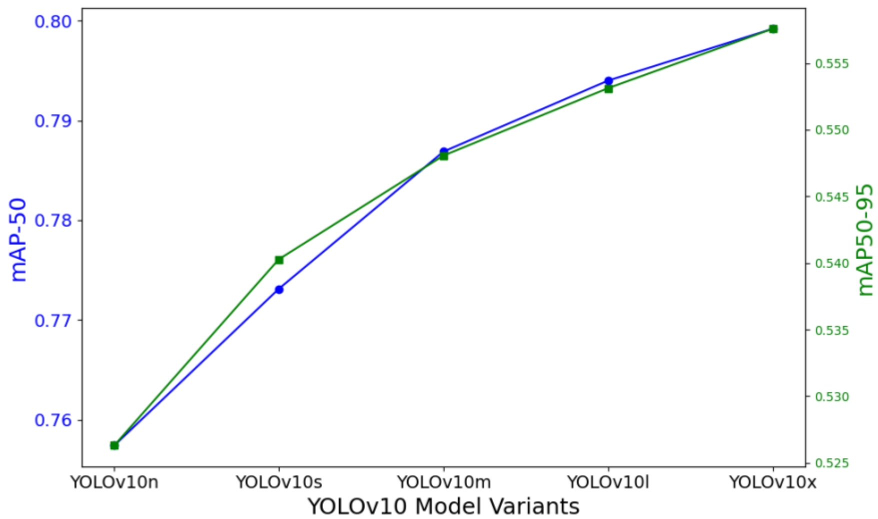

Fig. 3. Dual axes mean average precision plot for different YOLO variants

Fig. 4 is a precision versus recall plots for all variants of YOLOv10 which provides information on how well the model detects as many objects as possible while performing with maximum accuracy. YOLOv10m has the peak value of precision (precision= 0.808) but lower value of recall (recall= 0.807) whereas for YOLOv10x has the maximum value of recall (recall= 0.825) with a lower value of precision (precision= 0.801). The YOLOv10l and YOLOv10x model exhibits a tendency to experience difficulties in maintaining the value of precision at high recall levels, potentially due to a greater number of false positives. This suggests YOLOv10m as a potential saturation point in model scalability. Hence, the selection of such model depends exactly on the task performed. In detecting insect damage, we need a higher accuracy to know the presence and absence of damage through precision and after that the amount of damage could be considered through recall value so, YOLOv10m seems efficient in this insect damage detection task. The result plotted in the graph is the average of five-fold cross validation.

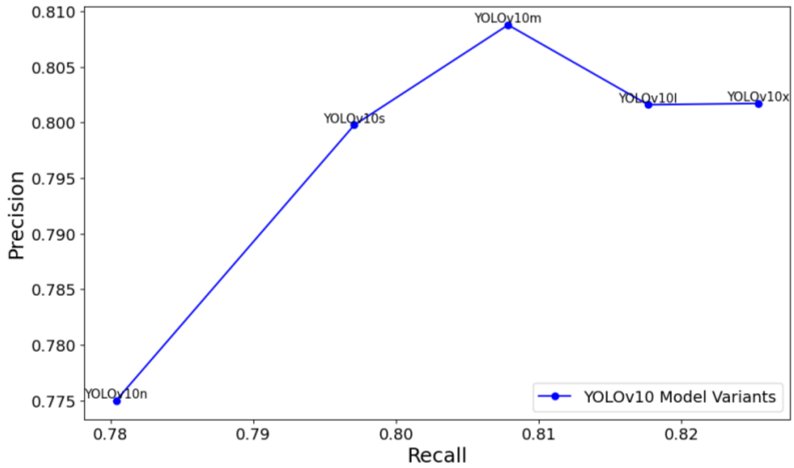

Fig. 4. Precision versus Recall (PR) curve for different YOLO variants

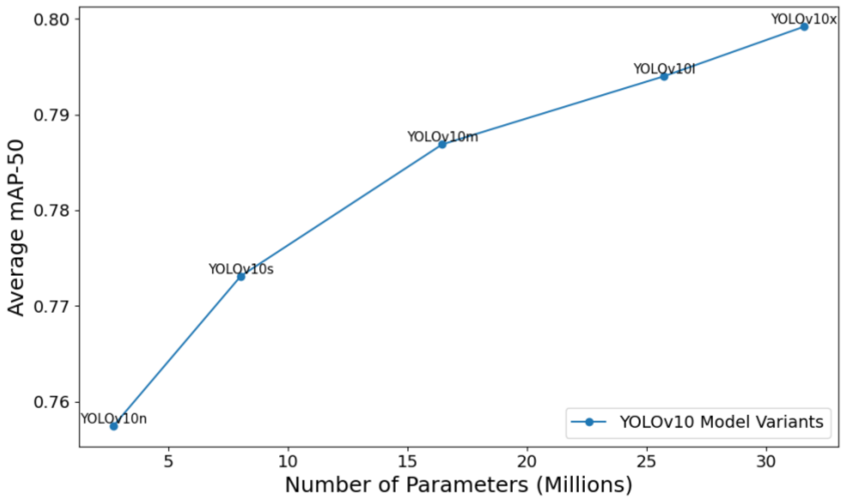

Fig. 5. Mean average precision with respect to number of parameters.

The parameter count provides brief insight into the model's ability to learn and its resource demand which is demonstrated in Fig. 5 with respect to the model's performance in terms of mAP50. Models with greater number of parameters have potential to extract deep features which have ability to increase in the value of mAP50. YOLOv10x has highest average mAP as well as highest number of parameters (no. of parameters= 31.59 millions) among its variants. Developing a lighter model maintains the mAP value could be efficient for deployment in edge devices that could be suitable for farmers. Concurrently, the analogous trend of the graph could be visualised in Figure 6, which demonstrates the latency of the YOLOv10 variants for testing each unseen image and their respective average precision values. YOLOv10n has lowest latency (latency= 27ms) with lowest average precision (precision= 0.757) while YOLOv10x has highest latency (latency= 39ms) with highest average precision (precision= 0.799). The best model among YOLOv10 variants can be chosen considering the trade-offs between accuracy, computational efficiency, and application requirements. Furthermore, lightweight model could be designed for detecting insect damage maintaining the performance [22, 23].

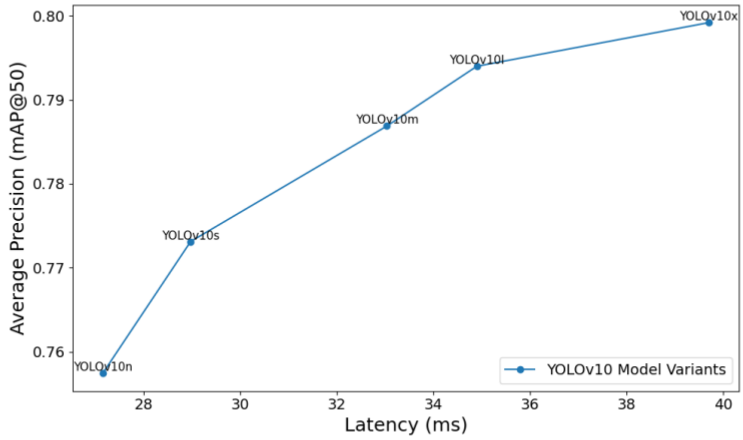

Fig. 6. Mean average precision with respect to latency for object detection.

The distribution of mAP can be better visualised using box plot which provides the statistical summary as in Fig. 7. Among 5 variants, YOLOv10m has highest mean mAP50_95. The mAP50_95 is symmetrically distributed around the mean indicating consistent performance across the 5 folds cross-validation. Hence, YOLOv10m is performing well with insect damage detection dataset.

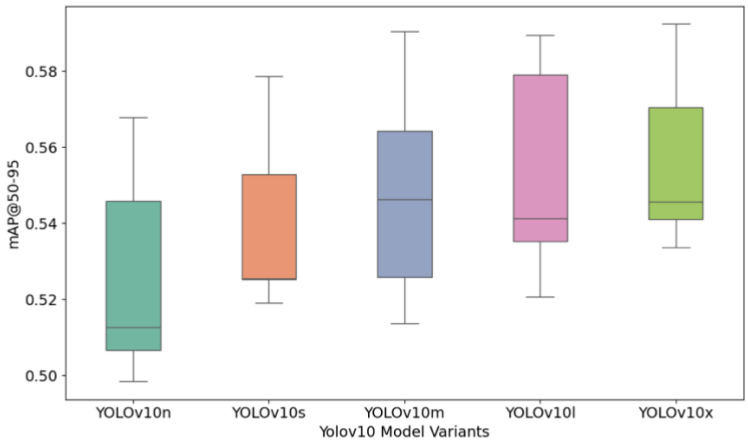

Fig. 7. Box plot representing Inter Quartile Range of mAP for all YOLOv10 variants.

Fig. 8 provides a sample result of unseen corn leaf image with bounding box detected when the test is applied with best weight model. Overall, all models can detect the insect damage with high accuracy. Fig. 8(a). is a test image for YOLOv10n model with object detected with high confidence for Fall armyworm damage and in the case of Grasshopper damage the detection is with low confidence and redundant bounding box is also observed. Fig. 8(b). is a test image for YOLOv10s model with Grasshopper damage detection with lower confidence value than YOLOv10n while Fall armyworm damage detection with higher confidence value than YOLOv10s. Fig. 8(c). is a test image for YOLOv10m which shows the model can detect Grasshopper and Fall armyworm damage with high confidence. Fig. 8(d). is test image for YOLOv10l indicating high detection capabilities for both type of insect damage but there exists a redundant bounding box for Grasshopper damage detection. Fig. 8(e). is test image for YOLOv10x which can detect Fall armyworm damage detection with high confidence while for Grasshopper it is not able to detect damage with high confidence. Considering these detection results on test image we can generalise that YOLOv10m model could be more generalised model among all model though YOLOv10x has provided high value of mAP values during training and validating a model.

(a). Test image with confidence score for YOLOv10n (b). Test image with confidence score for YOLOv10s (c). Test image with confidence score for YOLOv10m

(d). Test image with confidence score for YOLOv10l (e). Test image with confidence score for YOLOv10x

Fig. 8. Test image results with confidence score for all variants of YOLOv10

4. CONCLUSIONS

Timely identification of insects' pest on corn leaf is paramount for increasing the agricultural production. In this research, among the different YOLOv10 architectures implemented for the automated localization and detection of Grasshopper and Fall armyworm YOLOv10l and YOLOv10x has 0.794 and 0.799 values of mAP50 indicating highest among all other variants which could be beneficial with high resource availability. But the PR curve for YOLOv10m exhibits the balancing tendency for highest precision with increasing recall compared to YOLOv10x and YOLOv10l. YOLOv10m has symmetrically distributed mAP50_95 around the mean indicating consistent performance across the five-folds. This nature could potentially make YOLOv10m suitable among all other variants for detecting insect damage in corn leaves images with resource constraints.

Various detection architectures were popular for the pest recognition for instance, Faster R-CNN and R-CNN; however, these architectures are computationally intensive, making them unsuitable for the implementation on edge devices. Hence, in the future these resource intensive architectures could be optimised to lightweight models particularly with focus on maintaining the accuracy of pest identification for farmers with resource intensive model design. Moreover, the severity of pest damage in corn fields analysed through drone images and implementation of DL techniques could be potential future research directions.

ACKNOWLEDGEMENTS

We extend our appreciation for the assistance supported through the Project Based Learning-AI (PBL AI) program and the Ministry of Education, Culture, Sports, Science, and Technology (Monbukagakusho: MEXT) scholarship awarded to Suresh Timilsina and Sandhya Sharma. We are deeply grateful to Anup Neupane, Sarmila Gautam, Ashok Adhikari, Samit Poudel, Santosh Panth and Arjun Awasthi for their invaluable assistance in data collection in the field.

REFERENCES

[1] FAO. (2021). *World Food and Agriculture – Statistical Yearbook 2021*. Rome.

[2] Rouf, S.T., Prasad, K., Kumar, P., & Yildiz, F. (2016). Maize- A potential source of human nutrition and health: A review. Cogent Food & Agriculture, 2. https://doi.org/10.1080/23311932.2016.1166995

[3] Sapna, A.J., Dar, Z.A., & Qureshi, M.I. (2020). Maize fabric: leading to a new paradigm shift. International Journal of Current Microbiology and Applied Sciences, 9, 1341-1343. https://doi.org/10.20546/ijcmas.2020.905.150

[4] Savary, S., Willocquet, L., Pethybridge, S.J. et al. The global burden of pathogens and pests on major food crops. Nat Ecol Evol 3, 430–439 (2019). https://doi.org/10.1038/s41559-018-0793-y

[5] Crop Protection Network. (2021). Corn Disease Loss Estimates from the United States and Ontario, Canada. Crop Protection Network. Available at: https://cropprotectionnetwork.org/publications/corn-disease-loss-estimates-from-the-united-states-and-ontario-canada-2021 (Accessed: 13 February 2025).

[6] Mueller, D.S., Wise, K.A., Sisson, A.J., Allen, T.W., Bergstrom, G.C., Bissonnette, K.M., Bradley, C.A., Byamukama, E., Chilvers, M.I., Collins, A.A., Esker, P.D., Faske, T.R., Friskop, A.J., Hagan, A.K., Heiniger, R.W., Hollier, C.A., Isakeit, T., Jackson-Ziems, T.A., Jardine, D.J., et al. (2020). Corn yield loss estimates due to diseases in the United States and Ontario, Canada, from 2016 to 2019. Plant Health Progress, 21(4), pp. 238-247. https://doi.org/10.1094/PHP-05-20-0038-RS (Accessed: 13 February 2025).

[7] Reisig, D., Kesheimer, K., Bateman, N., Studebaker, G., Meyer, R., Reay-Jones, F., Wise, K., Owens, D., Buntin, G. D., Smith, J., Seiter, N., Hodgson, E., Sisson, A., Zukoff, A., Villanueva, R., Towles, T., Hamby, K., Cramer, M., DiFonzo, C., . . . Conley, S. (2023). Corn Invertebrate Loss Estimates from the United States and Ontario, Canada — 2022. https://doi.org/10.31274/cpn-20230421-0

[8] Chakrabarty, S., Shashank, P. R., Deb, C. K., Haque, M. A., Thakur, P., Kamil, D., Marwaha, S., & Dhillon, M. K. (2024b). Deep Learning-Based accurate detection of insects and damage in cruciferous crops using YOLOV5. Smart Agricultural Technology, 100663. https://doi.org/10.1016/j.atech.2024.100663

[9] Furuya, D. E. G., Ma, L., Pinheiro, M. M. F., Gomes, F. D. G., Gonçalvez, W. N., Marcato, J., Junior, De Castro Rodrigues, D., Blassioli-Moraes, M. C., Michereff, M. F. F., Borges, M., Alaumann, R. A., Ferreira, E. J., Osco, L. P., Ramos, A. P. M., Li, J., & De Castro Jorge, L. A. (2021). Prediction of insect-herbivory-damage and insect-type attack in maize plants using hyperspectral data. International Journal of Applied Earth Observation and Geoinformation, 105, 102608. https://doi.org/10.1016/j.jag.2021.102608

[10] Lee, S., Yun, C.M. A deep learning model for predicting risks of crop pests and diseases from sequential environmental data. Plant Methods 19, 145 (2023). https://doi.org/10.1186/s13007-023-01122-x

[11] Uygun, T., Özgüven, M. M., & Yanar, D. (2024). Detection of Damage Caused by Some Vineyard Pests with the YOLOv8x Model Using Deep Learning and Object Detection Methods. Turkish Journal of Agriculture - Food Science and Technology, 12(3), 386–396. https://doi.org/10.24925/turjaf.v12i3.386-396.6510

[12] Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2022). Explainable deep convolutional neural networks for insect pest recognition. Journal of Cleaner Production, 371, 133638. https://doi.org/10.1016/j.jclepro.2022.133638

[13] Kundur, N. C., & Mallikarjuna, P. B. (2022). Insect Pest Image Detection and Classification using Deep Learning. International Journal of Advanced Computer Science and Applications, 13(9). https://doi.org/10.14569/ijacsa.2022.0130947

[14] Venkateswara, S., Padmanabhan, J. Deep learning based agricultural pest monitoring and classification. Sci Rep 15, 8684 (2025). https://doi.org/10.1038/s41598-025-92659-5

[15] Zhao, S., Liu, J., Bai, Z., Hu, C., & Jin, Y. (2022). Crop pest recognition in real agricultural environment using convolutional neural networks by a parallel attention mechanism. Frontiers in Plant Science, 13. https://doi.org/10.3389/fpls.2022.839572

[16] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOV10: Real-Time End-to-End Object Detection. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2405.14458

[17] Wang, C., Liao, H. M., Yeh, I., Wu, Y., Chen, P., & Hsieh, J. (2019). CSPNet: A New Backbone that can Enhance Learning Capability of CNN. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1911.11929

[18] Lin, T., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2016). Feature pyramid networks for object detection. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1612.03144

[19] Ultralytics. (2025, March 17). Hyperparameter tuning. Ultralytics YOLO Docs. https://docs.ultralytics.com/guides/hyperparameter-tuning/

[20] R. Padilla, S. L. Netto and E. A. B. da Silva (2020). "A Survey on Performance Metrics for Object-Detection Algorithms," *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, pp. 237-242, https://doi.org/10.1109/IWSSIP48289.2020.9145130

[21] Timilsina, S., Sharma, S., & Kondo, S. (2025). Advancements in maize leaf disease detection, segmentation and classification: A review. Biosystems Engineering, 255, 104162. https://doi.org/10.1016/j.biosystemseng.2025.104162

[22] Jiang, T., Du, X., Zhang, N., Sun, X., Li, X., Tian, S., & Liang, Q. (2024). YOLOv8-GO: A Lightweight Model for Prompt Detection of Foliar Maize Diseases. Applied Sciences, 14(21), 10004. https://doi.org/10.3390/app142110004

[23] Timilsina, S., Sharma, S., Gnawali, S., Sato, K., Okada, Y., Watanabe, S., & Kondo, S. (2025). CNDD-Net: A Lightweight Attention-Based Convolutional Neural Network for Classifying Corn Nutritional Deficiencies and Leaf Diseases. *Electronics*, *14*(7), 1482. https://doi.org/10.3390/electronics14071482