Original Research Article

Deep learning-based multi-class pest and disease detection in agricultural fields

ABSTRACT :

Farmers and agricultural workers would manually inspect crops for signs of pests or use traps to monitor pest populations. The advent of deep learning algorithms such as vision transformers and FastAIResNet has brought about a significant transformation in pest detection practices. These advanced algorithms leverage the capabilities of artificial intelligence to process vast amounts of data and learn intricate patterns associated with different pest species and their impact on crops. Unlike manual methods, deep learning algorithms can analyze large datasets quickly and accurately, leading to more efficient and effective pest detection. Vision transformers and FastAIResNet stand out for their ability to continuously learn and adapt to new data, including changes in pest populations over time. This adaptability is crucial in agriculture, where pest dynamics can vary due to factors like climate conditions, environmental changes, and pest control interventions. FastAI ResNet-50 and Vision Transformers have demonstrated remarkable accuracy in classifying various disease classes, indicating their reliability and precision in detecting different pests and diseases affecting crops. Their high accuracies, ranging from 0.95 to 1.00, underscore their effectiveness in agricultural pest detection tasks. However, the study highlights challenges that arise when dealing with more classes in a classification task. Factors such as increased complexity, imbalanced data distributions, and higher-dimensional feature spaces can impact model accuracy. To address these challenges, the study recommends various strategies, including data augmentation, class balancing, robust model architectures, regularization techniques, and transfer learning. Implementing these strategies can help maintain or improve accuracy levels, ensuring that deep learning models remain effective and reliable for agricultural pest detection and disease management applications.

Keywords: Algorithms, Deep learning, Disease and Pest classification, ResNet, Vision transformer

1. INTRODUCTION

Crop production faces various challenges hindering plant growth and yield, including biotic and abiotic stresses caused by living organisms and external environmental variables (Anami et al.,2020). These factors, exacerbated by climate change, can significantly reduce agricultural productivity, posing economic risks to farmers and threatening global food security. Managing these stressors effectively has become increasingly urgent, as a global loss in food production due to pest and disease outbreaks can have profound implications. Plant diseases alone are responsible for 20–40% of crop yield losses (FAO, 2021), impacting the agriculture industry on a large scale and leading to environmental losses in both the quantity and quality of agricultural output. Thus, early-stage plant pest control is critical for sustainable agriculture and environmental stewardship.

AI and machine learning techniques can be used to automate pest identification, overcoming these barriers. Machine learning methods play a crucial role in the process, as they enable extracting relevant features from segmented images for classification purposes. To extract textural information from images, techniques such as Local Binary Patterns (LBP), Grey Level Co-occurrence Matrix (GLCM), and Spatial Grey Level Dependence Matrix (SGLDM) (Ngugi et al.,2021) have been utilized. These methods aid in accurate classification of pests and diseases. An approach to automating disease diagnosis on potato plants on a massive scale using feature extraction followed by Support Vector Machine classification (Islam *et al.,* 2017) achieved high accuracy. An Ensemble of Support Vector Machines and Artificial Neural Network Classifiers for grape leaf disease recognition attained higher accuracy (Padol and Sawant, 2016) compared to single

UNDER PEER REVIEW

classifier systems. Various other Machine Learning Classifiers such as KNN (k-Nearest Neighbour), DT (Decision Trees), RF (Random Forest), ANN (Artificial Neural Network), *etc.,* have been widely used in various pest identification studies (Qin et al., 2016; Krithika and Selvarani, 2017; Sabrol and Satish, 2016).

Researchers have explored various Machine Learning (ML) approaches such as Support Vector Machines (SVM) (Hou et al.,2021 and Hamdani et al., 2021), Artificial Neural Networks (ANNs) (Ramesh and Vydeki, 2020), Naive Bayes (Abdu et al., 2020), k-means clustering (Johannes et al., 2017), and simple linear iterative clustering (Sun et al., 2019). However, recent years have seen a shift towards Deep Learning (DL) due to increased data availability, computing power, and efficient training methods. Convolutional Neural Network (CNN) architectures, particularly those with attention mechanisms, have shown remarkable performance in plant disease detection, alongside standard models like AlexNet, GoogleNet, VGG16, and ResNet used with transfer learning (Mohanty et al., 2016). Customized CNN architectures have also been developed for specific plant disease detection tasks (Yadav et al., 2021), reflecting the ongoing evolution of DL methods in this domain. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have shown remarkable success in image-based pest detection (Panchal et al., 2021; Ozguven and Adem, 2019). The authors identified 13 different types of rice pests and illnesses using a pre-trained CaffeNet model, which had an accuracy rate of 87% (Alfarisy et al., 2018). A total of 26 different diseases in 14 different crop species were identified using GoogLeNet and AlexNet algorithms (Mohanty et al., 2016). Pretrained CNNs of different depths such as VGG16, VGG19, Inception-v3, and ResNet50 were used and fine-tuned to increase the performance in applications with small datasets. The VGG16 model performed the best with an accuracy of 90.4% (Wang et al., 2017).

Pre-designed architectures like ResNet, ViT, *etc*., are prevalent for computer vision tasks and have been used in agricultural applications (Reedha et al., 2022). These models usually have lots of trainable parameters, requiring a large dataset to find the optimal values for their parameters. Transformer is a novel attention structure that was initially used with a Natural Language Processing (NLP) technique to extract the semantics and global features of the context (Wang et al., 2022). These algorithms can analyze visual patterns within images with high accuracy, making them ideal for identifying subtle differences indicative of pest infestations or disease symptoms. Integrating deep learning algorithms into the mobile application enhances its pest and disease detection capabilities in crops. Keeping these points in view, the development of an AI-based crop pest detection mobile application represents a significant advancement in agricultural technology. By leveraging AI, machine learning, and deep learning techniques, this tool has the potential to revolutionize pest management practices, improve crop yields, reduce environmental impact, and ultimately contribute to global food security.

2. MATERIAL AND METHODS

Images of infected rice crops were taken to capture the images. These images are processed using image processing methods to perform automated classification and recognition tasks based on the patterns seen in the images. The step wise procedures adopted are briefed below:

1. Data Collection - Gather a diverse dataset that includes images of healthy plants, plants with various pests (like insects or mites), or plants affected by different diseases (such as fungal infections or viral diseases). Ensure that the dataset covers a wide range of plant species and conditions.

2. Data Pre-processing - Resize all images to a standard size suitable for the chosen deep learning models. Common sizes are 224x224 or 299x299 pixels for models like ResNet, Inception, and MobileNet. Normalize the pixel values of the images to a range between 0 and 1 by dividing the pixel values by 255.

3. Data Augmentation - Apply data augmentation techniques to increase the diversity of the dataset and help the model generalize better. Augmentation methods include random rotation, scaling, flipping, brightness adjustments, and adding random noise to the images.

4. Split Data into Training, Validation, and Test Sets - Divide the dataset into three sets: a training set (used for model training), a validation set (used for hyperparameter tuning and model evaluation during training), and a test set (used to evaluate the final model's performance). Typically, the split ratio is around 70-80% for training, 10-15% for validation, and 10-15% for testing.

5. Choose a Deep Learning Framework - Select a deep learning framework based on your familiarity and the specific requirements of your project. FastAI is a high-level API built on PyTorch that simplifies deep learning tasks.

6. Model Selection - Choose one or more deep learning models suitable for image classification tasks. For example:

UNDER PEER REVIEW

- ResNet-34, ResNet-50: Good balance between accuracy and computational complexity.
- MobileNetv3: Efficient for mobile and edge devices.
- InceptionV3: Known for handling complex patterns in images.
- SqueezeNet1_0: Lightweight model suitable for resource-constrained environments.
- DenseNet121: Dense connectivity structure beneficial for feature reuse.
- Vision Transformers (ViT): Effective for capturing long-range dependencies in images.

7. Model Training - Initialize the selected model(s) with pre-trained weights on large datasets to leverage learned features. Fine-tune the model(s) using the training set. Adjust hyperparameters such as learning rate, batch size, and optimizer choice to optimize model performance. Use techniques like learning rate scheduling and early stopping to prevent overfitting and improve convergence speed.

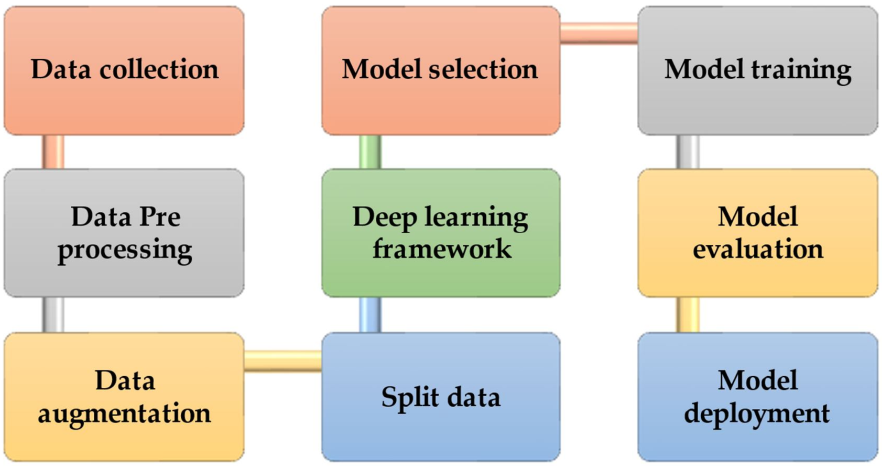

Figure 1. Methodology for Data Collection and Processing

8. Model Evaluation - Evaluate the trained model(s) on the validation set to monitor performance metrics such as accuracy, precision, recall, and F1 score (Pedregosa et al., 2011). Use tools like confusion matrices and ROC curves for detailed analysis. Perform model tuning based on validation results, such as adjusting regularization techniques or model architecture modifications.

$$\text{Accuracy} = \frac{TN + TP}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Precision} + \text{Recall}} \quad (4)$$

Where TP represents True Positive samples, TN represents True Negative samples, FP represents False Positive samples, and FN represents False Negative samples. These metrics are commonly used to gauge the performance of classification models.

9. Model Deployment - Once satisfied with the model's performance on the validation set, evaluate it on the test set to assess its generalization ability. Deploy the trained model in a production environment, integrating it into an application or system for real-time pest and disease identification. Implement monitoring and feedback mechanisms to continuously improve the model's performance based on user feedback and new data. Regularly update the model with new data to adapt to evolving pest and disease patterns. Explore advanced techniques like ensemble learning (combining multiple models for improved accuracy) and model distillation (compressing large models for deployment on resource-constrained devices) for further enhancement.

UNDER PEER REVIEW

3. RESULTS

This study utilized cutting-edge deep learning algorithms such as ResNet, Vision Transformer (ViT), *etc*. instead of the conventional machine learning models employed in pest detection. Rice plant diseases (Rice blast, Dead heart, Bacterial leaf blight) and nutritional disorders (Boron Deficiency, Brown Spot, False Smut, Leaf Folder, Sheath Blight, Sheath Rot, White Ear, and Zinc Deficiency) were identified using deep learning algorithms (Fig. 2)*viz.,* ResNet-34, MobileNetv3, ResNet-50, InceptionV3, squeezenet1_0, densenet121, FastAI and Vision Transformers (ViT).

Blast **False smut** **Leaf folder** **Zn deficiency**

Fig. 2. Pest and Disease

The performance of the algorithms *viz.,* ResNet-34, ResNet-50, and MobileNetV3 was measured for three disease class identification. The results showed that ResNet-34, ResNet-50, and MobileNetV3 (Table 1) achieved an accuracy range of 0.96-0.99, 0.38-0.64, and 0.94-0.98, respectively. For four disease class identification (Table 2), ResNet-34, MobileNetV3, InceptionV3, squeezenet1_0, and densenet121 performed with an accuracy ranging from 0.94 to 0.98, 0.74 to 0.99, 0.96 to 0.98, 0.59 to 0.90 and 0.55 to 0.97, respectively.

Table 1. Accuracy assessment for identifying three disease classes

Model	ResNet-34	MobileNetV3	ResNet-50
Blast	0.96	0.64	0.97
Dead Heart	0.99	0.78	0.98
Bacterial Leaf Blight	0.98	0.38	0.94

Table 2. Accuracy assessment for identifying four disease classes

Model	ResNet-34	MobileNetV3	InceptionV3	Squeezenet1_0	Densenet121
Blast	0.98	0.94	0.98	0.89	0.96
Dead Heart	0.98	0.95	0.97	0.59	0.55
Bacterial Leaf Blight	0.94	0.74	0.96	0.9	0.97
Others	0.98	0.99	0.98	-	-

FastAI ResNet-50 and Vision Transformers (ViT) performed well in classifying eleven disease classes, with an accuracy ranging from 0.95 to 1.00 (Figure 3). Both models had a proven architecture with pre-trained weights for efficient training and deployment. They excel in complex pattern recognition and scalability but may require additional resources for training and deployment. The model testing accuracy ranged from 97 to 100 *per cent* and 85 to 98 *per cent* for 6 and 11 disease class identification (Table 3).

UNDER PEER REVIEW

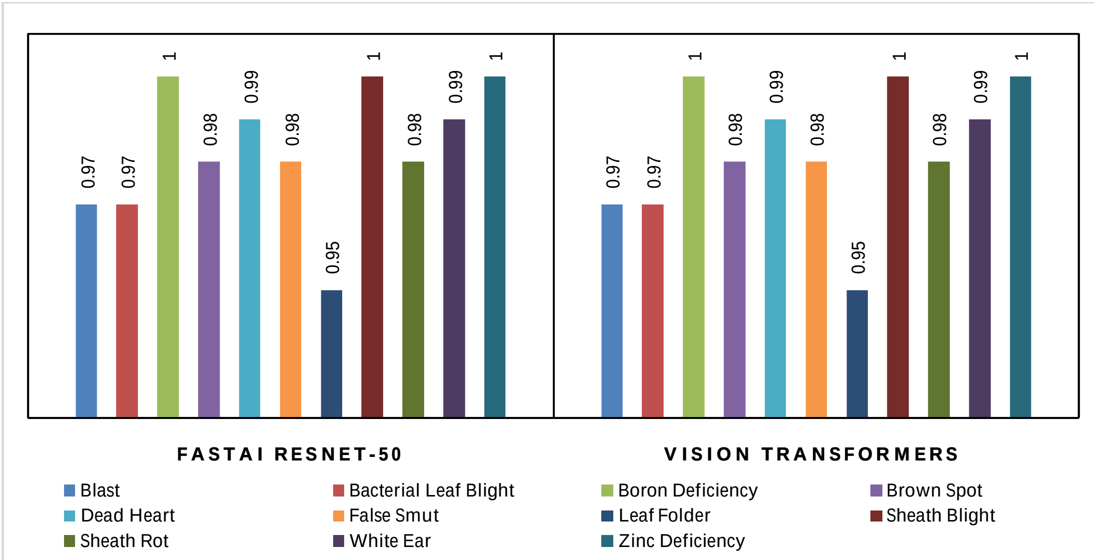

Figure 3. Accuracy assessment for identifying eleven disease classes

Table 3. Model testing accuracy for 6 and 11 disease classes

6 and 11 Classification Model Testing Accuracy			
Blast	98%	Bacterial Leaf Blight	85%
Brown Spot	98%	Blast	87%
Dead Heart	100%	Boron Deficiency	86%
False Smut	98%	Brown Spot	95%
Leaf Folder	97%	Dead Heart	94%
White Ear	98%	False Smut	98%
		Leaf Folder	89%
		Sheath Blight	93%
		Sheath Rot	90%
		White Ear	90%
		Zinc Deficiency	96%

4. DISCUSSION

For medium-sized datasets like 3 or 4 class disease identification, ResNet-34 outperformed other models. It was designed for mobile and edge devices, offering efficiency without significant loss in accuracy. The inclusion of more classes in a classification task makes it harder for the model to distinguish between them accurately. This complexity and issues like imbalanced data and higher-dimensional feature spaces can decrease accuracy. To maintain or improve accuracy with more classes, use strategies like data augmentation, class balancing, robust model architectures, regularization, and transfer learning. Fast AI ResNet-50 and Vision Transformers (ViT) are two advanced deep-learning architectures used in computer vision tasks. ResNet-50 has been widely used in various computer vision tasks and has shown strong performance in image classification, including plant disease classification (Zhao et al., 2022). Vision Transformers uses transformer-based architectures, processes images in patches, and is highly effective in capturing global context (Chen *et al.,* 2021). ResNet-50 is a robust choice for deep feature representation and transfer learning, while ViT excels in global context understanding and scalability. Their selection depends on task requirements, dataset characteristics, and the desired balance between model complexity, accuracy, and efficiency.

Ramalingam et al. (2020) deployed an IoT-based system to detect pests using a faster RCNN ResNet50 model. They tested the model with 150 images per class from the IP102 dataset, achieving an average accuracy of around 94% for the eight insect classes. The Vision Transformer (ViT) (Dosovitskiy et al., 2020) is a novel neural network architecture for image processing. It outperforms ResNet152 (He *et al.,* 2016) on datasets like ImageNet and CIFAR-100 and also surpasses Noisy Students (EfficientNet-L2). Li et al. (2023) proposed a computationally efficient deep learning

UNDER PEER REVIEW

architecture for real-time plant disease identification. The architecture is based on Mobile Vision Transformer (MobileViT) and is called the PMVT network. The results showed that PMVT network outperformed VGG-19, GoogLeNet, and ResNet-50, which are state-of-the-art CNN architectures, on both PlantVillage and Maize disease datasets. Boukabouya et al. (2022) reported that vision transformer-based models achieved top classification accuracies on various plant disease datasets. Specifically, the accuracies were 96.7%, 98.52%, 99.1% and 99.7%. Borhani et al. (2022) compared the performance of classical CNN methods, vision transformers, and their combination for plant disease classification. The results showed that while attention blocks in ViT increased accuracy, they also slowed down the prediction speed. However, combining CNN and ViT blocks could address this speed issue.

5. CONCLUSION

FastAI ResNet-50 and Vision Transformers demonstrated exceptional accuracy, ranging from 0.95 to 1.00, in effectively classifying eleven disease classes. Specifically, they achieved higher testing accuracies of 97 to 100 *per cent* for identifying six classes and 85 to 98 *per cent* for identifying eleven classes. This performance underscores their robustness and reliability in agricultural pest detection tasks. Traditionally, pest detection in agriculture has heavily relied on labor-intensive manual inspection methods or trapping techniques. However, this project introduced a significant shift by leveraging state-of-the-art deep learning algorithms like FastAIResNet and Vision Transformers. These advanced models not only outperform conventional methods but also continuously learn and adapt to evolving pest varieties. By harnessing the power of deep learning, particularly in tasks such as classification and object detection, these models offer a promising avenue for accurate pest identification. They excel in analyzing images or videos with labeled training data, making them valuable tools in addressing pest and disease challenges in crops. The strategic adoption of these cutting-edge techniques aligns seamlessly with the project's objectives, signaling a potentially transformative impact on agriculture. This approach not only enhances accuracy but also streamlines processes, ultimately contributing to more efficient and sustainable agricultural practices.

DEFINITIONS, ACRONYMS, ABBREVIATIONS

ANN	:	Artificial Neural Network
CNN	:	Convolutional Neural Network
DL	:	Deep Learning
DT	:	Decision Trees
GLCM	:	Grey Level Co-occurrence Matrix
KNN	:	k-Nearest Neighbour
LBP	:	Local Binary Patterns
NLP	:	Natural Language Processing
RF	:	Random Forest
SGLDM	:	Spatial Grey Level Dependence Matrix
ViT	:	Vision Transformers

REFERENCES

Abdu, A. M., Mokji, M. M., & Sheikh, U. U. (2020). Automatic vegetable disease identification approach using individual lesion features. Computers and Electronics in Agriculture, 176, 105660.

Alfarisy, A. A., Chen, Q., & Guo, M. (2018, April). Deep learning-based classification for paddy pests & disease recognition. In *Proceedings of 2018 International Conference on mathematics and Artificial Intelligence* (pp. 21-25).

Anami, B. S., Malvade, N. N., &Palaiah, S. (2020). Classification of yield affecting biotic and abiotic paddy crop stresses using field images. *Information Processing in Agriculture*, *7*(2), 272-285.

Borhani, Y., Khoramdel, J., & Najafi, E. (2022). A deep learning based approach for automated plant disease classification using vision transformer. *Scientific Reports*, *12*(1), 11554.

Boukabouya, R. A., Moussaoui, A., &Berrimi, M. (2022). Vision Transformer Based Models for Plant Disease Detection and Diagnosis. In *2022 5th International Symposium on Informatics and its Applications (ISIA)* (pp. 1-6). IEEE.

UNDER PEER REVIEW

Chen, X., Hsieh, C. J., & Gong, B. (2021). When vision transformers outperform ResNets without pre-training or strong data augmentations. *arXiv preprint arXiv:2106.01548*.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... &Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

FAO. New standards to curb the global spread of plant pests and diseases. http://www.fao.org/news/story/en/item/1187738/icode/, 2021. Accessed: 2021-09-04.

Hamdani, H., Septiarini, A., Sunyoto, A., Suyanto, S., &Utaminingrum, F. (2021). Detection of oil palm leaf disease based on color histogram and supervised classifier. *Optik*, *245*, 167753.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hou, C., Zhuang, J., Tang, Y., He, Y., Miao, A., Huang, H., & Luo, S. (2021). Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation. *Journal of Agriculture and Food Research*, *5*, 100154.

Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017, April). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th Canadian conference on Electrical and computer engineering (CCECE)* (pp. 1-4). IEEE.

Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A. D., & Ortiz-Barredo, A. (2017). Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. Computers and electronics in agriculture, 138, 200-209.

Krithika, N., & Selvarani, A. G. (2017, March). An individual grape leaf disease identification using leaf skeletons and KNN classification. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1-5). IEEE.

Li, G., Wang, Y., Zhao, Q., Yuan, P., & Chang, B. (2023). PMVT: a lightweight vision transformer for plant disease identification on mobile devices. *Frontiers in Plant Science*, *14*, 1256773.

Mohanty, S. P., Hughes, D. P., &Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 215232.

Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. *Information processing in agriculture*, *8*(1), 27-51.

Ozguven, M. M., & Adem, K. (2019). Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Physica A: statistical mechanics and its applications*, *535*, 122537.

Padol, P. B., & Sawant, S. D. (2016, December). Fusion classification technique used to detect downy and Powdery Mildew grape leaf diseases. In *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)* (pp. 298-301). IEEE.

Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based plant disease detection using deep learning. *Materials Today: Proceedings*, *80*, 3500-3506.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... &Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.

Qin, F., Liu, D., Sun, B., Ruan, L., Ma, Z., & Wang, H. (2016). Identification of alfalfa leaf diseases using image recognition technology. *PloS one*, *11*(12), e0168274.

Ramalingam, B., Mohan, R. E., Pookkuttath, S., Gómez, B. F., Sairam Borusu, C. S. C., Wee Teng, T., & Tamilselvam, Y. K. (2020). Remote insect trap monitoring system using deep learning framework and IoT. *Sensors*, *20*(18), 5280.

Ramesh, S., &Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. Information processing in agriculture, 7(2), 249-260.

Reedha, R., Dericquebourg, E., Canals, R., & Hafiane, A. (2022). Transformer neural network for weed and crop classification of high resolution UAV images. *Remote Sensing*, *14*(3), 592.

Sabrol, H., & Satish, K. (2016, April). Tomato plant disease classification in digital images using classification tree. In *2016 international conference on communication and signal processing (ICCSP)* (pp. 1242-1246). IEEE.

Sun, Y., Jiang, Z., Zhang, L., Dong, W., & Rao, Y. (2019). SLIC_SVM based leaf diseases saliency map extraction of tea plant. Computers and electronics in agriculture, 157, 102-109.

Wang, G., Sun, Y., & Wang, J. (2017). Automatic image-based plant disease severity estimation using deep learning. *Computational intelligence and neuroscience*, *2017*.

Wang, Y., Chen, Y., & Wang, D. (2022). Convolution network enlightened transformer for regional crop disease classification. *Electronics*, *11*(19), 3174.

Yadav, S., Sengar, N., Singh, A., Singh, A., & Dutta, M. K. (2021). Identification of disease using deep learning and evaluation of bacteriosis in peach leaf. Ecological Informatics, 61, 101247.

Zhao, P., Li, C., Rahaman, M. M., Xu, H., Yang, H., Sun, H., ... &Grzegorzek, M. (2022). A comparative study of deep learning classification methods on a small environmental microorganism image dataset (EMDS-6): from convolutional neural networks to visual transformers. *Frontiers in Microbiology*, *13*, 792166.

UNDER PEER REVIEW