Journal of Information Systems Engineering and Management
2025, 10(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Advanced CNN-Based Leaf Disease Detection for Sustainable Agriculture

Tanvi Deshmukh[1*], Anand Singh Rajawat [2], Amol Potgantwar[3]

[1]SOCSE, Sandip University, Nashik, Maharashtra, India. tanvi.panse19@gmail.com

[2]SOCSE, Sandip University, Nashik, Maharashtra, India. anandsingh.rajawat@sandipuniversity.edu.in

[3]SITRC, Sandip Foundation, Nashik, Maharashtra, India. amol.potgantwar@sitrc.org

Corresponding Author: tanvi.panse19@gmail.com

ARTICLE INFO

Received: 02 Oct 2024

Revised: 29 Nov 2024

Accepted: 15 Dec 2024

ABSTRACT

Crop diseases are one of the major challenges to global agricultural productivity, and efficient and accurate solutions for early detection are needed. This paper proposes a CNN-based model for crop leaf disease detection, integrated with the Crop Pro user interface for practical deployment. The model is trained on an extensive dataset that contains 8,264 images of crop leaves belonging to various species, classified into 14 disease categories, with preprocessing and augmentation techniques further improving it. The CNN model had a training accuracy of 94.73% and a validation accuracy of 88.75%, showing robustness and the ability to generalize over different crops and conditions. Users can upload their leaf images to the Crop Pro platform that preprocesses and classifies images and uses them for high-confidence disease predictions. The availability of an iterative feedback loop helps improve the model based on input received from the user. Scalability of the system and application in real-world applications may offer promising solutions for precision agriculture through early intervention for sustainable crop management.

Keywords: Leaf Disease Detection, Convolutional Neural Network, Machine Learning, Smart Farming, Crop Management System

INTRODUCTION

Cutting-edge technologies, including artificial intelligence and machine learning, are now paving the way for innovative solutions to pressing issues in agriculture, particularly in the identification and management of plant diseases. Timely and precise detection of crop diseases is crucial for mitigating yield losses, decreasing reliance on chemical pesticides, and enhancing overall agricultural efficiency. Traditional disease detection is labour-intensive and subjective, based on conventional inspection by human beings and, thus prone to errors (Li, Zhang, & Wang, 2021). The addition of ML techniques with a combination of classical ML algorithms and DL models such as CNNs provides a new alternative that has much promise regarding high accuracy and scalability in its use (Ngugi et al., 2024).

Deep learning techniques have found great success in image-based disease detection by identifying the most complex patterns of plant images and automatically detecting disease symptoms (Ouamane et al., 2024). Recently, it has been proven that CNNs can classify a vast array of diseases in crops such as tomatoes, apples, and rice using very large amounts of images (Panchal et al., 2023; Salih et al., 2020). The most recent advances that bring machine learning together with domain-specific applications, such as Vision Transformers and transfer learning, pave the way for further development of these models in precision agriculture (Hassan et al., 2021; Singh et al., 2024)

While current research illustrates the promise of CNN-based methods for identifying crop diseases, issues such as dataset constraints, model adaptability, and practical implementation remain insufficiently addressed. This research makes a valuable contribution by tackling these issues with an innovative approach that integrates hyperparameter optimization, transfer learning, and sophisticated data augmentation strategies. Our primary goals are to improve model scalability, guarantee accuracy across various crop environments, and incorporate practical functionality through the Crop Pro platform. This strategy presents a scalable, real-time solution designed to foster sustainable agricultural practices, representing a notable leap forward in precision farming technologies.

LITERATURE REVIEW

Machine learning techniques in agriculture have significantly revolutionized the detection of crop diseases. The conventional ML-based approaches like KNN, SVM, and Random Forest work very efficiently with feature-based classification of disease. These rely on handcrafted features of colour, texture, and shapes to create models that classify healthy and diseased plants (Abdu et al., 2020; Hatuwal et al., 2021). Usually, these suffer from scalability and require a great deal of domain knowledge.

Deep learning methods, which focus more on feature representation, best apply to more complex pattern recognition tasks, such as the ones found in plant disease identification with RGB images of leaves (Li, Zhang, & Wang, 2021; Panchal et al., 2023). These models can extract spatial hierarchies from image data utilizing convolutional and pooling layers to correctly classify diseases, such as late blight in tomatoes and bacterial spots in peppers, based on the works of Kim et al. (2023) & Salih et al. (2020).

The major reason for this problem is that effective models for disease detection are created while the availability of diverse datasets in quantity is limited. Transfer learning and data augmentation have been applied for model generalization for disease detection in cases employing pre-trained models (Saleki & Tahmoresnezhad, 2024). For instance, the rotation, crops, and scales of the data sets have been used to create artificial variants that increase the model's potential resistance towards new data (Ding et al., 2024). Transfer learning enables the models to use existing knowledge learned from related domains without requiring massive labelled data (Mohameth, Bingcai, & Sada, 2020).

Such innovations, especially in classification accuracy and computational techniques, have been due to designs of new architectures integrating CNNs with optimization algorithms like the Red Deer Optimization, developed recently by Reddy et al., 2023. Others, like the CNN model, hybridized with a combination of Vision Transformers and Bayesian networks, have proven beneficial across various agriculture settings by enhancing performance (Sachdeva et al., 2021; Singh et al., 2024).

Beyond classification, the practical utility of these models is also applied for severity estimation and real-time monitoring of diseases. Mobile applications are designed in CNNs to allow a farmer to upload images of leaves for quick diagnosis and management recommendations (Rajendra, Rajkumar, & Shetty, 2019). However, the computational requirements, lack of interpretability of the model, and the complexity in integrating environmental data are some of the issues that require further investigation (Arya & Singh, 2019; Tudi et al., 2021).

The existing literature emphasizes the groundbreaking impact of CNNs in the realm of plant disease detection while also revealing significant drawbacks, such as dependence on restricted datasets, computational inefficiencies, and obstacles in real-world implementation. Our study seeks to fill these voids by utilizing a comprehensive and varied dataset while incorporating practical usability through an interactive interface. The integration of ongoing feedback for continuous enhancement and the establishment of an easily accessible platform set our research apart as a unique contribution, effectively connecting cutting-edge technology with real-world agricultural practices.

METHODOLOGY

The following Figure 1 depicts a flow chart of the step-by-step process of the proposed crop leaf disease detection system using a Convolutional Neural Network (CNN). The flowchart starts from dataset collection and then continues with preprocessing and feature extraction. Thereafter, it divides the data into training and test datasets with different processes to train and test the CNN model. The system does leaf disease detection after training and evaluates the model. The final step produces results and accuracy, thus closing the loop.

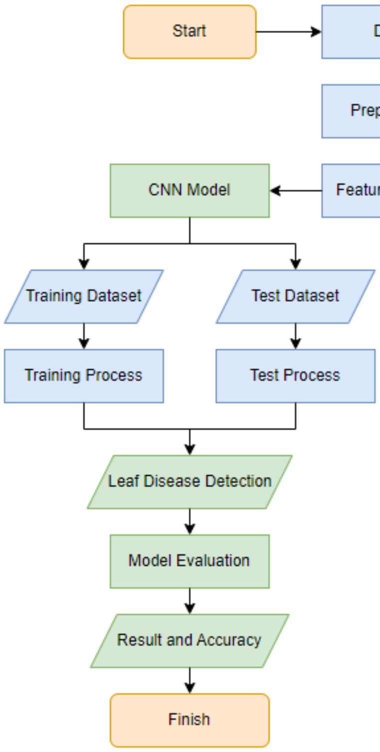

Figure 1. Flowchart of the Proposed Crop Leaf Disease Detection System

The research incorporates a CNN-based model to identify crop leaf diseases, combined with the Crop Pro user interface for the easy and interactive presentation of information to farmers. The methodology involves dataset preparation, preprocessing, model architecture, and evaluation. Figure 2 demonstrates the platform where users upload crop leaf images for disease analysis and receive predictions.

Figure 2. Crop Pro Interface

Dataset and Preprocessing

The dataset contains images of crop leaves that are a total of 8,264, including healthy and diseased leaves from crops like tomatoes, potatoes, rice, cotton, and grapes. There are 14 disease categories, including fungal, bacterial, and viral infections. The dataset is divided into 80% training and 20% validation, with an additional 33 test images used for final evaluation. Figure 3 below illustrates sample images from the dataset, highlighting potato and grape leaves. This figure showcases examples of healthy and diseased leaves, demonstrating the variety and complexity of the dataset used for training and validation.

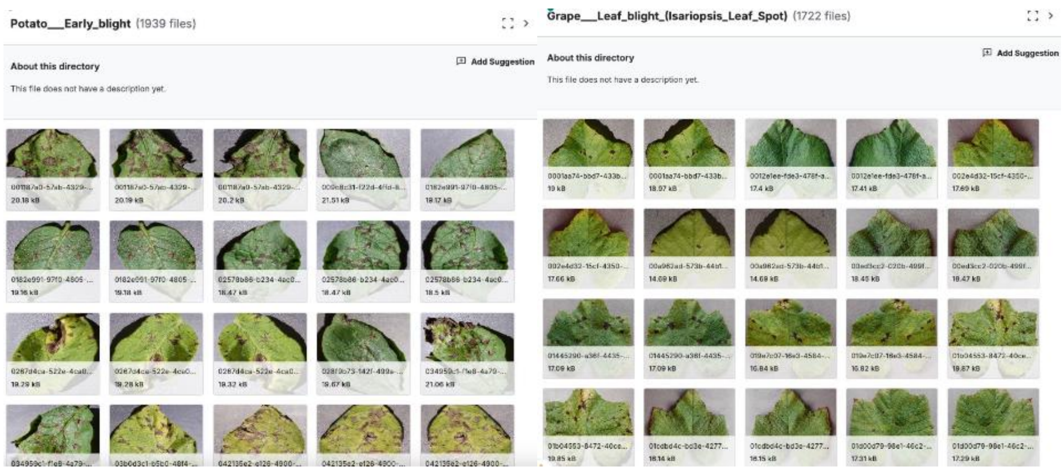

Figure 3. Sample Images of Potato and Grape Leaves in Dataset (Samir Bhattarai, n.d.)

To ensure consistency and improve model generalization, images undergo the following preprocessing steps:

- Resizing: Standardizes image dimensions for uniformity across the dataset.
- Normalization: Scales pixel values to the range [0, 1] to standardize input data.
- Data Augmentation: Includes transformations such as rotation, flipping, and cropping to expand dataset diversity and improve the model's robustness.

Feature Extraction

This crop leaf disease classification CNN uses multiple central layers to make a better prediction. The convolutional layers capture important features such as edges and textures from the images so that the model can notice distinct patterns related to a disease. Pooling layers help to reduce the dimension of the feature maps, decreasing computation but preserving the most useful information. Lastly, the extracted features get fully connected to the output layers with a final Softmax activation function used for classifying probabilities of diseases.

Figure 4 illustrates the layered structure of the CNN model, demonstrating the flow of data through convolutional, pooling, and fully connected layers. This architecture is optimized for identifying disease patterns in input crop leaf images.

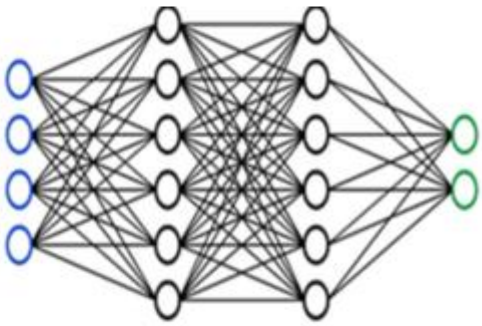

Figure 4. CNN Layers - Convolutional, Pooling, and Fully Connected Layers

Training and Evaluation

The model was trained using the training set and tested on the validation set. It utilized Adam's optimizer with a categorical cross-entropy loss function. Precision, accuracy, recall, and F1 score are the major metrics used to measure the model. The technology also consists of a feedback loop through the Crop Pro interface that enables farmers to affirm or correct predictions to thereby improve the model, continuously.

Disease Detection Process

Crop leaf conditions are classified through the detection procedure, which involves several steps. Crop Pro allows farmers to upload pictures of crop leaves. These pictures are preprocessed, scaled, normalized, and augmented to make them a better fit for analysis using the CNN model. The model looks at the input, extracts features, and classifies while determining whether the leaf is healthy or diseased. It finally gives a prediction with confidence scores, which aids in informed decision-making in crop management.

RESULTS & DISCUSSION

CNN-based crop disease detection model proved to be of high performance, with high accuracy in disease classification. The result confirms that the system has the capability to improve agricultural practices through timely and reliable disease diagnosis.

Model Performance

The Training Accuracy and Validation Accuracy were calculated using the formula:

$$Accuracy = \frac{Current\ predictions}{Total\ predictions} \times 100$$

The Training Loss and Validation Loss were computed using the cross-entropy loss function:

$$Loss = \frac{1}{N}\sum_{i=1}^{N} y_i \log p_i$$

Where y_i represents the actual labels, and p_i denotes the predicted probabilities.

The model was trained over five epochs, achieving the metrics as presented in Table 1:

Table 1: Model Performance Summary

Metric	Value
Training Accuracy	94.73%
Validation Accuracy	88.75%
Training Loss	0.1735
Validation Loss	0.4251

The high training accuracy reflects the model's ability to learn disease-specific features effectively, while the validation accuracy and loss indicate good generalization to unseen data.

The CNN model proposed here has a high training accuracy of 94.73% and a validation accuracy of 88.75%, using an ample dataset of 8,264 images taken from 14 categories of diseases for crop leaves, thus forming a very effective model for detecting diseases in crops. SVM, Random Forest, and KNN models were trained over the different datasets of plant leaves with much lower performances.

Table 2 shows a comparison between accuracy and datasets used by the models in the crop leaf disease detection process.

Table 2: Comparison of Model Accuracy and Data set used for Crop Leaf Disease Detection

Model/Method	Accuracy (%)	Dataset Used	Study
SVM	82%	Various plant leaf datasets	Abdu et al., 2020
Random Forest	85%	Various plant leaf datasets	Hatuwal et al., 2021
KNN	84%	Various plant leaf datasets	Hatuwal et al., 2021
Proposed CNN (Training)	94.73%	8,264 images of crop leaves (14 disease categories)	Current Study

Proposed CNN (Validation)	88.75%	8,264 images of crop leaves (14 disease categories)	Current Study

The bar graph in Figure 5 illustrates a comparison of accuracy among various models used for crop leaf disease detection. The proposed CNN model is marked along with the annotations showing an improvement of +4.73% in the training accuracy and +2.75% in validation accuracy from the best existing model, which is CNN. This graph emphasizes the performance of the proposed CNN model, which is fed on a large and diversified dataset of 8,264 images across 14 categories of diseases, resulting in very considerable improvements over models like SVM, Random Forest, and KNN trained on various plant leaf datasets.

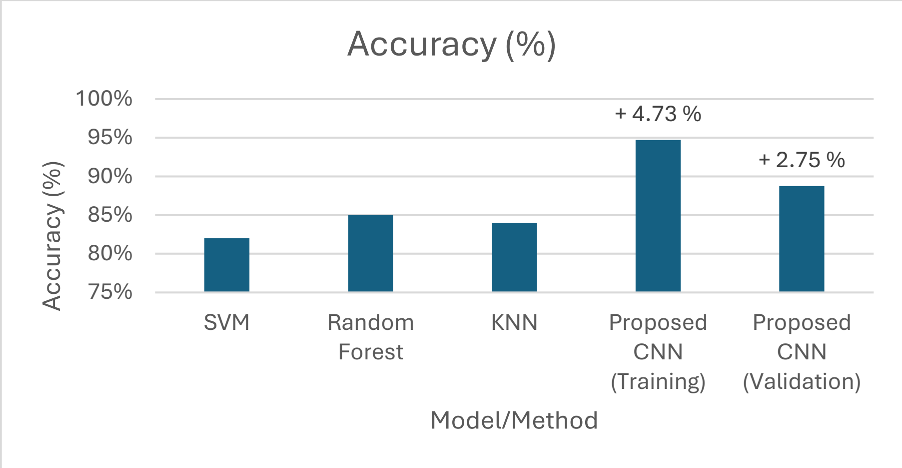

Figure 5. Performance Comparison of Crop Disease Detection Model

Disease Prediction Results

The CNN model demonstrated high accuracy in predicting various crop leaf diseases across the test dataset. Figure 6 displays the result of the model identifying Late Blight, a fungal disease, with high confidence, by recognizing characteristic spots and discoloration on the leaf surface.

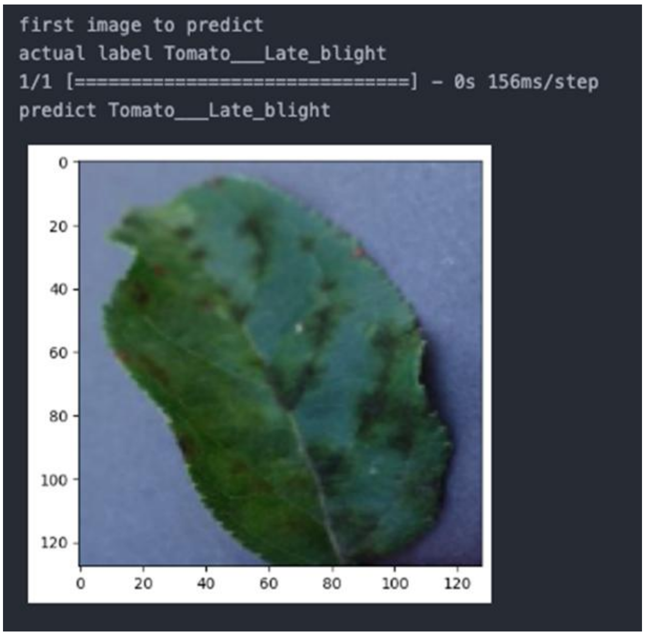

Figure 6. Prediction of Tomato Late Blight

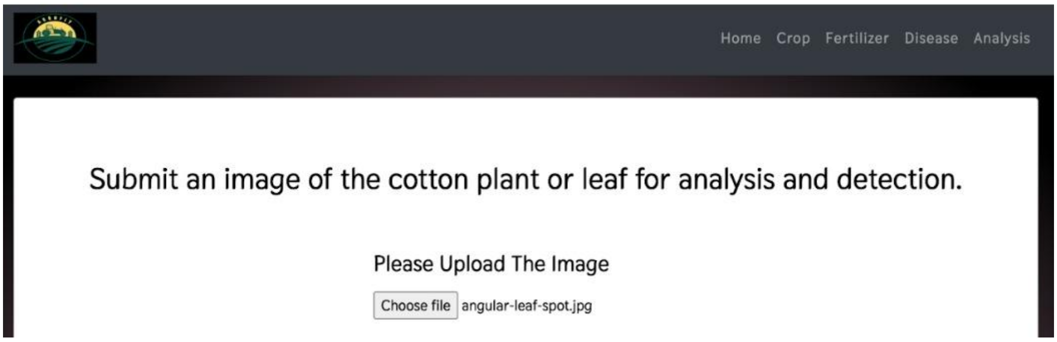

Figure 7a. Image upload

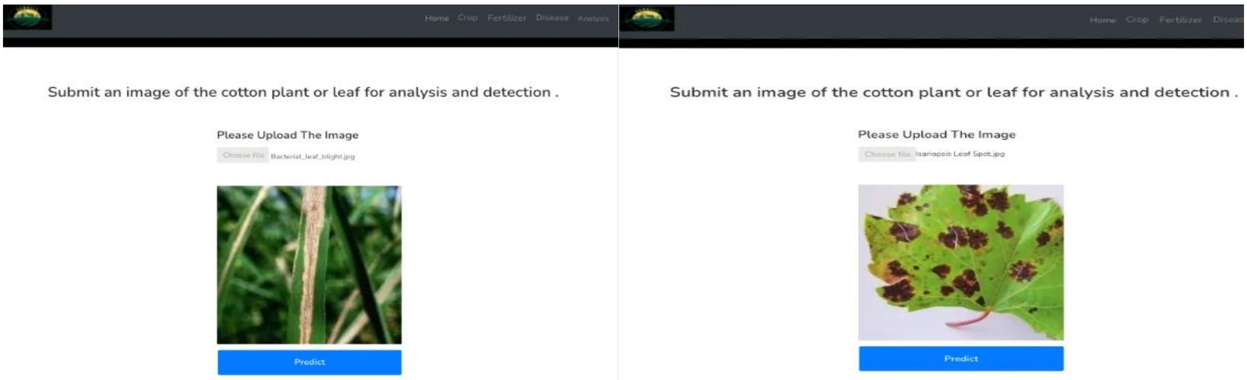

Figure 7b: Click predict

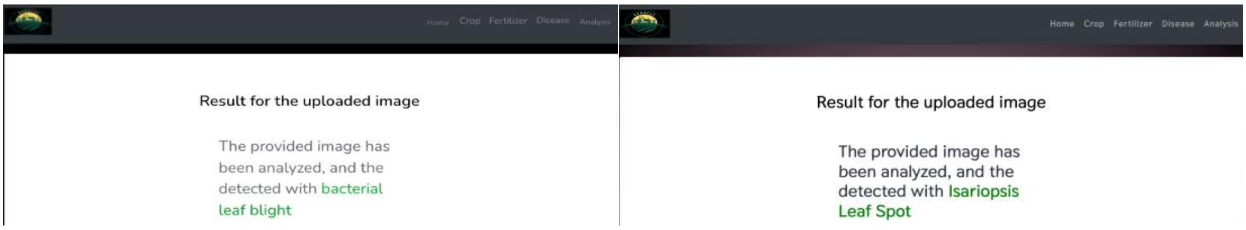

Figure 7c: Leaf Image Prediction

Figure 7 (a, b, and c) illustrates how efficiently the model processes a cotton leaf image and accurately predicts a fungal disease.

Generalization and Robustness

The model's robustness was enhanced through data augmentation, which allowed it to perform well under varied conditions. For example; the system demonstrated effective predictions for crops like Cotton and Grapes, showcasing its versatility.

Feedback Loop and Continuous Improvement

The Crop Pro interface also allows farmers to provide feedback on predictions of diseases. The iterative process improves the accuracy of the model by adding new data points to the dataset. For example, feedback from farmers who identified it as Tomato Leaf Disease was used to improve the model's accuracy for a similar case.

Discussion

Results of the CNN model indicate it to be a robust and scalable solution for crop disease detection. The training accuracy at 94.73% and validation accuracy at 88.75% validate that the model learns and generalizes the disease-specific patterns well. A slight difference in performance between training and validation calls for improvement in generalization, which can be achieved by increasing the dataset size and fine-tuning the model.

Being incorporated into the Crop Pro makes the system practical for the farmers. It enhances crop health management through the immediate detection of diseases. As more and more data are added to the database, it continually makes an effort to improve based on its continuous feedback against regional challenges of disease manifestation.

CONCLUSION

This paper presents a CNN-based crop leaf disease detection model integrated with the Crop Pro user interface for real-time use. The model performed well to have a training accuracy of about 94.73%, while the validation accuracy for it was 88.75%. It successfully generalizes to different crop types as well as disease categories. Since the model can classify diseases including Tomato Late Blight, and Bacterial Leaf Blight, it has excellent prospects for automating detection procedures in agriculture.

Crop Pro will allow farmers to interact with the model seamlessly, upload images of leaves, and receive a timely diagnosis of disease with associated confidence scores. The feedback loop built into the system ensures continuous improvement through updating the dataset with real-world observations that further enhance model accuracy and adaptability.

Future work will be carried out in the optimization of the model for real-time disease detection, integration with environmental factors like soil moisture and temperature, and an addition to the dataset of the availability of other crops and diseases. To address these issues, the proposed system would look to provide all-around precision agriculture solutions improving crop health management and promoting good agricultural practices.

REFERENCES

[1] Abdelmalik Ouamane, Ammar Chouchane, Yassine Himeur, Abderrazak Debilou, Abbes Amira, Shadi Atalla, Wathiq Mansoor, H. A. A. (2024). Enhancing Plant Disease Detection: A Novel CNN-Based Approach with Tensor Subspace Learning and HOWSVD-MD. Neural Computing and Applications.

[2] Abdu, A. M., Mokji, M. M. M., & Sheikh, U. U. U.Abdu, A. M., Mokji, M. M. M., & Sheikh, U. U. U. (2020). Machine learning for plant disease detection: An investigative comparison between support vector machine and deep learning. International Journal of Artificial Intelligence, 9(4), 670.

[3] Arya, S., & Singh, R. (2019). An analysis of deep learning techniques for plant leaf disease detection. International Journal of Computer Science and Information Security (IJCSIS), 17(7), 73–80.

[4] Ding, W., Abdel-Basset, M., Alrashdi, I., & Hawash, H. (2024). Next generation of computer vision for plant disease monitoring in precision agriculture: A contemporary survey, taxonomy, experiments, and future direction. Information Sciences, 665.

[5] G. Sachdeva, P. Singh, and P. K. (2021). Plant leaf disease classification using deep Convolutional neural network with Bayesian learning. Materials Today Proceedings, 45, 5584–5590.

[6] Hassan, S. M., Maji, A. K., Jasiński, M., Leonowicz, Z., & Jasińska, E. (2021). Identification of plant-leaf diseases using CNN and transfer-learning approach. Electronics, 10(12), 1388.

[7] Hatuwal, B., Joshi, B., Hatuwal, B. K., & Shakya, A. (2021). Plant leaf disease recognition using random forest, KNN, SVM and CNN. POLIBITS, 62, 13–19.

[8] Kim, C. H., Samsuzzaman, M. N. R., Ali, K. Y. L. M. R., Chung, S. O., & Kabir, M. S. N. (2023). Deep learning-based identification of Pepper (Capsicum annuum L.) diseases: A review. Precision Agriculture Science and Technology, 5(2).

[9] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. IEEE Access, 9.

[10] Mohameth, F., Bingcai, C., & Sada, K. A. (2020). Plant disease detection with deep learning and feature extraction using Plant Village. Journal of Computer Communications, 8(6), 10–22.

[11] Ngugi, H. N., Ezugwu, A. E., Akinyelu, A. A., & Abualigah, L. (2024). Revolutionizing crop disease detection with computational deep learning: A comprehensive review. Environmental Monitoring and Assessment, 196(3), 302.

[12] Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based plant diseases detection using deep learning. Materials Today Proceedings, 3500–3506.

[13] Rajendra, A. B., Rajkumar, N., & Shetty, P. D. (2019). Areca nut disease detection using image processing. In Soft Computing: Theories and Applications: Proceedings of SoCTA 2019. Singapore: Springer Singapore.

[14] Reddy, S. R. G., Varma, G. P. S., & Davuluri, R. L. (2023). Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification. Computers & Electrical Engineering.

[15] Saleki, S., & Tahmoresnezhad, J. (2024). Agry: A comprehensive framework for plant diseases classification via pretrained EfficientNet and convolutional neural networks for precision agriculture. Multimedia Tools and Applications, 83(24), 64813–64851.

[16] Salih, T. A., Ali, A. J., & Ahmed, M. N. (2020). Deep learning convolutional neural network to detect and classify tomato plant leaf diseases. OALib, 7(5), 1–12.

[17] Samir Bhattarai. (n.d.). New plant diseases dataset. https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset/data

[18] Singh, A. K., Rao, A., Chattopadhyay, P., Maurya, R., & Singh, L. (2024). Effective plant disease diagnosis using Vision Transformer trained with leafy-generative adversarial network-generated images. Expert Systems with Applications, 254, 124387.

[19] Tudi, M., Daniel Ruan, H., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., & Phung, D. T. (2021). Agriculture Development, Pesticide Application and Its Impact on the Environment. International Journal of Environmental Research and Public Health, 18(3), 1112.