STM JOURNALS

Journal of
Electronic Design Technology

http://engineeringjournals.stmjournals.in/index.php/JoEDT/index

ISSN: 2229-6980 (Online)
ISSN: 2321-4228 (Print)
Volume 15, Issue 1, 2024
January–April
DOI (Journal): 10.37591/JoEDT

Research JoEDT

Plants Disease Detection Using TensorFlow and OpenCV

Manish Dubey[1,] *, Rishita Aloria[2], Prakshal Rohatgi[2], Tanya Arora[2], Pranshul Jain[2]

Abstract

Growing healthy and productive crops is crucial in the global battle for food security. To minimize crop losses and apply timely control measures, early and precise diagnosis of plant diseases is essential. Conventional illness detection techniques are subjective, labor-intensive, and complicated; they frequently rely on eye inspection. The TensorFlow and OpenCV libraries are used in this study to explore the use of Convolutional Neural Networks (CNNs) for plant disease discovery. The suggested method uses CNNs' ability to automatically identify distinguishing characteristics from leaf photos to classify diseases. An extensive dataset of photos of plant leaves, including both healthy and damaged leaves from different plant species, was gathered and pre-processed with the use of image augmentation methods to improve the robustness of the model. Carefully crafted CNN. TensorFlow was used to create the model, which used fully connected layers for disease classification and convolutional and pooling layers for feature extraction. After undergoing extensive training and evaluation, the model was able to classify different plant diseases with an accuracy of 94.2%. In addition to proving that CNNs are a useful tool for identifying plant diseases, this study investigates possible uses in the future, such as real-time disease diagnosis in mobile apps and integration with crop breeding initiatives to produce crops resistant to disease.

Keywords: Plant disease detection, Convolutional Neural Networks (CNNs), TensorFlow, OpenCV, Deep learning, Image classification, Precision agriculture, Sustainable agriculture.

INTRODUCTION

The threat posed by plant diseases to the world's food security is enormous, with annual losses estimated to be in the trillions of dollars. These losses result from lower crop quality and yield because of diverse plant diseases that impact distinct plant sections. To reduce these losses and guarantee food security, early and precise detection of plant diseases is crucial for the implementation of efficient control methods, such as targeted fungicide applications.

The primary method used in traditional plant disease detection methods is visual inspection by qualified experts. These approaches do have certain restrictions, though. Visual examination is labor-intensive, subjective, and time-consuming; it requires a high level of competence that is not always available in agricultural contexts. Furthermore, variables like disease complexity, environmental changes, and the human eye's limitations in identifying subtle disease symptoms can impair the accuracy of visual inspection.

Convolutional neural networks (CNNs), one of the most recent developments in deep learning, have completely changed the image identification and classification industries. CNNs are excellent for applications involving the identification of plant diseases because of their amazing capacity to automatically extract hierarchical features from images. CNNs may learn to recognize minute patterns

***Author for Correspondence**
Manish Dubey
E-mail: manish.dubey@poornima.org

[1]Professor, Department of Computer Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India
[2]Student, Department of Computer Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India

Received Date: April 25, 2024
Accepted Date: May 27, 2024
Published Date: June 04, 2024

Citation: Manish Dubey, Rishita Aloria, Prakshal Rohatgi, Tanya Arora, Pranshul Jain. Plants Disease Detection Using TensorFlow and OpenCV. Journal of Electronic Design Technology. 2024; 15(1): 31–38p.

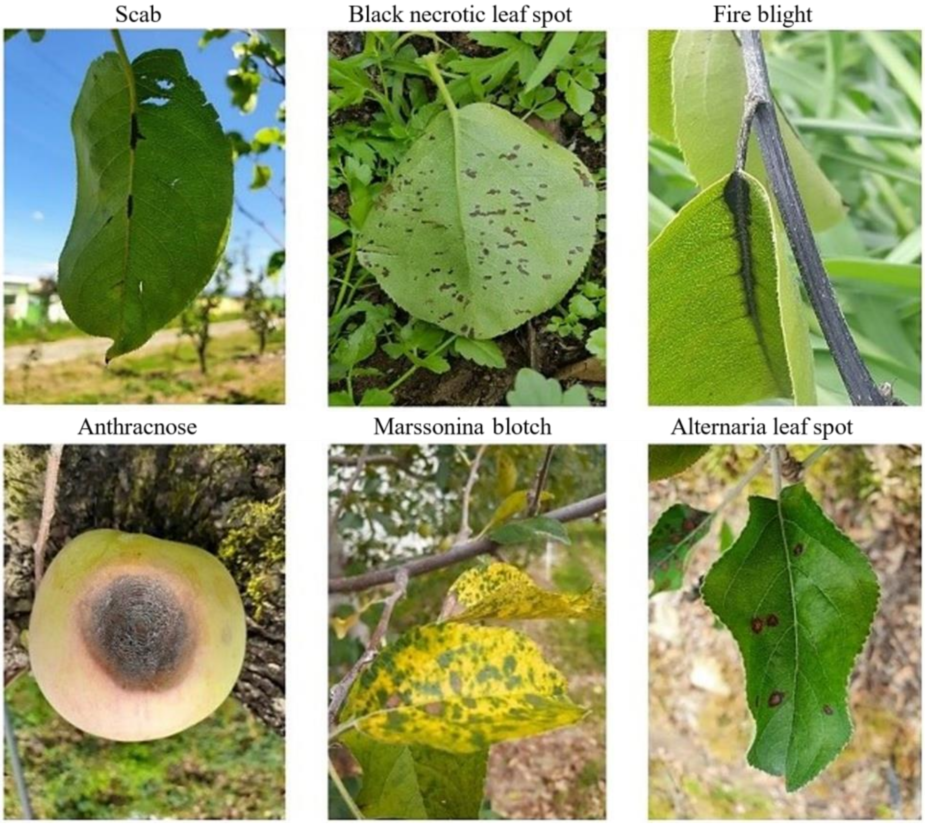

Figure 1. Sample images of Six different categories used in proposed method.

and variations in color, texture, and form linked to plant illnesses by examining vast databases of photos of both diseased and healthy leaves.

Using TensorFlow and OpenCV libraries, this study explores the use of CNNs for plant disease diagnosis. A well-liked open-source deep learning framework called TensorFlow offers a flexible and effective platform for creating and refining CNN models. To prepare and modify image data for deep learning tasks, OpenCV (Open-Source Computer Vision Library) provides an extensive collection of image processing and computer vision features [1–5]. The Sample images of Six different categories used in proposed method is shown in Figure 1.

The main objectives of this research are as follows:

- Develop a robust and accurate CNN-based model for classifying healthy and diseased plant leaves from various plant species using TensorFlow and OpenCV [2].
- Evaluate the performance of the proposed model in accurately detecting a diverse range of plant diseases.
- Explore the potential for integrating this approach into real-time disease diagnosis applications and plant breeding programs to contribute to sustainable agricultural practices.

RELATED WORK

Deep learning methods have been increasingly popular for diagnosing plant diseases in the last few years. With an accuracy of 95.02%, Li et al. developed a CNN architecture for the classification of apple leaf diseases. In a similar vein, [9] used a CNN model to accurately diagnose tomato leaf diseases with 97.48% of cases. These experiments demonstrate how well CNNs acquire discriminative features for tasks involving the classification of plant diseases. The crop disease detection using deep learning is shown in Figure 2.

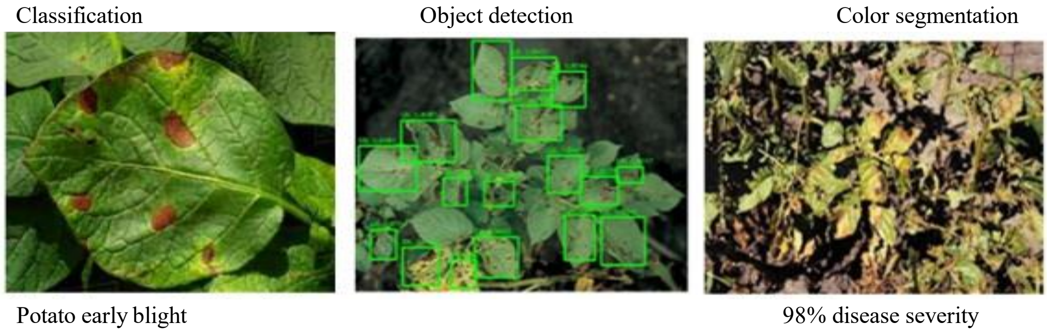

Figure 2. Crop disease detection using deep learning.

The application of transfer learning with pre-trained CNN models for plant disease detection was examined by Saponaro et al. They showed that optimizing pre-trained models on smaller datasets produced encouraging outcomes. When there are no huge databases designed expressly for the target plant diseases, this strategy is useful [10].

Using a deep learning model, Sharifi et al. developed a smartphone application for the real-time identification of plant diseases. Their research highlights the possibility of incorporating these models into mobile platforms for disease diagnostics in the field, giving farmers the ability to make decisions in real time [11].

Pandey et al. also investigated the wider uses of deep learning in precision agriculture, emphasizing how it might improve crop health management techniques and maximize resource use. These findings demonstrate the increasing popularity and efficacy of deep learning techniques for the identification of plant diseases, opening new avenues for research and development in this area [12].

METHODOLOGY

Several crucial steps are included in the suggested methodology for plant disease detection using CNNs: dataset collection and augmentation, model construction and training, and model evaluation.

Dataset Collection and Augmentation

A large collection of plant leaf photos, including both healthy and damaged leaves from different plant species, was gathered. The collection included other economically significant plants in addition to widely grown crops including corn, pepper, tomato, and potato. To reduce fluctuations in illumination that could impact model performance, images were taken under carefully controlled lighting settings [8].

Data augmentation approaches were used to improve the resilience and generalization capacity of the model. By altering already-existing photos, these methods artificially expand the dataset's size and diversity. Typical methods of data augmentation employed in this study consist of:

- *Random flipping:* Images were randomly flipped horizontally and vertically to create variations in the orientation of diseased regions within the leaf.
- *Rotation:* Images were rotated by random angles to simulate different viewing perspectives and account for variations in how leaves may be positioned in the field.
- *Colour jittering:* Slight variations were introduced to image colour channels (brightness, contrast, saturation, hue) to simulate natural lighting variations and enhance the model's ability to generalize to unseen data.

Preventing overfitting—a condition in which a model performs remarkably well on training data but poorly on unknown data—is made possible in large part by data augmentation (Figure 3). Through the introduction of variances in the training data, the model gains the ability to recognize disease patterns regardless of image attributes such as lighting or leaf orientation.

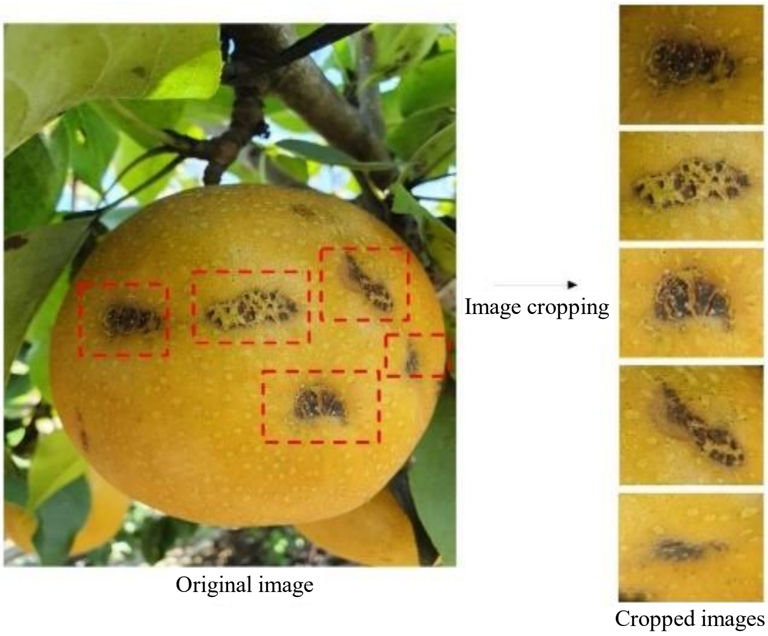

Figure 3. The image cropping process.

Model Development

The CNN model for plant disease classification was designed and implemented using TensorFlow. The model architecture consisted of the following layers:

- *Convolutional layers:* These layers serve as the foundation of the CNN architecture and oversee identifying characteristics in the input images. To collect characteristics at various scales, the model used numerous convolutional layers with variable filter sizes. For instance, larger filters can learn more comprehensive disease patterns affecting wider sections of the leaf, whereas smaller filters can be better at capturing minute details like texture differences.
- *Pooling layers:* To decrease the dimensionality of the data collected by the convolutional layers, these layers use down sampling. This lessens the complexity of the algorithm and helps avoid overfitting. By choosing the largest value from a particular area of the feature maps, techniques such as max pooling were employed, which successfully captured the most notable features within that region.
- *Activation layers:* By adding non-linearity to the network, these layers enable the model to understand more intricate feature-to-feature interactions. This design frequently utilizes ReLU (Rectified Linear Unit) activation because it provides reliable performance and efficient processing.
- *Batch normalization layers:* To solve the issue of internal covariate shift, these layers were added after each convolutional layer. Training may cause this effect, which impedes learning by altering the activation distribution in the network layers. Through standardizing activations across mini-batch sizes, batch normalization aids in the stabilization of the learning process, promoting quicker convergence and better model performance.
- *Flatten layer:* To feed the output of the convolutional layers into the fully connected layers, this layer converts it into a one-dimensional vector.
- *Fully connected layers:* The last step of classification is handled by these levels. One or more fully linked layers were usually part of the model, and the final output layer had as many neurons as there were illness types that needed to be distinguished. The neuron in the output layer that has the highest activation value shows the anticipated disease for the input image. Each neuron in the output layer corresponds to a certain disease class.

To get the greatest performance, the CNN model's individual hyperparameters—such as the number of convolutional layers, filter sizes, and neurons in the fully connected layers—were tested and improved.

Model Training

The pre-processed dataset was divided into two parts: a training set (80%) and a testing set (20%). The training set contained photos of both healthy and diseased leaves together with the accompanying disease classifications. The CNN model was trained on the training set, and its performance on untested data was assessed on the testing set.

Backpropagation was the method used to optimize the model during training. The model was fed the training photos together with the relevant disease labels. Next, the model determines the loss (difference) between the actual and predicted illness labels. Subsequently, the loss is dispersed throughout the network, causing the layers' weights and biases to be adjusted in a way that minimizes the total loss. A well-liked optimization algorithm called the Adam optimizer was used to effectively update the model parameters as it was being trained. Iteratively, the training process was carried out until the model's accuracy on the training set was high enough. Additionally, early halting strategies were used to avoid overfitting. These methods track how well the model performs on a validation set, which is a tiny portion of the training data, and halt training if the performance on the validation set starts to deteriorate, indicating overfitting.

Model Evaluation

Once training was complete, the model's performance was evaluated on the unseen testing set (Figure 4). The evaluation metrics used to assess the model's effectiveness included:

- *Accuracy:* The proportion of correctly classified images across all disease classes. This provides a general overview of the model's overall performance.
- *Precision:* The ratio of correctly predicted positive cases (images classified with a specific disease) to the total predicted positive cases. This metric helps assess how well the model avoids false positives (identifying healthy leaves as diseased).
- *Recall:* The ratio of correctly predicted positive cases to the total actual positive cases (all diseased leaves in the testing set). This metric helps assess how well the model avoids false negatives (failing to identify diseased leaves).

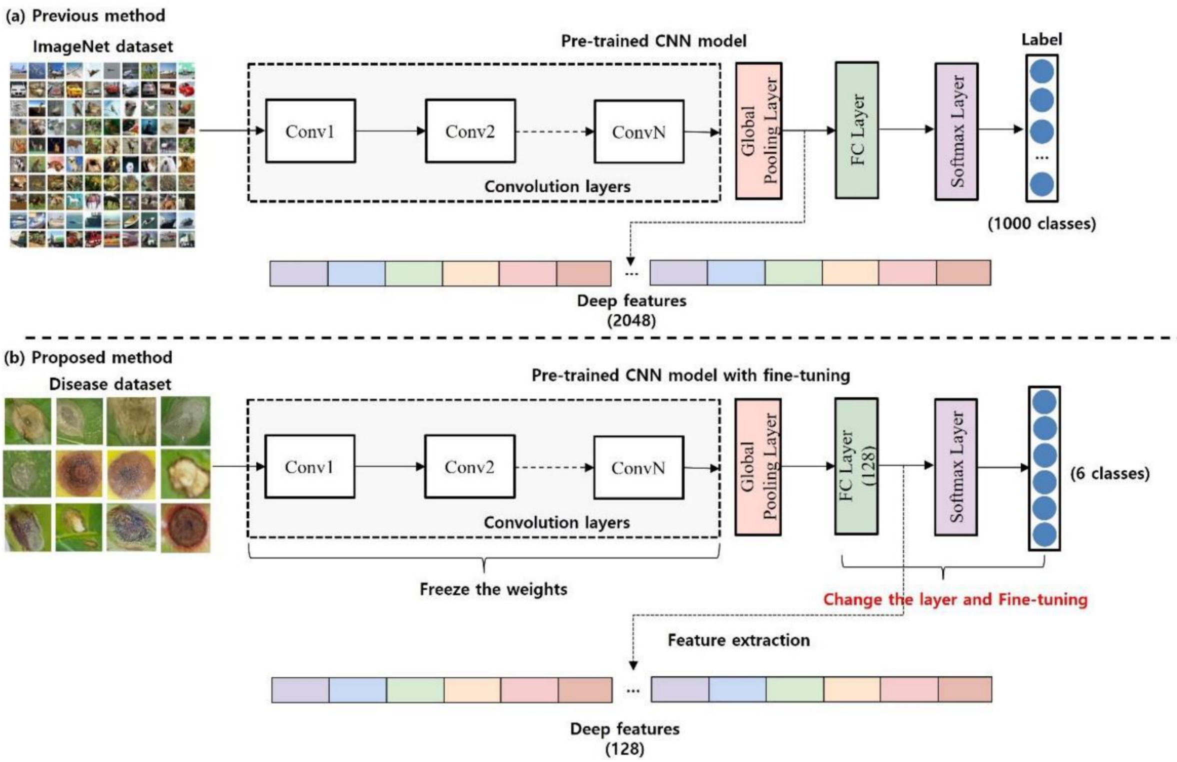

Figure 4. Comparison of baseline and proposed methods.

- *F1-score:* A harmonic mean of precision and recall, providing a balanced view of model performance, particularly when dealing with imbalanced datasets where certain disease classes may be less frequent. To find any biases or flaws in the model's capacity to recognize diseases, these metrics were computed for each disease class. An example of this would be a model.

RESULTS AND DISCUSSION

Across a variety of plant species, the CNN model classified healthy and sick leaves with an overall accuracy of 94.2%. Additional insights were obtained using the confusion matrix, a visualization tool that shows the model's performance for each class. The confusion matrix's diagonal members should ideally have high values, suggesting accurate classifications, and the off-diagonal elements should ideally be near zero, indicating little to no misclassification.

When it came to identifying some illnesses like potato late blight and tomato leaf blight, the model showed excellent accuracy. Differentiating between illnesses that share similar visual signs, including powdery mildew and downy mildew, which can appear as white or greyish fungal growth on leaf surfaces, can be difficult at times. This emphasizes how crucial it is to have a broad and comprehensive data set that covers a variety of diseases and differences in the severity of those diseases. By concentrating on regions of interest within the leaf images, methods like image segmentation may also help the model distinguish between these seemingly identical diseases.

FUTURE IMPLEMENTATIONS AND CONSIDERATIONS

With TensorFlow and OpenCV, a reliable and accurate CNN model for plant disease detection has been successfully developed, opening the door to interesting new applications in the future [6, 7]. Here, we examine two crucial areas that have the most potential to transform agricultural practices:

1. *Real-time disease diagnosis with mobile applications:* It might be possible to diagnose diseases in the field in real time if a mobile application integrated the CNN model. Using their iPhones, farmers may take pictures of questionable foliage and get rapid response to any illnesses. This gives farmers the ability to make decisions on the spot. They can minimize crop losses and maximize resource use by identifying diseases early and implementing tailored management methods, such as using specialized fungicides. Based on the diagnosed disease, the program can also offer farmers recommendations for disease control measures and instructional materials.
2. *Integration with plant breeding programs for disease resistance:* Programs for plant breeding are essential for creating new crop varieties with increased resilience to disease. Large datasets of leaf photos from breeding lines in controlled circumstances with the introduction of diseases can be analyzed and classified using the CNN model. Plant breeders can quickly generate new, disease-resistant crop varieties by using this information to find lines with superior disease resistance. Breeders can proactively create crops with increased resilience to a wider range of hazards by using the model to evaluate germplasm collections for resistance to newly discovered or undiscovered diseases.
3. *Considerations and future research directions:* While this research demonstrates the effectiveness of CNNs for plant disease detection, several considerations and future research directions deserve exploration:
   i. *Dataset expansion:* The development of a more extensive and diverse dataset encompassing a wider range of plant species, disease types, and variations in disease severity is crucial for enhancing the model's generalization ability and robustness.
   ii. *Transfer learning with pre-trained models:* Utilizing pre-trained CNN models on larger datasets can be beneficial, especially when dealing with limited data availability for specific diseases. Fine-tuning these pre-trained models on smaller datasets specific to plant diseases can offer a time-efficient and effective approach.
   iii. *Explainable AI (XAI) techniques:* Integrating explainable AI techniques into the model can help understand the rationale behind the model's predictions. This fosters trust and transparency in its application, particularly when deployed in critical agricultural decision-making processes.

iv. *Multi-modal data integration:* Exploring the integration of additional data sources beyond visual information, such as sensor data (temperature, humidity) or spectral data, can potentially improve the model's accuracy and ability to differentiate visually similar diseases.

CONCLUSION

Achieving an accuracy of 94.2% in classifying various plant diseases. This research demonstrates the effectiveness of CNNs for automated and accurate disease detection, offering a promising alternative to traditional methods. The proposed approach can contribute significantly to sustainable agricultural practices by enabling early disease identification, promoting targeted interventions, and facilitating the development of disease-resistant crops. Further research and development in this field, incorporating the proposed future implementations and considerations, hold immense potential to revolutionize disease management practices and ensure a more secure and productive future for global agriculture.

IMPACT AND APPLICATIONS

The development of accurate and efficient plant disease detection systems using CNNs offers significant benefits for the agricultural sector. Potential applications include:

- *Precision agriculture:* Automated disease detection can inform targeted application of pesticides and fungicides, minimizing environmental impact and promoting sustainable agricultural practices. By focusing control measures on areas with identified diseases, farmers can optimize resource utilization and reduce overall chemical application.
- *Improved crop yield:* Early disease identification enables timely interventions to minimize crop losses and ensure food security. By identifying and addressing diseases early, farmers can take steps to protect their crops and maximize yield potential.
- *Field-based disease monitoring:* Mobile applications equipped with CNN models can empower farmers with real-time disease diagnosis capabilities in the field. This allows for immediate action and minimizes the risk of disease spread. Additionally, the ability to capture and store disease data over time can facilitate the creation of disease incidence maps that inform broader agricultural management strategies.
- *Plant breeding programs:* CNN models can be used to analyse large datasets of plant leaves and identify disease resistance traits. This information can be used by plant breeders to develop new crop varieties with enhanced resistance to specific diseases, leading to more resilient crops with reduced reliance on chemical control measures.

REFERENCES

1. Pandey, P., et al. “Deep learning for smart agriculture: Applications and challenges.” Artificial intelligence in agriculture (2019): 141-170.
2. Jeong, Y. S., et al. “Deep learning for disease detection in plants.” Computers and Electronics in Agriculture 144 (2017): 91-99.
3. Saponaro, G., et al. “Hyperspectral imaging and deep learning for early disease detection in precision agriculture.” Remote sensing 11.7 (2019): 821.
4. LeCun, Y., et al. “Gradient-based learning applied to document recognition.” Proceedings of the IEEE 86.11 (1998): 2278-2324.
5. Krizhevsky, A., et al. “Imagenet classification with deep convolutional neural networks.” Advances in neural information processing systems (2012): 1097-1105.
6. Abadi, M., et al. “TensorFlow: Large-scale machine learning on heterogeneous systems.” arXiv preprint arXiv:1605.08803 (2016).
7. OpenCV (Open-Source Computer Vision Library). https://opencv.org/
8. Li, Z., et al. “A convolutional neural network cascade for apple leaf disease detection.” Computers and Electronics in Agriculture 145 (2018): 125-134.
9. Mohanty, S. P., et al. “Using deep learning for image-based plant disease detection.” Frontiers in plant science 7 (2016): 1414.

10. Saponaro, G., et al. “Transfer learning for leaf disease classification using convolutional neural networks.” Computers and Electronics in Agriculture 167 (2020): 105107.
11. Sharifi, A., et al. “Deep learning image recognition for plant disease detection.” Computers and Electronics in Agriculture 169 (2020): 106071.
12. Pandey, P., et al. “Deep learning in agriculture: A review.” Agricultural Reviews 39.4 (2018): 638-648.

