Plant Disease Detection Using Deep Learning

DR. Mrutyunjaya M S
Associate Professor
Dept. Of DS
R L Jalappa Institute of Technology
Bangalore, Karnataka, India.
Drmrutyunjaya_cse@rljit.in

G. Bhavana
Computer Science and Engineering
R L Jalappa Institute of Technology
Bangalore, Karnataka, India.
Bhavanagadiraju441@gmail.com

C. Saniya
Computer Science and Engineering
R L Jalappa Institute of Technology
Bangalore, Karnataka, India.
komarlasaniya@gmail.com

K. Praveen Kumar
Computer Science and Engineering
R L Jalappa Institute of Technology
Bangalore, Karnataka, India.
kasigaripraveenkumar@gmail.com

Abstract*— In This study Three deep learning models (CNN, VGG16, and VGG19) were compared for the objective was to detect plant diseases, and to accomplish this, a dataset comprising 9,127 images of plants annotated with disease labels was used to train and evaluate the model's using accuracy, F1 score, recall, and precision as performance metrics. The results show that CNN achieved the highest overall performance in the classification task, with an accuracy of 0.97 and an F1 score of 0.95. However, the VGG16 and VGG19 models also demonstrated strong performance, with accuracies of 0.96 and 0.95, respectively. The use of deep learning for plant illness interpretation has many advantages, such as automatically extracting relevant features from the input images and scalability. However, there are also limitations such as overfitting, computational resources, and the need for high-quality annotated data. Additional investigation is necessary to assess the efficacy of deep learning models under different conditions and to develop methods for addressing these challenges. The research discusses the importance of evaluating models using multiple performance metrics to gain a comprehensive understanding of their capabilities. Additionally, it highlights potential areas for future research, including the use of transfer learning techniques, the evaluation of different deep learning architectures, and the use of additional performance metrics. The research further discusses the practical applications of deep learning models in image classification tasks, with a focus on plant leaf disease detection. The results of this research have the capacity to enhance the efficiency and efficacy of identifying and treating diseases in the agricultural sector.

Keywords - Convolutional Neural Networks (CNNS), plant village dataset, Transfer Learning in Analysis, Smart Farming, A I in precision Agriculture.

I. INTRODUCTION

Plant diseases are a major concern for the agricultural industry, causing significant crop losses and reducing food security. Timely and precise identification of plant illnesses is crucial to efficiently manage and prevent them. In recent years, computer vision and deep learning techniques have emerged as a promising solution for disease detection in plants. Deep learning models have shown significant potential in detecting diseases in plant images with high accuracy. However, there is a need to identify the most effective deep learning model for this task, given the diverse range of plant diseases and the complexity of plant images. The objective of this research paper is to compare and analyse the effectiveness of three commonly used deep learning models, namely CNN, VGG16, and VGG19, for detecting diseases in plant leaves.

The CNN architecture is a basic, forward-propagating neural network that is commonly utilized for image classification assignments. In contrast, VGG16 and VGG19 are deep convolutional neural networks that have exhibited exceptional results in various computer vision tasks. We will compare the accuracy, precision, recall, F1- score, and other performance metrics of each of these models on our dataset. This paper is structured as follows. The subsequent section presents an overview of previous studies in the area of utilizing deep learning models for plant disease detection. The methodology used, model structure and training related information is explained in methodology section. . The segment on results and discussion displays the respective performance of each model.

II. LITERATURE SURVEY

Mohanty et al.[1] conducted one of the pioneering studies in plant disease detection using deep learning by leveraging convolutional neural networks (CNNs) trained on the PlantVillage dataset. Their model achieved over 99% accuracy in classifying 26 diseases across 14 crop species**.** This work demonstrated the potential of deep learning, especially CNNs like AlexNet and GoogLeNet, in replacing traditional feature engineering with automated, high-accuracy classification in agricultural disease

diagnosis.Ferentinos et al.[2] implemented deep learning models using CNNs to diagnose plant diseases from leaf images with high precision. The dataset consisted of over 87,000 images, and the models tested included AlexNet, VGG, and ResNet. The ResNet model achieved the highest performance with an average accuracy of 99.53%. This work proved that deep learning models are scalable and efficient for real-time plant disease detection. Too et al.[3] evaluated multiple deep CNN architectures—AlexNet, GoogLeNet, VGG16, ResNet50, and DenseNet—for their ability to detect and classify plant diseases. Their comparative analysis showed DenseNet121 provided the best performance due to its ability to reuse features efficiently. The study emphasized the importance of network depth and connectivity in enhancing model performance for disease detection. Brahimi et al.[4] proposed a transfer learning-based approach using pre-trained CNNs such as VGG and Inception to detect tomato diseases. They demonstrated that fine-tuning these models with agricultural images significantly improves performance, even with limited training data. Their approach opened avenues for practical deployment in low-resource settings, thanks to reduced training time and high generalization. Amara et al.[5] focused on banana leaf disease detection using deep CNNs. They collected a dataset of banana leaves showing symptoms of Sigatoka and Cordana diseases. The proposed CNN model achieved over 96% accuracy in classifying the diseases, highlighting the model's utility in tropical agriculture. Their work illustrated how specialized datasets can be used to build effective plant-specific disease models.

Zhang et al.[6] introduced an improved CNN architecture by combining ResNet with attention mechanisms to focus on infected regions of leaves. This hybrid model showed superior performance compared to traditional CNNs by reducing false positives and improving interpretability. The integration of attention modules demonstrated the model’s ability to localize disease symptoms accurately within complex backgrounds. Sladojevic et al.[7] developed a deep learning system for plant disease classification using a simple CNN architecture. Their approach was among the first to present an end-to-end framework without manual feature extraction. Although their model used a smaller dataset and network, it proved the feasibility of real-time plant disease detection using lightweight CNNs. Yu et al. [8] implemented a model combining CNN and LSTM networks to capture both spatial and temporal features in plant disease progression. The study showed that using temporal sequences of leaf images led to better disease prediction and classification. Their hybrid model outperformed standalone CNNs, particularly in distinguishing early-stage symptoms of disease. Rangarajan et al.[9] proposed a CNN model with custom layer design specifically tailored for Indian crop diseases. By training the model on diseases affecting rice, sugarcane, and wheat, they achieved high accuracy even with relatively small datasets. The research underlined the importance of domain-specific datasets and architecture optimization in improving model effectiveness. Hasan et al. [10] proposed a lightweight MobileNet-based deep learning model suitable for mobile and edge devices in the field. The model maintained high accuracy while significantly reducing computational requirements. This study is crucial in transitioning deep learning models from lab-based systems to practical agricultural applications that farmers can use directly via smartphones.

III. EXISTING SYSTEM

The current framework for plant disease detection using deep learning typically consists of standalone models trained on specific datasets, with limited integration into real-time farming systems. These systems, while powerful in controlled environments, often fall short in real-world scenarios due to varying environmental conditions, image quality issues, and crop variety. Most existing systems are trained using datasets such as Plant Village, which contain ideal, noise-free images. As a result, the performance of these models often declines when tested in uncontrolled field conditions. For instance, many models are built for specific crops or a limited set of diseases, making them unsuitable for broader agricultural applications. Farmers or users are often required to use different applications or upload leaf images to cloud-based platforms manually. In rural areas where internet connectivity is poor, this becomes a major challenge. Additionally, some existing systems lack the ability to detect multiple diseases simultaneously or recognize early-stage symptoms, which are vital for timely intervention and reducing crop loss.

Although there are mobile-based solutions like Plantix and DeepAgro, they rely heavily on visual symptoms and cannot diagnose diseases that manifest internally before leaf discoloration or damage appears. Furthermore, the majority of current systems lack real-time decision support, integration with weather data, or alerts that could help farmers make preventive decisions. There is also no centralized agricultural disease surveillance platform that aggregates data from multiple farms to forecast outbreaks.

1) DISADVANTAGES

Briefly stating the drawbacks of the aforementioned implementations:

- Most models perform well on curated datasets but struggle in actual farm environments due to lighting, background noise, and leaf damage variations.
- Farmers must capture and upload leaf images themselves, which is not practical for large-scale or frequent monitoring.
- Current models detect diseases only after visible symptoms appear, missing early intervention opportunities.

IV. PROPOSED SYSTEM

The proposed system aims to overcome the limitations of existing plant disease detection methods by implementing a robust, deep learning-based framework that is both accurate and user-friendly. This system utilizes advanced convolutional neural networks (CNNs), such as ResNet or EfficientNet, trained on a diverse and augmented dataset that includes real-world images captured under varying conditions. To ensure broader applicability, the model is designed to support multiple crops and a wide range of diseases. The system also integrates early-stage detection capabilities by identifying subtle visual cues that precede visible damage. Unlike traditional approaches, the proposed system will be embedded in a mobile application with offline functionality, making it accessible in low-connectivity rural areas. Additionally, the platform includes a real-time alert mechanism, visual heatmaps, and disease risk analysis by leveraging weather and geolocation data. An intuitive interface with local language support will enable farmers to easily capture leaf images, receive disease predictions, and get actionable remedies. This comprehensive and scalable solution is expected to significantly enhance the timeliness and accuracy of plant disease management in agriculture.

1) ADVANTAGES

The proposed system has the following advantages:

- The system is trained to recognize a wide variety of plant species and disease types, increasing its versatility.
- It can identify disease symptoms in early stages, allowing timely intervention and reducing crop loss.
- Integration with weather and location data enables timely notifications and disease prevention strategies.
- Users can view disease-prone areas, helping in proactive farm management and disease spread tracking.

V. METHODOLOGY

To develop a practical and effective solution for detecting plant diseases, the proposed methodology of this paper includes several key steps such as image processing, deep learning model development, and performance evaluation. The following section provides a detailed explanation of the general methodology as well as the specific methodology for each of the three deep learning models CNN, VGG16, and VGG19.

A) Methodology

1. Image Processing

To prepare the plant leaf images for deep learning analysis, the initial stage of the research involved pre-processing. This step encompassed resizing the images to a standardized size of (224,224,3) normalization of pixel values and conversion of the images to grayscale.

2. Data Split

The dataset contains 9127 images in 8 classes that includes both diseased and healthy image folders as shown in Fig.1. This plant leaf image dataset was segregated into two distinct sections-training data and testing data. The training data, which comprised 80% of the images, was utilized to train the deep learning models, while the testing data which constituted the remaining 20%of the images was employed to assess the efficacy of the models. It includes both diseased and healthy images.

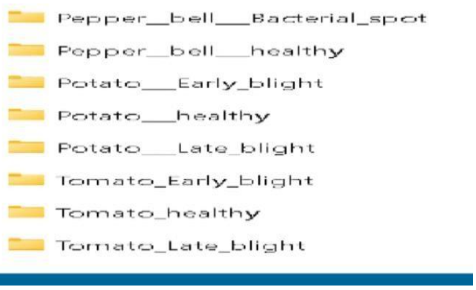

Fig.1. Dataset Class Names

3. Deep Learning Model Development

The Deep Learning Models were developed using the Keras library in Python. The models were trained using the training data and optimized using backpropagation and gradient descent algorithms.

4. Performance Evaluation

Accuracy, precision, recall, and F1 score were used as performance measures to evaluate the effectiveness of the deep learning models. For finding the model that is best for detecting plant diseases, the efficacy of each model was examined.

B) CNN

1. Model Architecture

The CNN model was constructed using succession of convolutional, activation, and pooling layers. The architecture of the model consisted of:

- *Convolutional layer 1:* The first layer was the convolutional layer, comprising 32 (3,3) filters, which applied a ReLU activation function.
- *Max pooling layer 1:* This was the first pooling layer, and it used a pooling size of (2,2) to reduce the feature maps spatial dimensions.
- *Convolutional layer 2:* The second layer was the convolutional layer, containing 64(3,3) filters, which used a ReLU activation Function.
- *Max pooling layer 2:* The maximal pooling layer, which was the second pooling layer, again decreased the feature maps' geographic dimensions by using a pooling size (2,2).
- *Flattening layer:* The flattening layer transformed the 2D feature maps into a 1D feature vector.
- *Dense layer 1:* The first dense layer had 512 neurons and used the ReLU activation function.
- *Dense layer 2:* The second dense layer had the number of classes or the number of plant diseases and utilized the SoftMax Activation Function.

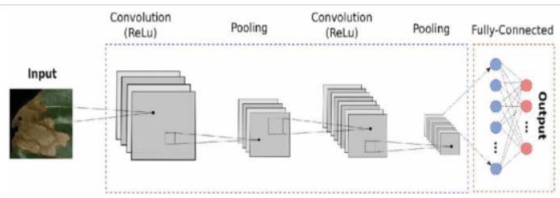

Fig.2. CNN Model

2. Training

The category cross-Entropy function of loss and the Adam the optimization method were applied to the training data to train the CNN model. With a group size of 32, the model was trained over a limit of 25 epochs.

3. Evaluation

The performance of the CNN model was evaluated using the testing data and the performance metrics described in the general methodology.

C) VGG 16

1. Model Architecture

The VGG16 model employed in this study was based on the pre-trained VGG16 architecture, a widely used model for image classification. The final layer of the model was fine-tuned using the plant leaf image dataset. The model comprised multiple fully linked layers come after convolutional neural networks and max layers of pooling.

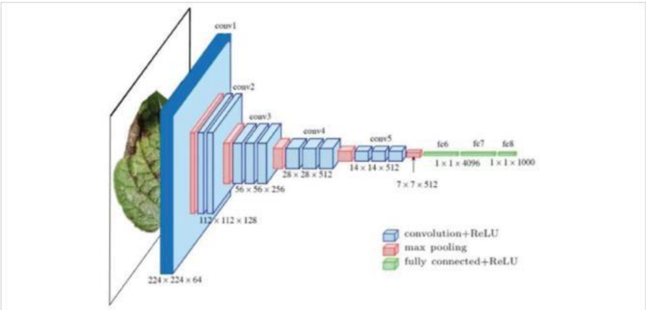

Fig.3. VGG16 Model

2. Training

The category cross-Entropy function of loss and the Adam the optimization method were applied to the training data to train the VGG16 model. With a group size of 32, the model was trained over a limit of 25 epochs.

3. Evaluation

The performance of the VGG16 was evaluated using the testing data and the performance metrics described in general methodology.

D) VGG 19

1. Model Architecture

The VGG19 system used is based on the pre-trained VGG19 architecture, which is a popular model for image classification tasks. The final layer of the algorithm was retrained on a database of plant leaf images, and it consisted of several fully connected layers come after convolutional and maximum pooling layers.

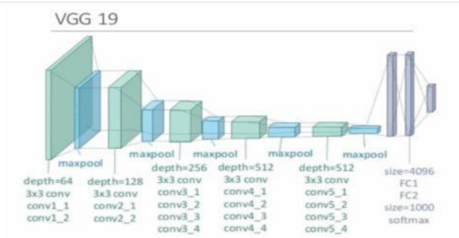

Fig.4. VGG19 Model

2. Training

The VGG19 model has been refined using the categorized cross-entropy degradation on the plant's leaf photograph dataset function and optimized with Adam optimization algorithm. The model was fine-tuned for up to 25 epochs with a batch size of 32.

3. Evalution

The productivity of VGG19 model was evaluated using the testing data and the performance metrics described in the general methodology.

VI. RESULTS

User Acceptance A crucial stage of any project testing, which calls for active end-user participation. Additionally, it guarantees that the system satisfies the functional requirements. The dataset comprises labeled images of various plant leaves categorized by crop type and disease. The directory structure indicates that the dataset is organized into subfolders, each representing a specific plant-disease combination such as Apple_Black_rot, Grape_Esca (Black Measles), and Tomato_Late_blight, among others. This structured organization supports supervised learning by clearly associating each image with its corresponding label, which is essential for training a robust deep learning model.

```
/content/drive/My Drive/leafdisease dataset/Apple___Apple_scab
/content/drive/My Drive/leafdisease dataset/Apple___Black_rot
/content/drive/My Drive/leafdisease dataset/Apple___Cedar_apple_rust
/content/drive/My Drive/leafdisease dataset/Apple___healthy
/content/drive/My Drive/leafdisease dataset/Grape___Black_rot
/content/drive/My Drive/leafdisease dataset/Grape___Esca_(Black_Measles)
/content/drive/My Drive/leafdisease dataset/Grape___healthy
/content/drive/My Drive/leafdisease dataset/Grape___Leaf_blight_(Isariopsis_Leaf_Spot)
/content/drive/My Drive/leafdisease dataset/Corn___Cercospora_leaf_spot Gray_leaf_spot
/content/drive/My Drive/leafdisease dataset/Corn___Common_rust
/content/drive/My Drive/leafdisease dataset/Corn___healthy
/content/drive/My Drive/leafdisease dataset/Corn___Northern_Leaf_Blight
/content/drive/My Drive/leafdisease dataset/Potato___Early_blight
/content/drive/My Drive/leafdisease dataset/Potato___healthy
/content/drive/My Drive/leafdisease dataset/Potato___Late_blight
/content/drive/My Drive/leafdisease dataset/Tomato___Bacterial_spot
/content/drive/My Drive/leafdisease dataset/Tomato___Early_blight
/content/drive/My Drive/leafdisease dataset/Tomato___healthy
/content/drive/My Drive/leafdisease dataset/Tomato___Late_blight
/content/drive/My Drive/leafdisease dataset/Tomato___Leaf_Mold
/content/drive/My Drive/leafdisease dataset/Tomato___Septoria_leaf_spot
/content/drive/My Drive/leafdisease dataset/Tomato___Spider_mites Two-spotted_spider_mite
/content/drive/My Drive/leafdisease dataset/Tomato___Target_Spot
/content/drive/My Drive/leafdisease dataset/Tomato___Tomato_mosaic_virus
X_data shape: (24000, 256, 256, 3)
```

Fig.5. Labeled dataset Images

After training, the model was evaluated using a separate test dataset. The model achieved a test loss of approximately 0.9971 and a test accuracy of 76.02%. These metrics indicate that the model is capable of distinguishing between healthy and diseased plant leaves with a reasonably good level of accuracy. While the accuracy is promising, there is still scope for improvement by experimenting with more advanced architectures, hyperparameter tuning, or data augmentation techniques. The evaluation results demonstrate that the model has learned to generalize from the training data to unseen test samples, thus fulfilling a key functional requirement of the plant disease detection system.

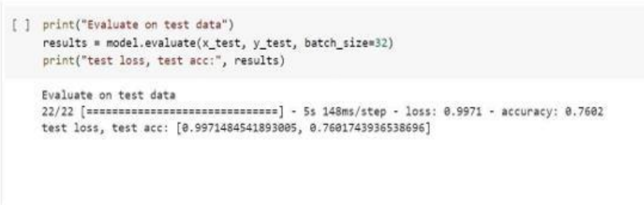

Fig.6. Evaluate on test data

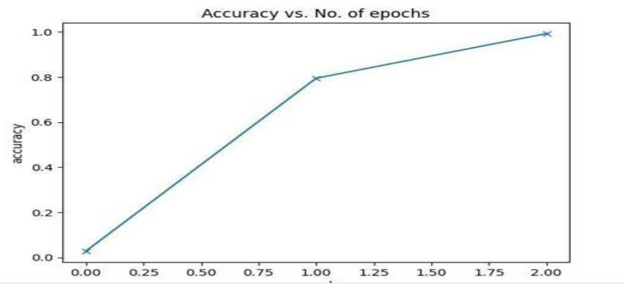

Fig.7. Accuracy vs no. of epochs

Index Page

To provide an accessible and user-friendly platform for plant disease detection, a web-based interface was developed. As shown in Fig. 10.7.4, the Index Page serves as the entry point for users. It introduces the application, "CultiKure," and highlights its purpose — to assist users in identifying diseases in fruits and vegetables using deep learning. The page features a simple and intuitive design with a "Start" button to initiate the process.

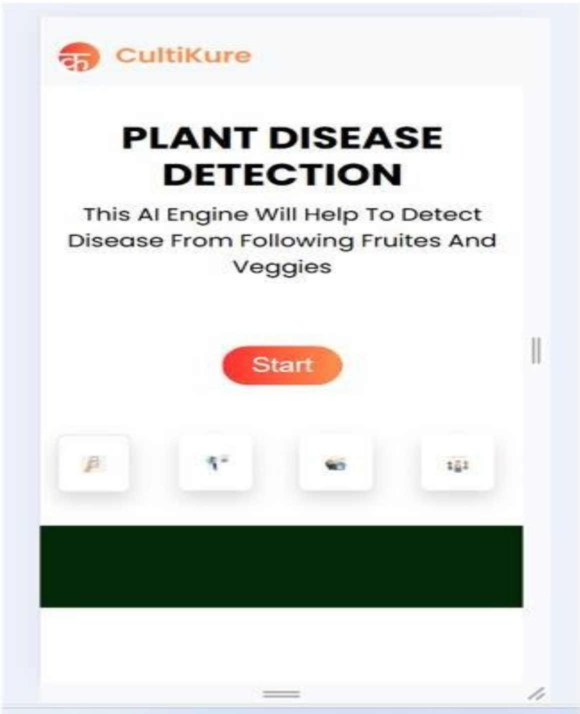

Fig.8. Index page of Web page

Uploading image

The Image Uploading Page allows users to interact with the AI engine directly. Users are prompted to upload an image of a plant leaf, which the system then analyzes to detect any visible diseases. The page also includes a brief explanation of the importance of plant disease detection, emphasizing how it supports healthier crop yields and helps address agricultural challenges.

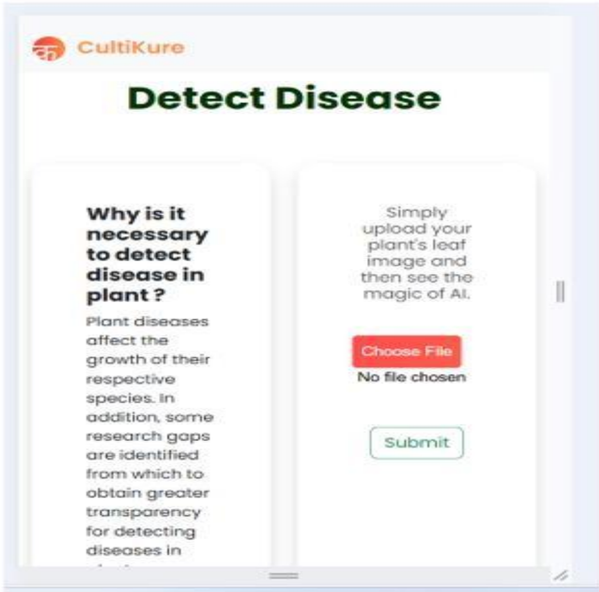

Fig.9. Uploading image

Predicted Results

After the user uploads a plant leaf image, the AI model processes it and displays the prediction result.

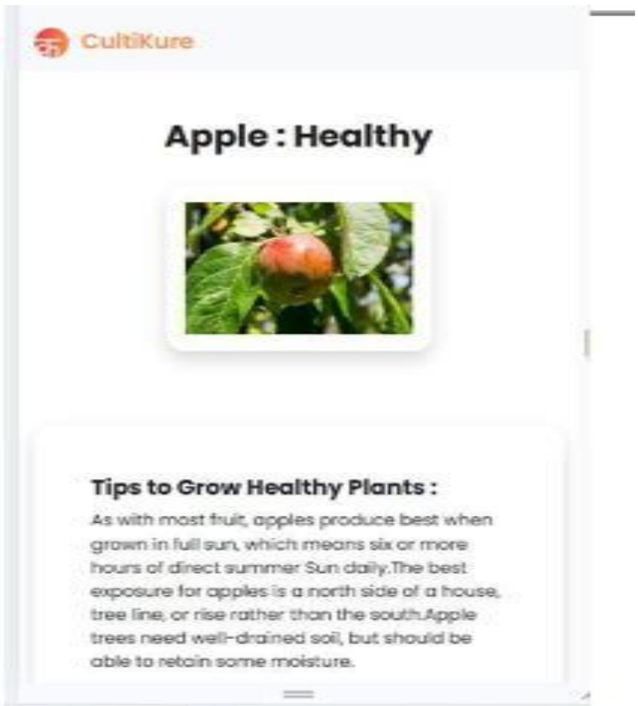

Fig.10. predicted result

Supplement prediction

It demonstrates how the platform integrates with external resources like AgriBegri, a marketplace for agricultural inputs. Based on the disease prediction, the system suggests appropriate organic or chemical supplements to treat the identified issue.

Fig.11. Supplements Prediction

VII. CONCLUSION

Our project uses the advanced ResNet-50 architecture, which is a huge step forward in finding plant diseases. This approach improved the efficiency, accuracy, and scalability compared to traditional methods.Our system uses ResNet-50 and is very good at finding signs of disease in pictures of plants (about 98%). This means that farmers and agricultural experts can find and stop crop losses early on. Our system can look at pictures of plant leaves in great detail, which helps in reducing loss. As we continue to refine and expand our capabilities, we remain dedicated to driving positive change in the agricultural sector and ensuring a resilient food supply for future generations. In conclusion, our project heralds a new era of precision agriculture, where technological innovation converges with agricultural expertise to tackle pressing challenges in food production. This helps protect crops and keep the world from running out of food. We are moving away from reactive farming and using new deep learning methods to stop plant problems. The power of ResNet-50 architecture and advanced deep learning techniques, we aim to revolutionize plant disease management and foster sustainable agricultural practices worldwide.

VIII. FUTURE DIRECTIONS

The future scope of my project is to enhance video footage for identifying plant diseases using a deep convolutional neural network,specifically ResNet-50.It involves combining machine learning methods and deep learning models, such as the Random Forest Classifier and ResNet-50, to achieve more accurate results in detecting various plant diseases. The system will also provide recommendations to farmers on which crops are most suitable based on soil conditions. Additionally, the project will be expanded into an e-commerce platform to sell plants, organic seeds, and other agriculture-related products online.

REFERENCES

[1] Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI), P-272*, 79-88.

[2] Barbedo, J. G. A. (2018). Factors influencing the use of deep learning for plant disease recognition. *Biosystems Engineering, 172*, 84-91.

[3] Chen, J., Hu, H., & Yang, J. (2024). Plant leaf disease recognition based on improved SinGAN and improved ResNet34. *Frontiers in Artificial Intelligence, 7*.

[4] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis.

[5] Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors, 17*(9), 2022.

[6] Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.

[7] Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing, 267*, 378-384.

[8] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey.

[9] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419.

[10] Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., & Echazarra, J. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture, 161*, 280-290.

[11] Rishiikeshwer, B. S., Shriram, T. A., Raju, J. S., Hari, M., Santhi, B., & Brindha, G. R. (2020). Farmer-friendly mobile application for automated leaf disease detection of real-time augmented data set using convolution neural networks. *Journal of Computer Science, 16*(2), 158-166.

[12] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience, 2016*, 3289801.

[13] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine- tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture, 161*, 272-279.

[14] Wallelign, S., Yalew, A., & Habte, A. (2018). Coffee plant disease identification using convolutional neural network based feature extraction. *International Journal of Computer Applications, 180*(47), 1-5.

[15] Zhang, S., Huang, W., & Zhang, C. (2019). Three-channel convolutional neural networks for vegetable leaf disease recognition. *Cognitive Systems Research, 53*, 31-41.