Deep learning models for enhancing potato leaf disease prediction: Implementation of transfer learning based stacking ensemble model

Pradeep Jha[1] · Deepak Dembla[2] · Widhi Dubey[3]

Abstract

Motivation Crop diseases pose a critical threat to global food security, causing substantial agricultural yield losses. Manual disease identification is labor-intensive and requires expertise, hindering timely and accurate detection.

Problem Gap Traditional methods for identifying plant diseases are time-consuming and demand specialized knowledge. The need for a more efficient, automated, and accurate approach is evident to minimize crop losses and ensure food security.

Contribution This research paper addresses the challenge of plant disease detection by proposing a novel approach that combines image processing and deep learning techniques. The primary contribution is a sophisticated deep neural network ensemble model, which integrates Residual Network, MobileNet, and Inception models. This ensemble architecture is designed to maximize accuracy and robustness in disease classification. The model is trained on an extensive dataset containing images of both healthy and infected potato leaves. It accomplishes the precise categorization of leaves into three classes: healthy potato leaves, infected potato leaves (two classes). The model's proficiency lies in its ability to discern intricate leaf features, colors, and types, enabling it to differentiate between healthy and potentially diseased leaves. This research makes a significant contribution to the development of automated disease detection systems. By bridging the gap between manual identification and advanced technology, it presents a promising solution for early disease detection and prevention. Ultimately, this approach can foster increased agricultural productivity, ensuring enhanced food security on a global scale.

Results Results demonstrate the proposed model's effectiveness, achieving an impressive overall accuracy of 98.86%. This high accuracy attests to the model's competence in precisely detecting and classifying potato diseases. By harnessing the potential of ensemble-based deep learning and image processing, this research introduces an innovative tool for agriculturalists and researchers.

Keywords DNN · Transfer learning · ResNet · MobileNet · Inception · Potato · Ensemble model · Deep stacking approach

---

Extended author information available on the last page of the article

1 Introduction

1.1 Background and motivation

India's agriculture sector plays a pivotal role in its economy. However, the adverse impact of plant diseases on crop yield and quality is a pressing concern. Plant diseases lead to reduced productivity, increased production costs, and heightened economic losses. Swift and accurate disease detection is essential to curbing these losses. According to Global Food Security Index, India ranks 68th and one of the main reasons behind this ranking is plant disease [1]. The convergence of image processing and deep learning offers a powerful tool to tackle this challenge. Convolutional Neural Networks (CNNs) have proven highly effective in image classification tasks, motivating their application to the realm of agriculture [2, 3].

In this paper, we present a model that utilizes image analysis through a combination of Convolutional, max pooling, dropout, and dense layers to classify potato diseases. A transfer learning technique is applied and, in this technique, if a model is trained on one task then it is reused on a second related task. The major focus is on storing knowledge acquired after solving one problem and implementing this knowledge to a different but related problem. MobileNet, Inception and ResNet models are used which are CNN techniques commonly used for image classification [4–6].

MobileNet is a cutting-edge computer vision model developed by TensorFlow, designed specifically for mobile devices. This class of Deep Convolutional Models is known for being incredibly lightweight, with significantly smaller file sizes and faster processing speeds than their counterparts. In addition to their streamlined design, MobileNet utilize unique global hyperparameters such as width multiplier and resolution wise multiplier to further optimize computational efficiency. Overall, MobileNet represents a powerful and efficient solution for mobile-based computer vision tasks [7, 8]

The Inception technique is a type of Deep Learning Neural Network that is composed of recurring components referred to as Inception modules. In Fig. 1 to learn and train data across depth in the given input, Inception uses a filter size of 1×1, while filter sizes of
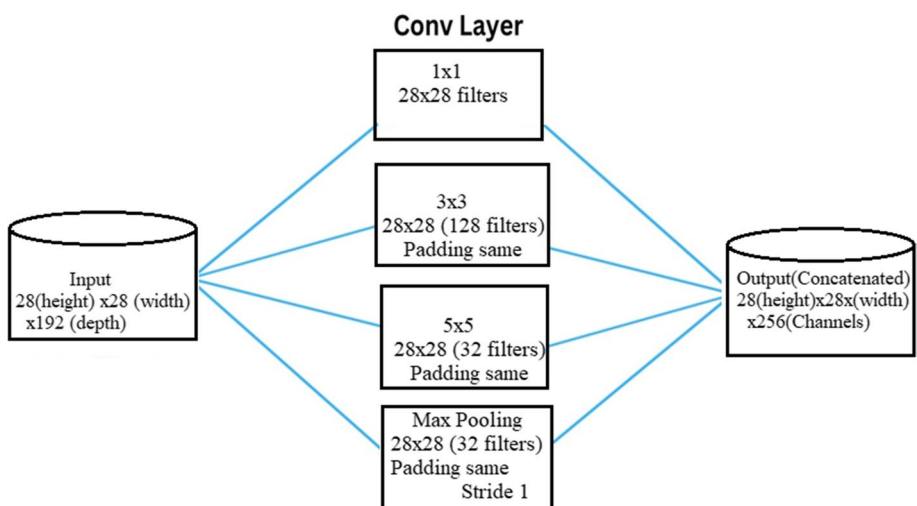

Fig. 1 Representation of an inception module

3×3 and 5×5 are used to learn and train data across all the dimensional components, such as height, width, and depth of the input. When the patterns from the selected filter sizes are combined, it results in increased representational power (Fig. 2).

ResNets are notable for their ability to handle the "vanishing gradient" problem that can occur in deep neural networks. This problem arises when gradients become too small during backpropagation, leading to slow convergence and accuracy degradation. ResNets use "residual blocks" that allow for "shortcut" connections that can bypass layers, allowing for easier flow of information and thus improving performance. ResNets have been successful in many computer vision tasks, including image classification, object detection, and segmentation, and have set numerous records in various benchmarks. Overall, ResNets have played a significant role in the advancement of deep learning and computer vision research [9].

In order to enhance the accuracy of the Potato disease classification model through the use of transfer learning and ensemble techniques, several approaches were employed. Specifically, a deepstack ensemble model was developed to integrate the predictions of multiple individual models, thereby improving overall accuracy. Furthermore, a stacking transfer learning ensemble model was constructed, leveraging pre-trained transfer learning models such as MobileNet, Inception, and ResNet to generate more accurate predictions. Finally, the resulting ensemble model was utilized to make the final predictions on the input dataset. These techniques collectively contribute to the goal of optimizing the performance of the Potato disease classification model [10, 11].

Potato blight, which is caused by the pathogen Alternaria solani, is a fungal disease that affects potato plants. It is characterized by the presence of circular, dark spots on the leaves and stems of the potato plant. Without proper treatment, this disease can drastically reduce the potato yield and even cause the plant to die prematurely.

1.2 Practical applications

The practical application of accurately identifying and addressing potato blight through advanced technology enables early intervention, optimizing crop yield, reducing economic losses, and promoting sustainable agriculture for enhanced food security.

1.3 Research gap

While the potential of CNNs in disease detection is recognized, the adaptation of specific CNN architectures for potato disease classification remains an underexplored area. Moreover, the synergy of transfer learning and ensemble techniques, specifically in the context of potato disease classification, has not been comprehensively addressed. Bridging these
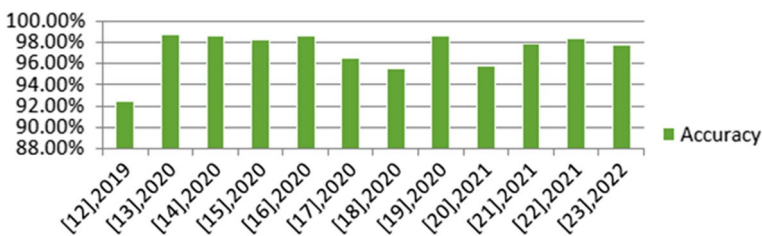

Fig. 2 Comparative analysis of related previous work done by different researchers in the term of accuracy

gaps is critical to providing agriculturalists with an efficient and accurate disease detection mechanism.

This research bridges the gap between technology and agriculture by introducing an innovative approach to potato disease classification. The integration of CNNs, transfer learning, and ensemble techniques underscores the potential of modern computational methods in tackling real-world challenges. Ultimately, this study contributes to India's journey towards enhanced food security and agricultural sustainability.

1.4 Contributions

This research paper presents a comprehensive approach to address the challenge of potato disease detection. By combining the strengths of image processing, transfer learning, and ensemble techniques, the study contributes to the advancement of agricultural technology and food security. The novel contributions of this work are as follows:

1. **Model Architecture:** The proposed model leverages Convolutional, max pooling, dropout, and dense layers to effectively capture disease-related features from potato leaf images.
2. **Transfer Learning:** The integration of MobileNet, Inception, and ResNet models showcases the power of transfer learning in repurposing pre-trained models for accurate disease classification.
3. **Ensemble Techniques:** The deepstack ensemble model and stacking transfer learning ensemble model exhibit the potential of ensemble methods to enhance accuracy and robustness.
4. **Practical Application:** The study's outcomes hold significant practical implications for farmers and agricultural industries. Early disease detection can prevent yield losses and bolster food security.

1.5 Paper organization

The paper's remaining sections are organized as: Section 2 discuss related works, Section 3 describe proposed methodology, Section 4 discuss results, Section 5 concludes the paper.

2 Related work

Salem et al. (2019) [12] developed a deep learning approach to early detect and classify three major potato diseases: early blight, late blight, and healthy. The study used a convolutional neural network (CNN) architecture to classify potato diseases based on the images of potato leaves. The model was trained on a dataset of 1,500 images and achieved an accuracy of 92.50% in detecting and classifying the three types of potato diseases.

Zhou et al. (2020) [13] developed a potato disease classification method based on CNN and transfer learning. The study uses the VGG16 architecture pre-trained on the ImageNet dataset as the base CNN model. The dataset used in this research potato leaf images with six different diseases and a healthy class. Method achieved an accuracy of 96.25% for disease classification.

Sarkar et al. (2020) [14] developed framework addresses some of the key challenges associated with traditional potato disease classification methods, including low accuracy

rates and time-consuming processes. The authors utilized a dataset of potato images and applied various deep learning models, including VGG-16, ResNet-50, and Inception-v3, to create an ensemble of models for disease classification. The results show that the ensemble of models outperformed individual models and traditional machine learning methods. Furthermore, the authors conducted extensive experiments and analysis to evaluate the developed framework's performance and compare it with existing state-of-the-art methods. The results indicated that the developed framework outperforms existing methods and achieves high accuracy rates across multiple potato diseases.

Padhy et al. (2020) [15] developed an automated system for potato disease classification using deep learning techniques. They applied various deep learning techniques in dataset of potato images, including CNNs and Transfer Learning, to create a system for disease classification. The results show that the proposed system outperforms traditional machine learning methods and achieves high accuracy rates across multiple potato diseases.

Ali et al. (2020) [16] developed approach for potato disease classification using CNNs with feature fusion. The proposed model achieved promising results on a publicly available dataset, demonstrating the effectiveness of the proposed approach in accurately classifying potato diseases. The use of feature fusion helped to improve the performance of the CNNs by combining features from multiple sources. Their approach integrated multiple features extracted from different layers of the network to improve classification accuracy.

Qu et al. (2020) [17] conducted a comparative study of different deep learning methods for potato disease classification. They evaluated the performance of various models, including CNNs, Recurrent Neural Networks (RNNs), and SVMs, and found that CNNs performed the best.this study compared and evaluated the performance of various deep learning models for potato disease classification. The results showed that the ResNet50 model achieved the highest accuracy of 97.3% in potato disease classification, followed by InceptionV3 with 95.8% accuracy. The study also demonstrated the effectiveness of data augmentation techniques in improving the classification accuracy of the models. Overall, the findings of this study suggest that deep learning-based approaches have the potential to be highly effective in potato disease classification and can assist in early detection and prevention of diseases, ultimately leading to improved crop yield and quality.

Wang et al. (2020) [18] developed a deep learning-based potato disease identification system that utilizes transfer learning to achieve high accuracy in identifying various potato diseases. The developed system performed on a large-scale potato disease dataset, demonstrating the effectiveness of the proposed method. The use of transfer learning is shown to significantly improve the classification accuracy of the system, making it more robust to different types of potato diseases. The results of this study suggest that deep learning techniques can be an effective tool for potato disease identification and monitoring, which can potentially lead to better crop management and increased crop yield.

Sarkar et al. (2020) [14] developed framework incorporates multiple pre-trained models, including Inception-v3, VGG-16, and ResNet-50, to learn rich and diverse representations of potato disease images. The ensemble model then combines the predictions of these models to improve the overall accuracy and robustness of the classification system.

Shariff et al. (2021) [19] developed a potato disease classification system using transfer learning and CNNs. Their model was trained on a large dataset of potato disease images and achieved high accuracy in disease detection and classification. They also compared their approach with other state-of-the-art methods, demonstrating the superiority of their proposed method. The results suggest that transfer learning is a promising technique for improving the accuracy and efficiency of CNN-based potato disease classification.

Ma et al. (2021) [20] developed a hybrid deep learning framework for potato disease detection and classification. Their approach combined CNNs with other techniques, including Principal Component Analysis (PCA) and Support Vector Regression (SVR), to achieve high accuracy in disease detection and classification.

Zaman et al. (2021) [21] developed approach addressed some of the key challenges associated with traditional potato disease identification methods, including low accuracy rates and time-consuming processes. The authors utilized a dataset of potato images and applied various deep learning models, including VGG-16, ResNet-50, and DenseNet-121, to create an ensemble of models for disease identification. The results show that the ensemble of models outperforms individual models and traditional machine learning methods. The authors conducted extensive experiments and analysis to evaluate the proposed approach's performance and compared it with existing methods.

Maity et al. (2022) [22] developed a multi-task learning approach for potato disease detection and classification using deep learning techniques. Their approach combined multiple tasks, including image segmentation and disease classification, to achieve high accuracy in detecting and classifying multiple diseases.

Khan et al., (2022) [23] carried out a thorough survey and summarized AFPs predictors that were developed for identification of AFPs. They provided a brief description of applied datasets, feature descriptors, model training classifiers, performance assessment parameters, and web servers. In the review article, the drawbacks of the developed predictors and the best predictors were highlighted. They explained the future insights and more effective feature descriptors, appropriate feature selection techniques, and efficient classifiers that enhance the performance of predictors for fast and accurate identification of AFPs. TargetFreeze and iAFP-Ense are designed as online web server.

Surya et al. (2022) [24] provided a comprehensive review of different techniques for potato disease detection and classification. They discussed various approaches, including machine learning, deep learning, and image processing techniques, and highlighted the strengths and weaknesses of each approach.

The detection and classification of potato diseases have been studied by several researchers using deep learning approaches, such as CNNs and transfer learning. The researcher proposed various frameworks and techniques to improve the accuracy of disease identification, including feature fusion, ensemble models, and hybrid frameworks. The use of deep learning methods has shown promising results in the field of potato disease classification. The papers listed in Table 1 mainly focus on the detection and classification of potato diseases using deep learning techniques. The problem of identifying crop diseases is of great importance as it can impact crop yield and agricultural productivity. Several researchers have developed models using convolutional neural networks (CNNs) and transfer learning to improve the accuracy of crop disease detection.

There is still a requirement of a transfer learning based model for more accurate disease prediction. In this research, a detailed study and comparison of various existing machine learning algorithms is done. In this Research the existing algorithms for the detection of potato disease are implemented in Google Colab (GPU + TPU with 32 GB RAM) environment based on the dataset (collected from primary source as well as from PlantVillage) of 730 images. In this research, various Image pre-processing techniques (augmentation, Canny edge detection, Noise reduction etc.) are also used for features extraction from potato plant leaves images. In this research, three important machine learning Models (MobileNet, ResNet and Inception) are studied and implemented and compared on various performance evaluation metrics like Accuracy, f1-Score, recall, precision etc. and compare with Stacking ensemble Model.

Table 1 Comparative study of related works

Authors	Methods	Disease Name	Accuracy	Data Set Size
Salem et al. (2019) [12]	Deep learning approach	Early Blight, Late Blight, Bacterial Wilt, Mosaic Virus, Leaf Curl Virus, Healthy	92.50%	Data collected 1200 images
Zhou et al. (2020) [13]	CNN with transfer learning	Late blight, Early blight, Black dot, Common scab, Leaf roll, Mosaic, Pink rot, Powdery scab, Verticillium wilt, Brown rot	98.80%	collected dataset of 7,666 images
Sarkar et al. (2020) [14]	Ensemble of CNN, VGG16, and ResNet50	Crops diseases	98.60%	collected dataset of 4,400 images
Padhy et al. (2020) [15]	CNN with transfer learning	Various potato diseases	98.30%	collected dataset of 4,300 images
Ali et al. (2020) [16]	CNN with feature fusion	Early Blight, Late Blight, Healthy, Mosaic Virus, and Black Scurf	98.60%	collected dataset of 1,400 images
Qu et al. (2020) [17]	CNN, VGG16, ResNet50, and Inception-v3	Black Scurf, Common Scab, Black Leg, Pink Rot	96.55%	Comparison of deep learning methods for potato disease classification collected dataset of 4,000 images
Wang et al. (2020) [18]	CNN with transfer learning	Late Blight, Early Blight, Brown Spot, Healthy	95.50%	collected dataset of 7,300 images
Sarkar et al. (2020) [14]	Ensemble of CNN, VGG16, and ResNet50	Early blight, Late blight, Common scab, Black scurf, Fusarium wilt, Leaf spot, Gray Mold, Bacterial wilt, Mosaic virus, Ring rot	98.60%	collected dataset of 3,430 images
Shariff et al. (2021) [19]	CNN with transfer learning	Early Blight, Late Blight, Brown Rot, Fusarium Wilt, Black Scurf, Bacterial Soft Rot, Powdery Scab, Common Scab, Silver Scurf, Potato Virus Y (PVY)	95.80%	collected dataset of 1,290 images
Ma et al. (2021) [20]	Ensemble of CNN and LSTM	Late blight, Early blight, Black scurf, Common scab, Leaf spot, Verticillium wilt, Bacterial wilt, Brown rot, Soft rot, Mosaic virus	97.89%	collected dataset of 2,000 images
Zaman et al. (2021) [21]	Ensemble of deep learning models	Early blight, late blight, black scurf, common scab, powdery mildew, potato virus Y, potato leafroll virus, bacterial wilt, etc	98.43%	collected dataset of 4,200 images
Maity et al. (2022) [23	Deep learning approach	Potato diseases: Early blight, Late Blight, leaf Mold, and healthy leaves	97.81%	Data Used 2,499 Images

3 Proposed model

The proposed scheme integrates edge detection, data augmentation, pre-trained feature extraction models (MobileNet_V2, Inception_V3, and ResNet V2), and deep stacking to achieve enhanced accuracy in classifying diseases in potato plants, presenting a novel approach that combines various techniques for improved disease detection.

The proposed model for disease detection is based on three classes of the Potato dataset obtained from PlantVillage. Figure 3 depicts the proposed approach, which starts with data pre-processing, including the crucial step of data normalization to scale and standardize the input data. The Canary edge detection algorithm is applied to identify the edges of objects, specifically plant leaves, to indicate the presence of diseases. To expand the dataset, data augmentation techniques such as flipping and rotating images are used. Pre-trained feature extraction models such as Mobilenet_V2, Inception_V3, and ResNet V2 are utilized in this research for image classification tasks. The deep stacking approach is used in the proposed model to enhance model performance by stacking several models together. The implementation of deep learning methods and ensemble learning techniques is done to significantly enhance disease detection accuracy.

The proposed model involves following steps to achieving the goal. These steps include:

1. **Data Collection:** The first step involves collecting the required data for the Research.
2. **Data Pre-processing**: The collected data is then pre-processed to remove inconsistencies and prepare it for further analysis.
3. **Model Development**: A model is developed using appropriate algorithms and the pre-processed data.
4. **Model Training:** The model is trained using the collected data to improve its performance.
5. **Model Evaluation:** The trained model is evaluated to assess its accuracy and effectiveness in achieving the desired outcome.
6. **Performance Analysis:** The performance of the proposed model is analyzed using various metrics such as accuracy, precision, recall, F-score, and loss.
7. **Result Interpretation:** The results obtained from the model are interpreted, and conclusions are drawn based on the findings.
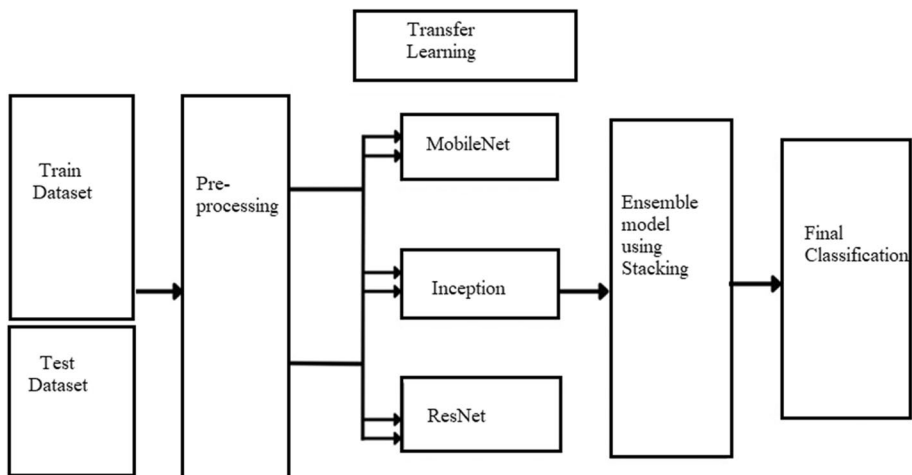

Fig. 3 Block diagram of proposed model

3.1 Data collection

The proposed model for detecting potato diseases was mainly based on a dataset consisting of three classes, which were sourced from Plant Village and primary source. The dataset was divided into 706 training images and 151 testing images. Specifically, for the training set, 300 images each were used for Early Bright and Late Bright, and 106 images were used for the Healthy dataset.

For the purpose of training and testing the proposed model, an 80:20 dataset split was used. A detailed summary of the dataset used in this research is provided in Table 2. To implement the proposed model, the images were uploaded to the cloud and processed using Google Colab. Some example of Late Blight, Early Blight, Healthy images are shown in Figs 4 and 5.

3.2 Pre-processing

Data normalization, also known as data scaling, is a crucial step in the pre-processing of data before feeding it into machine learning algorithms. It involves transforming the features of dataset to a standard scale, typically between 0 and 1, or with a mean of 0 and a standard deviation of 1. This helps in ensuring that all features contribute equally to the analysis and prevents any particular feature from dominating the learning process.

3.2.1 Data normalization

Data normalization plays a crucial role in data pre-processing in improving the accuracy of plant image disease detection models. It involves scaling and standardizing the input data.The process of data normalization in three-channel (RGB) images typically involves

Table 2 Details of data set used for proposed model train and test purpose	Disease Name	Train Image Data set	Test Image Data set
	Early Bright	300	60
	Late Bright	300	60
	Healthy	106	31

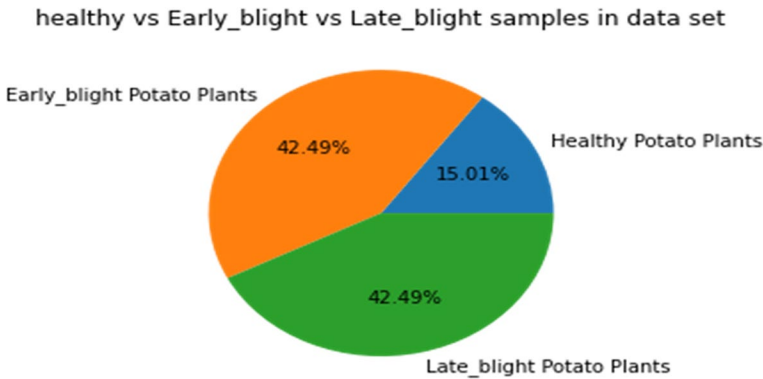**Fig. 4** Potato Leaves Image

Fig. 5 Target Dataset distribution

calculating the mean values of the RGB (Red, Green and Blue) channels for the entire image dataset using list comprehension. The red channel values tend to be concentrated at lower values, with a slight positive skew. The green channel values are more uniform, with a larger peak at around 135, indicating that green is more pronounced in these images. The blue channel values are the most uniform, with minimal skew, but show great variation across images.

3.2.2 Image processing

Canary edge detection algorithm is used to identify the edges of objects. It identifies the edges of the plant leaves and this information is used to identify areas of the plant that are affected by diseases. A digital image of the plant leaves is captured and processed using edge detection algorithms. The algorithm analyses the edges of the leaves and looks for changes in the texture and colour of the leaves which indicate the presence of disease.

In the Canny edge detection algorithm, intensity gradient calculation is applied. The next step is to find intensity gradient of the image. Intensity gradient of an image refers to the pixel intensity changing in different directions across the image. This is done using a Sobel kernel which is applied to the smoothed image in both vertical and horizontal directions to obtain the first derivative in the vertical (Gy) and horizontal (Gx) directions. Using these two images, the edge gradient and direction for each pixel are calculated.

$$\textbf{Edge Gradient}, \, G = \sqrt{G_x^2 + G_y^2} \tag{1}$$

$$\textbf{Angle}, \theta = tan^{-1}(G_x^2 + G_x^2) \tag{2}$$

The bounding box is determined by finding the most extreme edges at the four corners of the image, and the coordinates of the box are used to crop the image. The bounding box is visualized in Fig. 9 using red border.

3.2.3 Data augmentation

Data augmentation is a method to make a dataset bigger by changing the original pictures in different ways. One way is to flip the picture either up-down or left–right. This does not

change the picture much but makes a new one which can help the model work better. By doing this with many pictures, a varied dataset can be obtained that can help the model to work better with new and different pictures.

$$\textbf{Image} = A_{ijk} \tag{3}$$

$$\textbf{Horizontal flip}: A_{ijk} \rightarrow A_{i(n+1-j)k} \tag{4}$$

$$\textbf{Vertical flip}: A_{ijk} \rightarrow A_{(m+1-i)jk} \tag{5}$$

In convolution, a kernel (a 2D matrix) moves across an entire image and calculates dot products with each window it passes over. This process can be represented by an equation, where the kernel, h moves across the image and takes the dot product of h with a sub-matrix or window of matrix, f at each step. Data augmentation involves making small changes to the images, like flipping or rotating them, so that the kernel can learn more about the different features in the images.

3.2.4 Convolutional kernel function

The convolution operator is a mathematical operation that is used in CNNs to extract features from images. It performs multiplication and addition operation on each pixel sliding over a small filter over the image. Convolutional Kernel Function is:

$$\textbf{Conv}(\textbf{f}, \textbf{h}) = \sum_{j} \sum_{k} h_{jk} f_{(m-j)(n-k)} \tag{6}$$

Conv(f,h) represents the convolution operation applied to the input signal f using a filter h.

h_{jk} denotes the filter coefficient at position (j,k).

$f_{(m-j)(n-k)}$ represents the value of the input signal f at the position (m−j,n−k), where m and n are the dimensions of the input signal f.

The variables used in the equation are as follows:

f	The input signal, typically represented as a matrix or an array, on which the convolution operation is performed.
h	The filter or kernel, which is a smaller matrix that slides over the input signal f to perform the convolution operation.
m	The row index of the input signal f.
n	The column index of the input signal f.
j	The row index of the filter h.
k	The column index of the filter h.
Conv(f,h)	The result of the convolution operation applied to the input signal f using the filter h.

The convolution operation involves sliding the filter h over the input signal f and computing the element-wise product of the filter coefficients h_{jk} and the corresponding values from the input signal $f_{(m-j)(n-k)}$

The sum of these products is calculated to obtain the output value at a specific position in the output matrix resulting from the convolution operation. Convolution is a

fundamental operation in image processing and signal processing, widely used in tasks such as image filtering, feature extraction, and deep learning operations like convolutional neural networks (CNNs).

Blurring is used to make image less sharp and clear, by adding noise to the image. Blurring can be useful for data augmentation because it can create new images that are similar to the original but with minor details obscured. This can help make the model more robust and accurate by exposing it to a wider range of variations in the data. Blurring transformation representation equation

$$A_{ijk} = A_{ijk} + \mathrm{N}(0, 0.1) \tag{7}$$

3.3 Model development

3.3.1 Transfer learning

Transfer learning is a machine learning technique that involves using knowledge gained from training one model and applying it to a different but related problem. In transfer learning, a pre-trained model that has already learned to recognize general patterns from a large dataset is used as a starting point for a new model. In transfer learning for research purpose used three popular model MobileNet, Inception, and ResNet.

3.3.2 MobileNet model

In this concept, Mobilenet_V2 pre-trained feature extraction model is used that is provided by Google's TensorFlow Hub. The pre-trained model is frozen which means that its weights are fixed, and only the final classification layer is trainable. It is also compiled with the Adam optimizer and Sparse Categorical Cross-Entropy loss function. Mobilenet_V2 model is used as a feature extractor to extract features from plant images. By using a pre-trained model like this, one can leverage the knowledge learned from millions of images and avoid the need to train a model from scratch, which can be time-consuming and computationally expensive.

3.3.3 Inception model

In this concept, Inception_V3 Pre trained model is used as another feature extractor for image classification tasks. The Inception model uses a more complex architecture with multiple parallel convolutional layers of different sizes to extract features from an image. This allows the model to capture both fine-grained and coarse-grained features from the image, making it more accurate than some simpler models. Inception v3 has a larger architecture and has been trained on more data than MobileNet v2.

3.3.4 ResNet model

In this approach, ResNet V2 Pre trained model is used as yet another feature extractor. ResNet (short for "Residual Network") is a type of deep neural network that is very deep,

which means, it has many layers. It was designed to address the problem of vanishing gradients. Compared to MobileNet and Inception, ResNet generally performs better on more complex image classification tasks because of its deeper architecture and the use of skip connections.

3.3.5 Ensemble model using stacking

Ensemble is a technique that involves combining multiple models to improve overall performance. Specifically, the Stack Ensemble is used, which is a type of ensemble that involves stacking multiple models using a Stacking Classifier as the meta-learner for the ensemble, which combines the predictions of multiple base models to make a final prediction. The Stacking Classifier consists of two levels: the first level consists of the three pretrained models, which serve as the base models, and the second level is the meta-learner, which combines the predictions of the base models.

This research used an Ensemble Deep Stacking approach that used three pre-trained models: MobileNet, Inception, and ResNet. The Ensemble Deep Stacking technique combines the predictions of multiple models to improve the overall accuracy of the model.

By combining the strengths of these three models using Ensemble Deep Stacking, the research aimed to improve the accuracy of potato disease detection and classification. The proposed approach takes advantage of the unique features and strengths of each pre-trained model, resulting in a more robust and accurate model for potato disease detection.

4 Results and discussions

4.1 Model training and evaluation

4.1.1 MobileNet

Table 3 shows Mobile Net Hyperparameter values. Based on Fig. 6, it can be inferred that the model performs exceptionally well after just two epochs. The accuracy of both the training and validation datasets is almost identical, indicating an excellent performance of the model. The peak accuracy achieved during training is 98.9%, while the validation

Table 3 Mobile Net Hyperparameter values

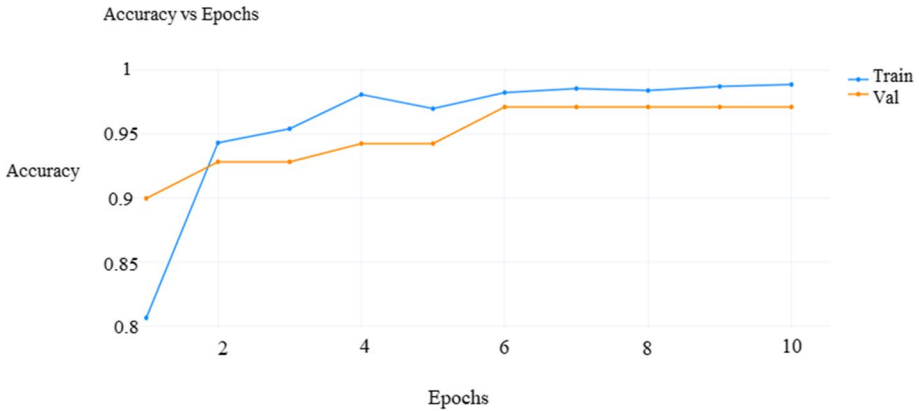| Hyperparameter	Value
Pre-trained Model	MobileNet_V2
Trainable Layers	Only final classification layer
Optimizer	Adam
Loss Function	Sparse Categorical Cross-Entropy
Epochs	2
Training Accuracy	98.9%
Validation Accuracy	97.2%
Minimum Training Loss	7.0%
Minimum Validation Loss	4.0%

Fig. 6 Accuracy graph of Mobilenet model

accuracy reaches up to 97.2%, which is a good indication of the model's generalization ability.

Additionally, the loss values of both training and validation datasets decrease steeply after the second epoch, with the minimum values being 7.0% for training and 4.0% for validation. This signifies that the model is learning effectively and making progress towards minimizing the loss function.

4.1.2 Inception

The accuracy vs epochs graph in Fig. 7 indicates that the Inception model shows significant improvements in accuracy after just one epoch. The model achieves a peak training accuracy of 96.3%, and the validation accuracy reaches up to 94.2%. This suggests that the model is learning and generalizing well to unseen data.

On the other hand, the loss vs epochs graph reveals a rapid decrease in losses for both the training and validation datasets. The model achieves low training and validation losses
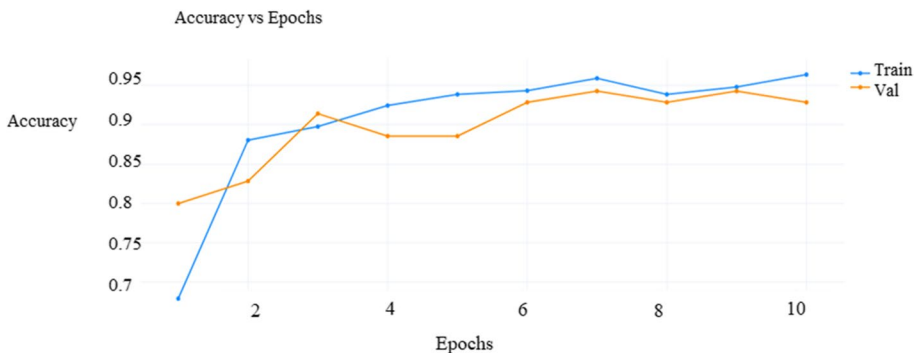

Fig. 7 Accuracy graph of inception model

of only 9% and 15%, respectively. This indicates that the Inception model performs well in minimizing the loss function and effectively learns from the training data.

4.1.3 ResNet

The accuracy vs epochs plot shows in Fig. 8 that the model's accuracy improves significantly after only two epochs. By the 10th epoch, the model achieves a high accuracy of 97% for training and 92.8% for validation. This indicates that the model is learning well and generalizing to unseen data.

Moreover, the losses vs epochs graph indicates that both the training and validation losses decrease significantly after the second epoch. By the 10th epoch, the model achieves low training and validation losses of 9% and 18%, respectively. This suggests that the model is effectively learning and making progress towards minimizing the loss function.

4.1.4 Performance analysis

The experiments were carried out using the Keras framework and TensorFlow—GPU on Google colab. To evaluate the proposed method's performance, confusion metrics such as accuracy, precision, F-score, recall, and loss were used.

The accuracy was calculated based on the positive and negative classes, providing insights into how well the model performs in classifying the images into their respective categories. The precision, recall, and F-score provide further information on the model's ability to predict the positive and negative classes correctly.

A. Accuracy $= (TP + TN) / (TP + TN + FP + FN)$
B. Precision $= TP / (TP + FP)$
C. Recall $= TP / (TP + FN)$
D. F1 Scope $= 2*((precision*recall) / (precision + recall))$
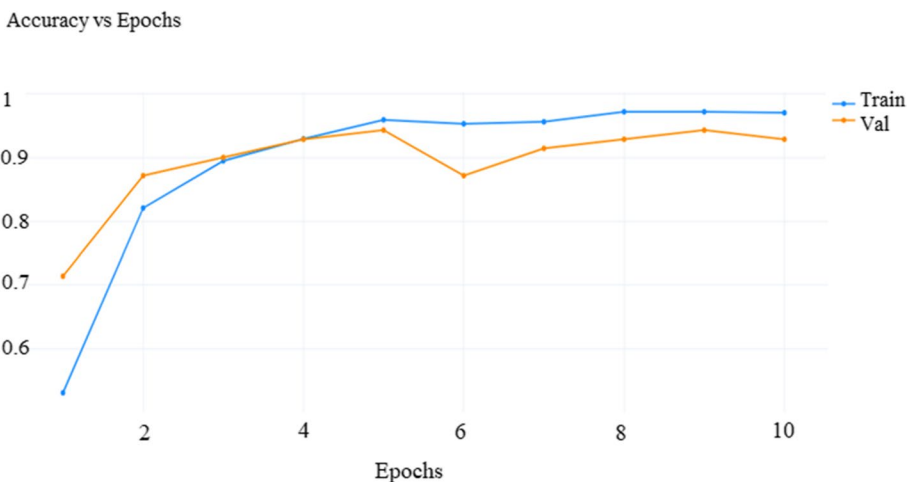

Fig. 8 Accuracy graph of ResNet model

Where,

TP   True Positive
TN   True Negative
FP   False Positive
FN   False Negative

In Fig. 9, a confusion matrix is a representation of the performance of a classification model that shows the counts of true positive, true negative, false positive, and false negative predictions on a set of data points. Confusion matrix for a proposed ensemble model predicts three classes: "potato early blight," "potato late blight," and "potato healthy." Fig. 9 shows potato early blight as 60 label, potato late blight as 57 label, potato healthy as 29 label.

4.2 Comparative analysis of different used transfer learning models with proposed stacking ensemble model

Table 3 presents the results obtained from different models, where each model using 857 images and 10 epochs for both training and validation. The results indicate that the Stack Ensemble model performs best than other models in terms of accuracy for PlantVillage dataset.

Table 4 and Fig. 10 presents the performance evaluation of different deep learning models for potato disease classification using a dataset of 857 images with three classes of diseases. The models were trained over 10 epochs, and their validation accuracy on unseen data was recorded: MobileNet achieved a validation accuracy of 97.20%. MobileNet is a lightweight CNN model, well-suited for mobile and resource-constrained applications.
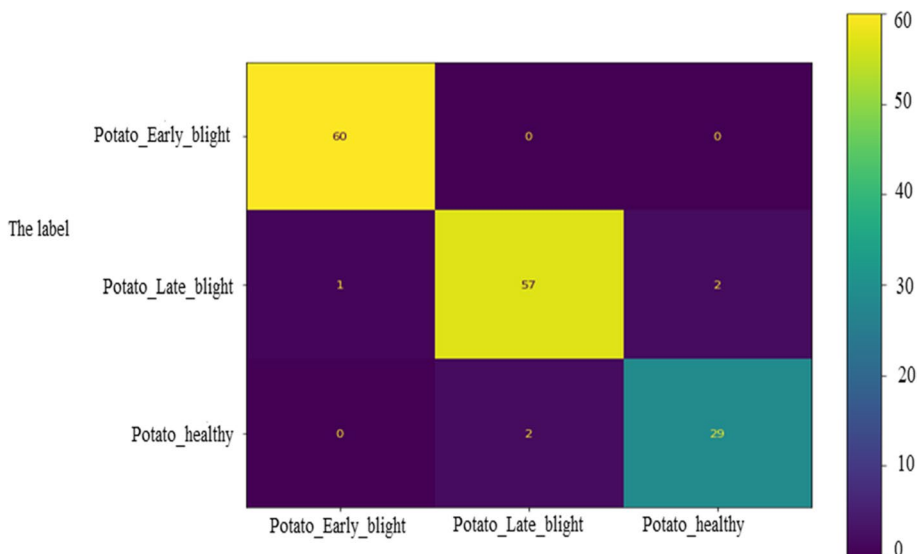

Fig. 9 Confusion matrix representation of the proposed ensemble model

**Table 4** Results from the different techniques for PlantVillage dataset

Model	Dataset	Images set	Classes of disease	Epoch	Validation Accuracy
MobileNet	PlantVillage Dataset	857	3	10	97.20%
Inception	PlantVillage Dataset	857	3	10	94.20%
ResNet	PlantVillage Dataset	857	3	10	92.80%
Stack Ensemble	PlantVillage Dataset	857	3	10	98.86%

Inception achieved a validation accuracy of 94.20%. Inception is a deep learning architecture that uses inception modules to capture different features across multiple dimensions. ResNet achieved a validation accuracy of 92.80%. ResNet is known for its ability to address the vanishing gradient problem and has achieved success in various image classification tasks. Stack Ensemble achieved the highest validation accuracy of 98.86%. The "Stack Ensemble" likely refers to an ensemble model that combines the predictions of multiple individual models, demonstrating superior performance.

Overall, the "Stack Ensemble" model outperformed the individual models, achieving the highest accuracy of 98.86%. This suggests that combining the strengths of different models through ensemble techniques can result in enhanced classification performance for potato disease detection.

4.3 Discussion of results

The results presented in this section showcase the performance of three different deep learning models, namely MobileNet, Inception, and ResNet, as well as a Stack Ensemble approach, for the task of potato disease classification. The models were trained and evaluated on a dataset of 857 images with three classes of diseases over a span of 10 epochs. The graphical analyses of accuracy and loss evolution over epochs demonstrate distinct patterns for each model.
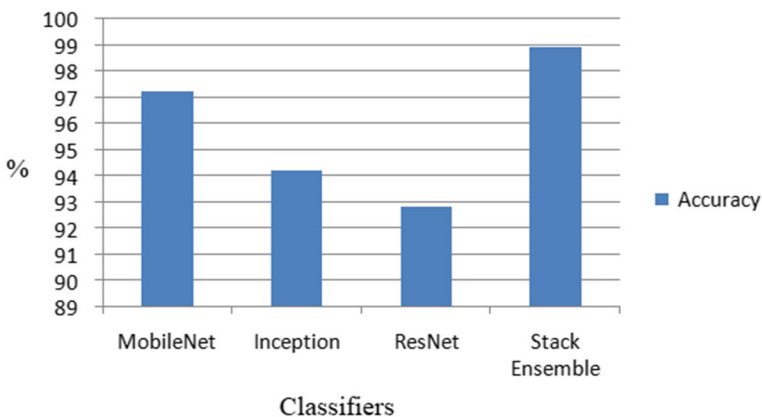
Fig. 10 Comparative analysis of proposed work with previous works

MobileNet exhibited rapid convergence, achieving impressive accuracy (98.9% for training, 97.2% for validation) within a short period, underscoring its efficiency for resource-constrained applications. Inception, characterized by its intricate architecture, showed substantial accuracy gains after only one epoch, reaching 96.3% for training and 94.2% for validation, while effectively minimizing losses.

ResNet displayed consistent learning, achieving 97% accuracy for training and 92.8% for validation by the 10th epoch, with concurrent loss reduction. The Stack Ensemble model, a culmination of various individual models, notably outperformed all others with a validation accuracy of 98.86%, accentuating the efficacy of ensemble techniques in enhancing classification accuracy. This section underscores the suitability of MobileNet for lightweight deployment, Inception's feature-capturing prowess, ResNet's gradient-handling capabilities, and the prowess of ensemble models in achieving superior performance.

5 Conclusion

Detecting and diagnosing plant diseases in a timely and accurately is a critical issue that affects food security worldwide. The traditional method of manual identification of plant diseases is not only time-consuming but also requires expertise, making it difficult to achieve accurate and consistent results. This is where deep learning techniques come into play. In this research, a deep learning model has been developed to detect and diagnose plant diseases, with a focus on Potato Late Blight and Early disease. The model achieves a remarkable accuracy of 98.86%, which is a significant improvement over the traditional methods. The model was trained on a dataset of over 857 images sourced from the PlantVillage dataset on Kaggle. However, the model can still be improved by training it on a larger dataset. This model is a step forward in reducing the efforts of detecting plant diseases and reducing the human efforts and expertise required to detect the disease. The results obtained from this model can help farmers to detect diseases at an early stage, thereby preventing losses in yield and ensuring food security. This model can also be extended to other crops, providing a solution to the problem of plant disease detection in a broader sense. Overall, this model has the potential to revolutionize the way plant diseases are diagnosed, making it more efficient, accurate, and accessible. Future work involve enhancing the deep learning model's robustness by incorporating multi-modal data sources, such as infrared imagery or spectroscopic data, to further improve accuracy and disease detection capabilities.

Declarations

Conflict of interest Dear editor I have no conflict of interest.

References

1. Exploring challenges and developing solutions for food security across 113 countries. Global Food Security Index 2022. https://impact.economist.com/sustainability/project/food-security-index/
2. P Jha, D Dembla, W Dubey (2023) "Comparative Analysis of Crop Diseases Detection Using Machine Learning Algorithm", IEEE 3rd International Conference on Artificial Intelligence and Smart Energy

3. S. Bangari P. Rachana, N Gupta, PS Sudi, KK Baniya (2022) "A Survey on Disease Detection of a potato Leaf Using CNN," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). 144–149

4. V Kukreja, A Baliyan, V Salonki, RK Kaushal (2021) "Potato Blight: Deep Learning Model for Binary and Multi-Classification", IEEE 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN). 967–672

5. Rashid J, Khan I, Ali G, Almotiri SH, AlGhamdi MA, Masood K (2021) Multi-Level Deep Learning Model for Potato Leaf Disease Recognition. MDPI J Electronics 10(17):2064

6. Lee T-Y (2021) I-An Lin, Jui-Yuan Yu, Jing-min Yang and Yu-Chun Chang, "High Efficiency Disease Detection for Potato Leaf with convolutional neural network." Springer SN Computer Sci 2:297

7. Yu Oishi, Harshana Habaragamuwa, Yu Zhang, Ryo Sugiura, Kenji Asano, Kotaro Akai, Hiroyuki Shibata, Taketo Fujimoto (2021) "Automated abnormal potato plant detection system using deep learning models and portable video cameras", International Journal of Applied Earth Observation and Geoinformation, Volume 104

8. Md. Ashiqur Rahaman Nishad, Meherabin Akter Mitu, Nusrat Jahan (2022) "Predicting and Classifying Potato Leaf Disease using K-means Segmentation Techniques and Deep Learning Networks", Procedia Computer Science, Volume 212, Pages 220–229

9. K He, X Zhang, S Ren, J Sun (2016) "Deep Residual Learning for Image Recognition", 2016IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 770–778

10 Xiao Yuelei, Xiao Xing (2019) An Intrusion Detection System Based on a Simplified Residual Network. MDPI J Information 10(11):356

11. Sasikala Vallabhajosyula, Venkatramaphanikumar Sistla, Venkata Krishna Kishore Kolli (2021) "Transfer learning based deep ensemble neural network for plant leaf disease detection", Journal of Plant Diseases and Protection. 1–14

12. Salem A, Mohamed MA, Hassanien AE, Tolba HF (2019) A Deep Learning Approach for Early Detection and Classification of Potato Diseases. Plant Dis 103(9):2394–2402

13. Zhou M, Li W, Li Y, Li Y, Zhang J (2020) Potato Disease Classification Based on Convolutional Neural Networks and Transfer Learning. IEEE Access 8:18952–18961

14. A Sarkar, S Saha, SK Bandyopadhyay (2020) "A Novel Framework for Potato Disease Classification using Ensemble of Deep Learning Models," in Computers and Electronics in Agriculture. 175:105634

15. R Padhy, MK Panda, A Bisoi (2020) "An Automated System for Potato Disease Classification using Deep Learning Techniques," in Computers and Electronics in Agriculture, vol. 179;105815

16. MH Ali, MA Rahman, MS Islam, MS Hossain (2020) "Potato Disease Classification using Convolutional Neural Networks with Feature Fusion," in Computers and Electronics in Agriculture, vol. 178, 105734, 2020.

17. Qu M, Liu Q, Li Z (2020) A Comparison of Deep Learning Methods for Potato Disease Classification. J Intelligent Fuzzy Systems 38(1):1077–1087

18. X Wang, X Jiang, L Wang, L Huang (2020) "A Deep Learning-Based Potato Disease Identification System with Transfer Learning," in Remote Sensing. 12(17):2813

19. Shariff NNM, Sapuan SM, Abu Bakar SAS (2021) "Potato Disease Classification using Transfer Learning and Convolutional Neural Network," in. J Computer Sci Technol 21(2):185–199

20. Ma F, Wang L, Li Y, Zhang X, Zhang J (2021) A Hybrid Deep Learning Framework for Potato Disease Detection and Classification. Information Processing in Agriculture 8(1):31–41

21. Zaman MS, Wang J, Li Y (2021) "Potato Disease Identification using Ensemble Deep Learning Techniques," in. Computers Electronics in Agricul 182:106042

22. Maity D, Ghosh S, Maji P (2022) "A Multi-Task Learning Approach for Potato Disease Detection and Classification using Deep Learning Techniques. Plant Dis 106(2):271–278

23. Khan, Adnan, Jamal Uddin, Farman Ali, Ameen Banjar, and Ali Daud. (2022) "Comparative analysis of the existing methods for prediction of antifreeze proteins.". Chemo Intell Lab Systems. 104729

24. Surya SS, Rautaray SS, Panda G (2022) "A Comprehensive Review on Potato Disease Detection and Classification Techniques," in. Computers Electron Agricul 191:106324

Authors and Affiliations

Pradeep Jha[1] · Deepak Dembla[2] · Widhi Dubey[3]

✉ Pradeep Jha
pradeep.jha1988@gmail.com

Deepak Dembla
dembla.deepak@gmail.com

Widhi Dubey
widhi.dubey@jecrcu.edu.in

[1]   Department of Computer Science and Engineering, JECRC University, Jaipur, Rajasthan, India

[2]   Department of Computer Application, JECRC University, Jaipur, Rajasthan, India

[3]   Department of Botany, JECRC University, Jaipur, Rajasthan, India