COTTON PLANT DISEASE DETECTION

INTRODUCTION AND ALGORITHM

Detecting cotton plant diseases is crucial for agriculture to protect crop health and yield. Early disease detection enables timely interventions like pesticide use or irrigation adjustments. Deep learning methods have gained popularity for their accuracy and efficiency in this task.

Cotton plant disease detection relies on deep learning models like ResNet-50, Inception, and VGG19 due to their pretrained feature extraction and transfer learning capabilities. These models capture critical leaf characteristics for accurate disease identification. Ensemble learning, specifically stacking, is favored for cotton plant disease detection as it combines predictions from multiple models, enhancing generalization, robustness, and accuracy by leveraging their diverse strengths and mitigating individual model errors.

THE DATASET

Fresh Cotton Plant & Fresh Cotton Leaf

Disease Cotton Plant & Disease Cotton Leaf

The cotton plant disease detection dataset is a collection of images capturing various diseases and health conditions affecting cotton plants. It comprises labeled images of cotton leaves with symptoms like spots, discoloration, and deformities, as well as healthy cotton leaves.

PRETRAINED MODEL TRANSFER LEARNING AND FINE TUNING

ResNet-50: ResNet-50 is a deep convolutional neural network known for its residual architecture, which helps in training very deep networks. It performs well in a wide range of image classification tasks and is often used as a strong baseline.

InceptionV3: InceptionV3, also known as Google Net, is famous for its inception modules that capture features at multiple scales. It's designed to be computationally efficient and can handle large image resolutions effectively.

VGG19: VGG19 is characterized by its simple and uniform architecture, with 19 weight layers. Although it's not as deep as some other models, it's known for its simplicity and strong performance in various image recognition tasks.

Transfer learning uses pretrained neural network models to accelerate training for new related tasks. It leverages learned features, reduces training time, and improves generalization by **fine-tuning** the model's specific task layers while keeping the pretrained layers fixed.

ResNet-50 achieved a final test accuracy of 77.78% by employing data augmentation and custom classification layers in a 20-epoch training with the Adam optimizer. InceptionV3 attained a final validation accuracy of 88.89% with similar settings but using a larger image size (299x299 pixels). VGG19 outperformed the others, achieving a superior final validation accuracy of 94.44% with the same training parameters and a standard image size of 224x224 pixels.

Among the pretrained models used for cotton plant disease detection, VGG19 performed the best with a final validation accuracy of 94.44%. ResNet-50 achieved a final test accuracy of 77.78%, while InceptionV3 reached a final validation accuracy of 88.89%. VGG19's superior performance makes it the preferred choice.

ENSEMBLE LEARNING ALGORITHM

Ensemble learning is a machine learning technique that combines the predictions of multiple models (learners) to improve overall predictive accuracy and robustness. By leveraging the diverse strengths of individual models, ensemble methods aim to achieve better results than any single model alone. Common ensemble techniques include bagging, boosting, and stacking.

Individual Model Prediction: Three pretrained models (ResNet-50, InceptionV3, VGG19) are loaded, and predictions are made on a validation dataset of cotton plant images.

Stacking Predictions: The predictions from these individual models are stacked horizontally into a single matrix, creating a feature set where each column corresponds to a model's prediction for a given data point.

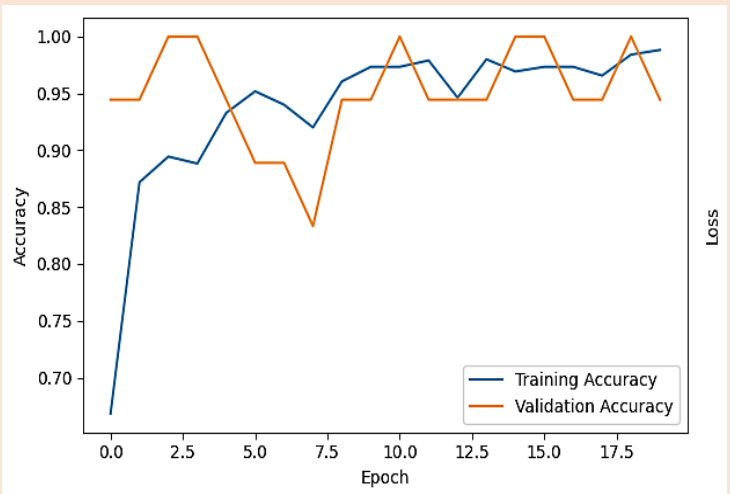
Final Model: A final model, in this case, a Logistic Regression classifier, is trained on the stacked predictions. It learns how to combine the individual model predictions to make a final prediction.

Ensemble Prediction: The final model is used to make predictions on the stacked data, creating an ensemble prediction.

Accuracy Evaluation: The accuracy of the ensemble prediction is calculated and printed. This represents the performance of the stacked ensemble in comparison to individual models.

In the previous code, individual pretrained models (ResNet-50, InceptionV3, and VGG19) were evaluated for cotton plant disease detection. Each model achieved specific accuracies, with the highest being 94.44% obtained by VGG19. In the new code, an ensemble learning technique known as stacking was applied. This involved combining predictions from the three models and using a logistic regression model as the final classifier. The stacking ensemble achieved an accuracy of 94.44%, matching the performance of the best individual model, VGG19.

GRAPH



Vgg19 Pretrained model Accuracy

CONCLUSION

In cotton plant disease detection, pretrained deep learning models like ResNet-50, InceptionV3, and VGG19 play a vital role due to their ability to capture essential leaf features. VGG19 stands out with the highest accuracy of 94.44%. Ensemble learning, particularly stacking, further enhances accuracy by combining predictions from these models, resulting in an ensemble accuracy of 94.44%. Among the individual models and ensemble methods, VGG19 and stacking prove to be the most effective choices for accurate cotton plant disease detection.

Course – Artificial Intelligence and Machine Learning (Honor/Minor)
Guide – Prof. Diptee Ghusse

Presented by

Aniket Bhavar – ETX – B236023
Ashish Gaikwad – Mech – B239107