DEEP LEARNING-BASED DISEASE PREDICTION IN APPLE AND GRAPE LEAVES

Sunil Kumar Balla

CAP 6545 Machine Learning Methods for Biomedical Data

ABSTRACT

Predicting diseases in plant leaves plays a key role in agricultural sustainability. As the diseases vary from leaf to leaf of different plants, traditional methods have trouble in correctly identifying and predicting diseases in apple and grape leaves. Deep learning techniques are used in this work to help solve these problems. Using transfer learning, deep learning models like VGG16 and ResNet18 offer a promising answer by using weights that have already been trained on large datasets. This method gets around the problems with older ones by making it possible to pull out complex patterns and features from pictures of leaves. Using these deep learning models not only improves the accuracy of predictions, but it also makes it easier to find diseases quickly and easily. Using deep learning to predict disease in apple and grape leaves works, as shown by this study. This shows the potential for scalable and reliable agricultural disease management systems.

1. INTRODUCTION

The global food supply chain is an important part of agriculture, which is a key part of global businesses. Diseases that affect plants are major threats to crop yield, which has a huge effect on the agricultural industry. The effects go beyond local crops and affect economies around the world. For example, when the output of staple crops is interrupted, food imports and exports often change, which affects prices and availability on markets around the world.

Recent progress in technology, especially in machine learning, shows scope for improving the effectiveness and durability of farming methods. Deep learning methods, especially Convolutional Neural Network (CNN) architectures, have shown a lot of scope for changing the way plant diseases are found.

To solve the problem of finding diseases in apple and grape leaves, this project focuses on using CNN designs, especially VGG16 and ResNet18. Transfer learning, which uses models that have already been pretrained and then tweaked for specific tasks, is one of the most important parts of this project. VGG16 and ResNet18, which are well-known for their performance in transfer learning tasks, are used to find and label complex disease-related patterns in pictures of plant leaves. This makes disease prediction models more accurate and useful.

By combining new developments in machine learning, especially in CNN structures and transfer learning methods, this study aims to make a difference in the creation of strong and scalable ways to predict disease in apple and grape leaves.

2. DATASET AND RESOURCES

The dataset has more than 50,000 leaves images of both healthy and sick leaves from different types of plants. These pictures come from the online site PlantVillage and are a good starting point for using machine learning and crowdsourcing to help crop plants that are losing their yields because of diseases. The dataset can be downloaded from Kaggle.

Dataset Link:

https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset

The dataset includes pictures from several different crop groups. Apple and grape leaves are taken from the dataset for this project. There are 4 groups of apple leaf images: Apple_scab has 630 images, Black_rot has 621 images, Cedar_apple_rust has 275 images, and Healthy has 1645 images. There are 4 types of grape leaf pictures. It has 1180 images for black rot, 1383 images for esca (black measles), 1076 images for leaf blight (Isariopsis leaf spot), and 423 image files for healthy leaves. Mobile apps can be developed which can diagnose diseases by using these images and machine learning techniques. These tools could make a big difference in reducing the amount of food yield loss caused by infectious diseases, especially in places with few resources. This collection of images can be used in computer vision and mobile technology to make farming more productive around the world. It shows how important it is for everyone to work together to fight crop pests and make sure there is enough food for the world's growing population.

FIGURE 1: APPLE LEAF

FIGURE 2: GRAPE LEAF

This project uses PyTorch framework using Python programming language. The model is trained on Nvidia GeForce GTX 1050, 4GB Graphics card. Apple image dataset with 3171 images and Grape image dataset with 4062 images are trained using the GPU.

3. BACKGROUND

Predicting plant diseases, which is mostly an image classification job, has come a long way to different deep learning techniques. Besides VGG16 and ResNet18, several other important methods have shown a lot of promise in image classification tasks for predicting plant diseases.

When deep learning came along, it opened the door to new ways of doing image classification jobs. Deep learning architectures like Xception, which have depth-separable convolutions, have shown amazing skill in figuring out complex patterns in images. The Inception V3 design, which uses inception modules to pull out features at different sizes, has also become a powerful tool for classifying images. The Mobilenet framework, which is designed to work best with mobile and edge computing devices, also offers a quick and accurate way to classify images.

Each of these different deep learning architectures has its own benefits when it comes to accuracy, model complexity, and computational speed. While VGG16 and ResNet18 have proven to be reliable models for transfer learning in predicting plant diseases, looking into other options such as Xception, Inception V3, and Mobilenet increases the number of possible answers. The different design details and optimization methods of each architecture give researchers and practitioners a wide range of options when they want to make disease prediction models for apple and grape leaves that are accurate and scalable.

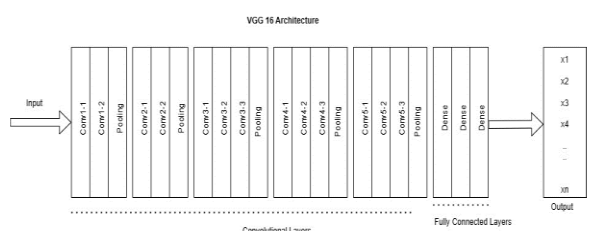
It is important to investigate a range of deep learning methods besides VGG16 and ResNet18 to find the best ones and make the architectures fit the specifics of tasks like finding plant diseases. By looking into these different approaches, this project aims to see how well transfer learning can work and if they can be used to predict plant diseases. This will help make models for managing crop diseases more accurately and efficiently in the long run.

4. IMPLEMENTATION METHODS

Deep learning models, like VGG16 and ResNet18, perform well for predicting diseases in plant leaves, especially in apples and grapes. These architectures, which are known for how well they work at classifying images, are very helpful for finding subtle disease-related patterns in leaf images.

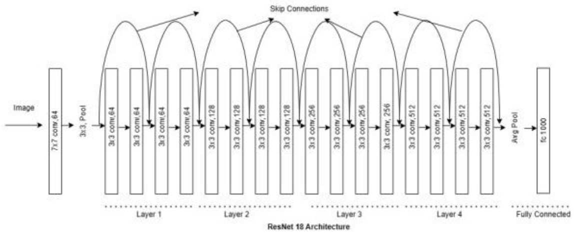
Data preprocessing is a crucial phase of training a deep learning model. This project introduces several data augmentation techniques like cropping the regions of images, flipping the images at diverse viewpoints, introducing variations in brightness, contrast, saturation and hue. These image transformations improve the generalization capabilities of the model and make it robust.

VGG16 Architecture:



VGG16 has 16 layers, with 13 convolutional layers, 3 pooling layers, and 3 fully connected layers on top. It is known for its depth and uniform design. Because this design is simple and consistent, it can be used for transfer learning very easily. VGG16 can be fine-tuned to quickly pull-out hierarchical features from images of apple and grape leaves by using weights that have already been learned on large image datasets. By doing this, the model learns to tell the difference between healthy and sick leaves using what it has learned, which makes it possible to accurately predict diseases.

ResNet18 Architecture:



ResNet18 is designed based on residual learning that fixes the problem of vanishing gradients in deeper networks by adding skip links. With 18 layers, this architecture's residual blocks make it easier for information to move, which makes it possible to train deeper networks more accurately. The design of ResNet18 improves feature spread and makes it easier to get detailed disease-related features from pictures of apple and grape leaves. ResNet18 quickly learns how to recognize different types of plant diseases by using transfer learning methods. This means that the model starts out with weights that have already been trained on different datasets.

Transfer learning is used in both the VGG16 and ResNet18 designs, and it is a key part of the implementation process. Transfer learning uses models that have already been trained on very large datasets (like ImageNet) and tweaks them on data that is specific to the topic (apple and grape leaf images). This process lets the models take on learned traits and representations from different pictures, which then lets them be tailored to the specifics of plant diseases.

Using the VGG16 and ResNet18 architectures to predict diseases in apple and grape leaves involves preprocessing the leaf images, feeding them into the right architectures, fine-tuning the models on the given dataset, and using the learned representations to decide whether leaves are healthy or sick. Because these designs are strong and good at picking up disease-related information, they are very useful for making accurate and scalable systems for predicting diseases in agriculture.

5. RESULTS

Training and validation accuracies of ResNet18 and VGG16 models over Apple and Grape leaves images are calculated. Precision, Recall and F1 score are calculated for different image class categories.

Precision: Precision measures the accuracy of the positive predictions made by the model. It is the ratio of true positive predictions to both true positives and false positive predictions.

Recall: Recall, which is also known as sensitivity or true positive rate, can be defined as the model's ability to identify all the positive instances. It calculates the ratio of true positive predictions to true positives and false negative predictions.

F1 Score: It is the harmonic mean of precision and recall. It is determined to find the balance between precision and recall, giving a single score that considers both metrics. It's especially useful when dealing with imbalanced classes.

ResNet Model: Apple Image Dataset

Training Accuracy: 98.14%

Validation Accuracy: 98.08%

TABLE 1

Disease Class	Precision	Recall	F1-Score
Apple_scab	0.97	0.99	0.98
Black_rot	1	0.99	0.99
apple_rust	0.98	0.98	0.98
healthy	0.99	0.99	0.99

ResNet Model: Grape Image Dataset

Training Accuracy: 98.09%

Validation Accuracy: 95.95%

TABLE 2

Disease Class	Precision	Recall	F1-Score
Black_rot	0.88	0.98	0.93
Esca	0.99	0.89	0.94
Leaf_blight	0.99	1	1
healthy	1	1	1

VGG Model: Apple Image Dataset

Training Accuracy: 97.24%

Validation Accuracy: 96.79%

TABLE 3

Disease Class	Precision	Recall	F1-Score
Apple_scab	0.91	0.99	0.95
Black_rot	0.99	0.91	0.95
apple_rust	0.91	1	0.95
healthy	0.99	0.97	0.98

VGG Model: Grape Image Dataset

Training Accuracy: 95.84%

Validation Accuracy: 95.66%

TABLE 4

Disease Class	Precision	Recall	F1-Score
Black_rot	0.96	0.89	0.92
Esca	0.90	0.99	0.94
Leaf_blight	1	0.94	0.97
healthy	0.96	1	0.98

Plots:

FIGURE 3

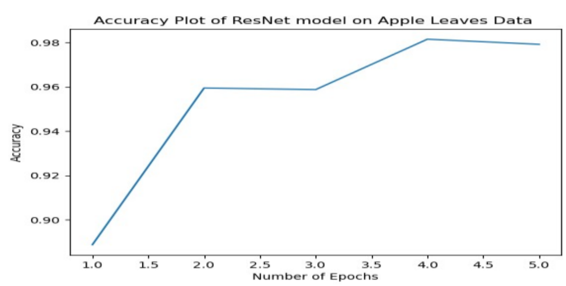


FIGURE 4

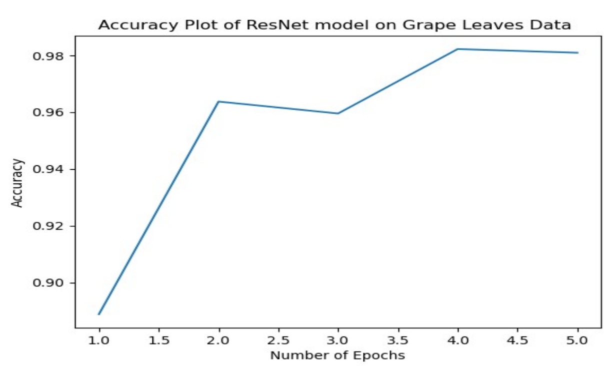


FIGURE 5

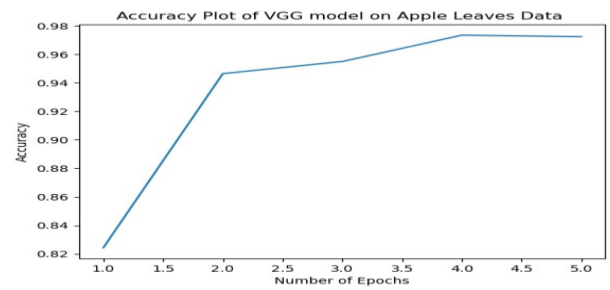
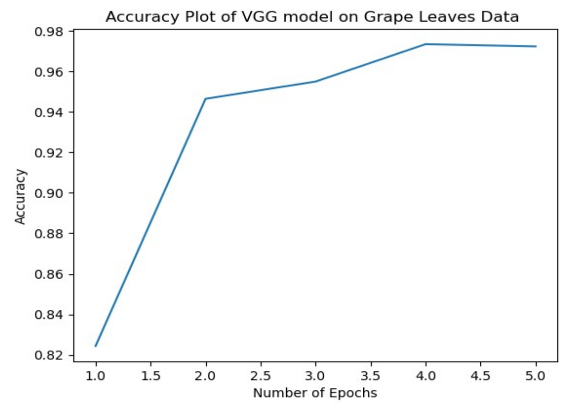


FIGURE 6



Both the apple and grape datasets produced high training and validation accuracies which shows that the ResNet model has learned from the training data and does well on validation data that it has not seen before. This means that the model seems to have been able to catch all the details and complexities in the images. Similarly, both the training and test datasets for the VGG model show high accuracy, though not quite as high as those for ResNet. It is important to keep in mind, though, that the variance, or changes in how well the model works with new data, could be shown by how well it handles the training and validation sets.

Low Bias: Due to their high accuracy on both the training and validation sets, both the ResNet and VGG models showed low bias. Low bias means that the models were able to find the real patterns and connections in the datasets, which led to correct forecasts.

Low Variance: Both ResNet and VGG models have low variance because the differences between their training and validation accuracy are low. Low variance means that the models work well with data they haven't seen before, which keeps them from being too perfect. ResNet, on the other hand, does a little better at generalization than VGG because the difference between its training and evaluation accuracy is quite larger.

6. CONCLUSION AND INSIGHTS

The utilization of transfer learning with VGG16 and ResNet18 models has proven to be a reliable approach for plant disease prediction in apple and grape leaves. These pre-trained models, with their learned features from extensive datasets, showcase the potential to generalize and identify patterns in new, unseen data. The application of data augmentation techniques such as cropping the regions of images, flipping the images at diverse viewpoints, introducing variations in brightness, contrast, saturation and hue has significantly enhanced the model's adaptability. These methods have enabled the model to perform well against variations in lighting and other environmental conditions, making it more robust and reliable in real-world scenarios where ideal conditions may not exist. Training the model exclusively on data containing grape and apple leaf images could affect its performance when introduced to other plant species. While the model exhibits proficiency in identifying diseases in grape and apple leaves, its effectiveness might vary when trained with diverse leaf structures and disease patterns present in other plant species. There exists an opportunity for further advancement in this project. The future scope involves increasing the size of the dataset to deal with a broader variety of plant leaves, thereby enhancing the model's adaptability and generalization capabilities. Additionally, using more sophisticated preprocessing methods and augmenting the dataset with real-world environmental variations can contribute to improving the model's robustness and accuracy.

Finally, combining transfer learning techniques with the right preprocessing methods shows potential for making disease prediction in plant leaves much more accurate. As it continues to be improved and used in more places, this method could become an important part in agricultural practices for early disease detection and mitigation.

7. REFERENCES

1. V. Menon, V. Ashwin and R. K. Deepa, "Plant Disease Detection using CNN and Transfer Learning," 2021 International Conference on Communication, Control and Information Sciences (ICCISc), Idukki, India, 2021, pp. 1-6, doi: 10.1109/ICCISc52257.2021.9484957.

2. A. Lakshmanarao, M. R. Babu and T. S. R. Kiran, "Plant Disease Prediction and classification using Deep Learning ConvNets," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1-6, doi: 10.1109/AIMV53313.2021.9670918.

3. A. Förster, J. Behley, J. Behmann and R. Roscher, "Hyperspectral Plant Disease Forecasting Using Generative Adversarial Networks," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 1793-1796, doi: 10.1109/IGARSS.2019.8898749.

4. X. Liu, W. Min, S. Mei, L. Wang and S. Jiang, "Plant Disease Recognition: A Large-Scale Benchmark Dataset and a Visual Region and Loss Reweighting Approach," in IEEE Transactions on Image Processing, vol. 30, pp. 2003-2015, 2021, doi: 10.1109/TIP.2021.3049334.

5. A. Hidayatuloh, M. Nursalman and E. Nugraha, "Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model," 2018 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia, 2018, pp. 199-204, doi: 10.1109/ICITSI.2018.8696087.

6. M. A. Moid and M. Ajay Chaurasia, "Transfer Learning-based Plant Disease Detection and Diagnosis System using Xception," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1-5, doi: 10.1109/I-SMAC52330.2021.9640694.

7. N. Gobalakrishnan, K. Pradeep, C. J. Raman, L. J. Ali and M. P. Gopinath, "A Systematic Review on Image Processing and Machine Learning Techniques for Detecting Plant Diseases," 2020 International Conference on

Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 0465-0468, doi: 10.1109/ICCSP48568.2020.9182046.

8. P. K. V, E. G. Rao, G. Anitha and G. K. Kumar, "Plant Disease Detection using Convolutional Neural Networks," 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2021, pp. 1473-1476, doi: 10.1109/ICOEI51242.2021.9453045.

9. X. Guan, "A Novel Method of Plant Leaf Disease Detection Based on Deep Learning and Convolutional Neural Network," 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2021, pp. 816-819, doi: 10.1109/ICSP51882.2021.9408806.

10. T. Sanida, D. Tsiktsiris, A. Sideris and M. Dasygenis, "A Heterogeneous Lightweight Network for Plant Disease Classification," 2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 2021, pp. 1-4, doi: 10.1109/MOCAST52088.2021.9493415.

11. Deepa, R. N and C. Shetty, "A Machine Learning Technique for Identification of Plant Diseases in Leaves," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 481-484, doi: 10.1109/ICICT50816.2021.9358797.