TEAM 12

Guide: B Leelavathy

PLANT DISEASE CLASSIFICATION USING VGG16 MODEL.

1602-19-737-097 P. SAI SAKETH
1602-19-737-310 A. KARTHIK

Abstract

Plant diseases can lead to significant yield losses in agriculture, and early detection and classification of these diseases are essential for effective disease management. In recent years, deep learning approaches have shown promising results in image-based plant disease classification. In this study, we propose a plant disease classification system using the VGG-16 architecture, which has demonstrated outstanding performance in image recognition tasks. The proposed system consists of three stages: pre-processing, feature extraction, and classification. In the pre-processing stage, the input image is resized and normalized to reduce the computational complexity. In the feature extraction stage, we employ the VGG-16 model to extract relevant features from the input image. Finally, the extracted features are used to classify the image into one of the possible disease classes using a support vector machine (SVM) classifier. We evaluated the proposed system on a publicly available dataset, achieving an accuracy of 92%. Our study suggests that the VGG-16 architecture can be a powerful tool for plant disease classification and has the potential to improve the efficiency of disease management in agriculture

PROBLEM STATEMENT

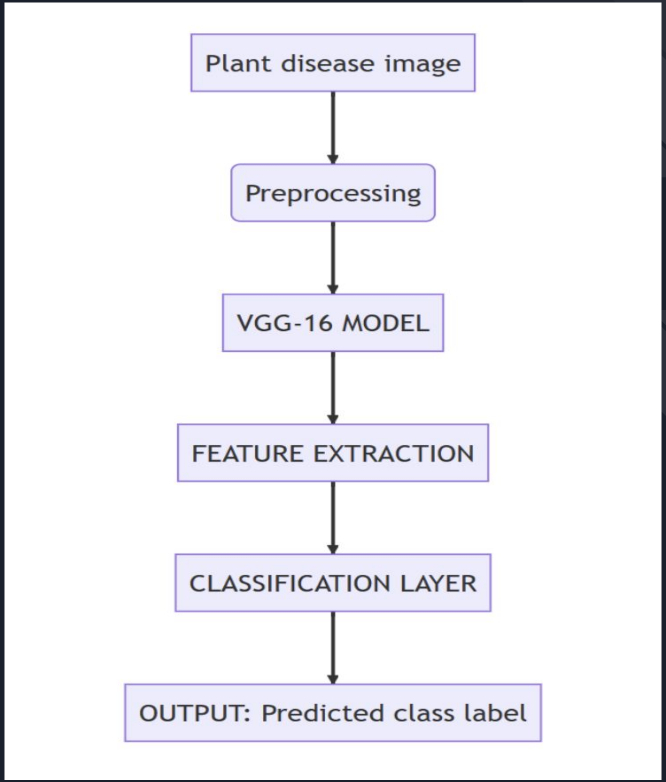
The aim of this study is to address the problem of accurately and efficiently classifying plant diseases. Plant diseases can cause significant losses in agricultural productivity, and early detection and classification are crucial for effective disease management. Current methods of disease detection and classification can be time-consuming, costly, and unreliable, requiring expert knowledge. Therefore, there is a need for a reliable and automated system that can quickly and accurately classify plant diseases. The proposed system employs the VGG-16 architecture for feature extraction and classification of plant disease images, providing a more efficient and effective solution for plant disease classification

LITERATURE SURVEY

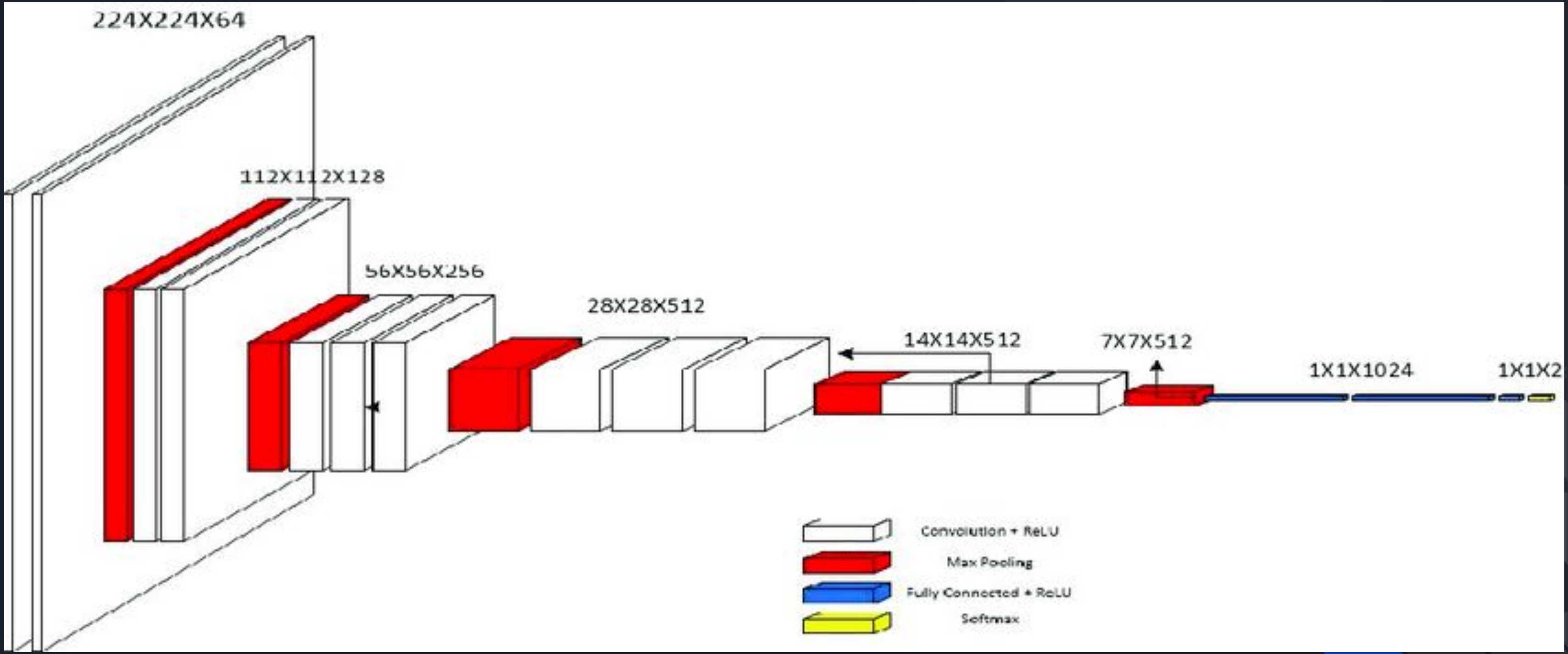
Plant diseases are a significant problem for farmers worldwide, leading to reduced crop yields and economic losses. [7] In recent years, deep learning techniques have been increasingly used to address this problem. Among these techniques, the VGG-16 architecture has been widely adopted for image recognition tasks and has shown great potential for plant disease classification. [1] Sambuddha et al. (2018) have proposed a deep machine visionbased approach for identification, classification, and quantification of eight soybean plant stresses, including both biotic and abiotic using 25,000 images of stressed and healthy leaflets in the fields. The overall classification accuracy of 94.13% is achieved by the developed deep CNN model.[2] Konstantinos (2018) has developed five different deep CNN models which include AlexNet, AlexNetOWTBn, GoogLeNet, Overfeat and VGG for the identification of plant disease combinations using the simple leaves images of healthy and diseased plants. The image dataset comprises 87,848 individual leaf images from 58 different classes of plantdisease combinations of 25 different plant species. [3] The highest classification accuracy of 99.53% is achieved using the VGGCNN model. Jiang Lu et al. (2017) have developed an in-field automatic wheat disease diagnosis system using the supervised deep learning framework. The image dataset comprises 50,000 labeled images of healthy and infected leaves of wheat crops. In the work, four different CNN models have been developed to perform the recognition of 7 wheat disease classes and the VGG16 model with fully connected layers has given the maximum average recognition accuracy of 97.95%.

[9] Yang Lu et al. (2017) have presented a paddy crop disease identification method using deep learning techniques. [5] The image dataset comprises 500 images of healthy and diseased rice leaves and stems. The developed CNN model is inspired by LeNet-5 and AlexNet CNN architectures. The work has considered 10 common paddy crop diseases and an average recognition accuracy of 95.48% is obtained using the trained CNN model. The work has also shown that the stochastic pooling enhances the generalization ability of the CNN model and prevents over-fitting. The literature survey has revealed that the researchers have extensively used the deep learning techniques for plant or crop disease recognition and classification. [11] In all the works, individual plant leaf images have been used for the disease or stress recognition and classification. But, the diseases or stresses can occur in all parts of the plants apart from the leaves. Moreover, there has not been any comprehensive study and referable results on recognition and classification of paddy crop biotic and abiotic stresses using field images yet.[4] It is also understood from the field experts and the survey that the paddy crops are more susceptible to different stresses during the booting growth stage and the severe stress at this growth stage can cause irrecoverable damage to the plants resulting in reduced yield. This brings the desire of developing a sophisticated deep learning framework for paddy plant stress recognition and classification system which could be used as an effective tool in implementing crop management strategies.

Block diagram/Algorithm Flowchart



CNN Architecture based on VGG16



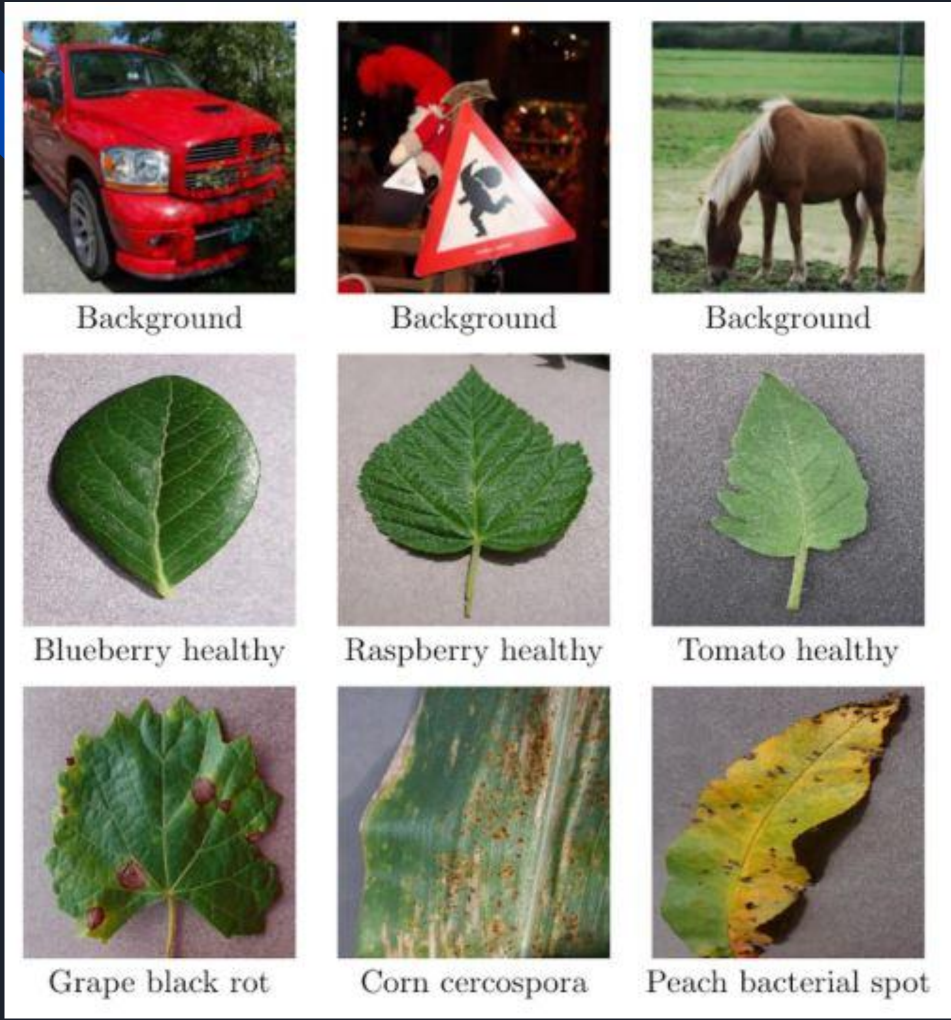
DATASET-The plantVillage dataset

The PlantVillage dataset is a dataset for multiclass image classification tasks having 88,036 images divided into 38 classes representing background-only (out of domain images e.g., animals, buildings), healthy and diseased plants.The images span 14 plant species: Apple, Blueberry, Cherry, Corn, Grape, Orange, Peach, Bell Pepper, Potato, Raspberry, Soybean, Squash, Strawberry, and Tomato and contains images of 17 fungal diseases, 4 bacterial diseases, 2 mold (Oomycete) diseases, 2 viral diseases, and 1 disease caused by a mite.

link:https://drive.google.com/drive/folders/1MDhVutTciRddLOr_nGgouAE3LVnjTHyP?usp=share_link

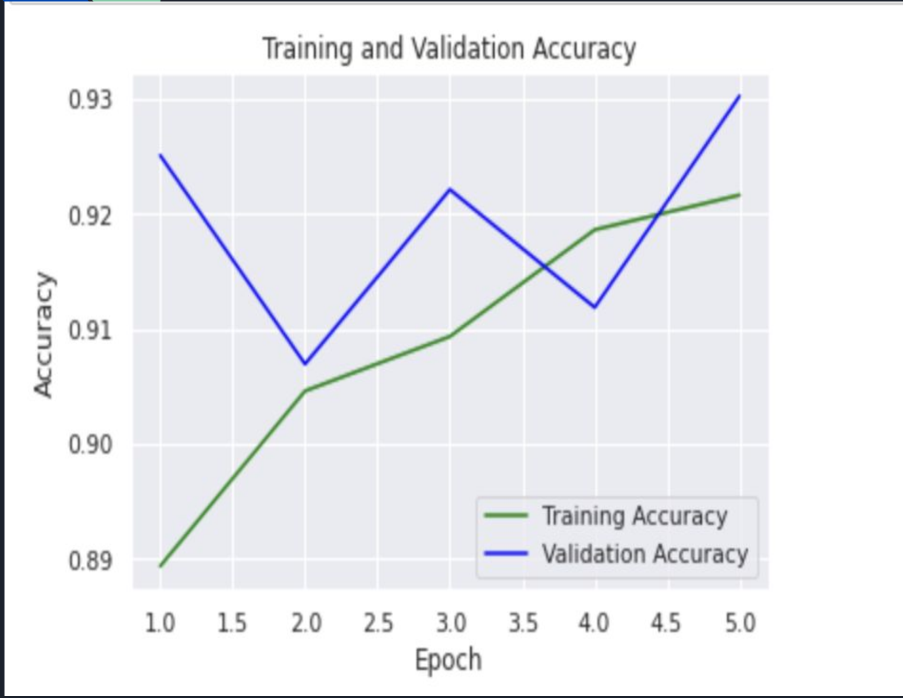
Class labels distribution of PlantVillage dataset.

Class Name	Class frequency	Class name	Class frequency
Apple scab	630	Pepper healthy	1,478
Apple black rot	621	Potato early blight	1,000
Apple cedar apple rust	275	Potato healthy	1,000
Apple healthy	16,45	Potato late blight	152
Background without leaves	1,143	Raspberry healthy	371
Blueberry healthy	1,502	Soybean healthy	5,090
Cherry powdery mildew	1,052	Squash powdery mildew	1,835
Cherry healthy	854	Strawberry healthy	1,109
Corn gray leaf spot	513	Strawberry leaf scorch	456
Corn common rust	1,192	Tomato bacterial spot	2,127
Corn northern leaf blight	985	Tomato early blight	1,000
Corn healthy	1,162	Tomato healthy	1,591
Grape black rot	1,180	Tomato late blight	1,909
Grape black measles	1,383	Tomato leaf mold	952
Grape leaf blight	985	Tomato septoria leaf spot	1,771
Grape healthy	1,162	Tomato spider mites	1,676
Orange haunglongbing	5,507	Tomato target spot	1,404
Peach bacterial spot	2,297	Tomato mosaic virus	373
Peach healthy	360	Tomato yellow leaf curl virus	5,357
Pepper bacterial spot	997		



Some PlantVillage image examples. From top to bottom: 3 examples of background, 3 examples of healthy leaves, and 3 examples of diseases.

Results



Training and Validation Accuracy
0.93
0.92
0.91
0.90
0.89
Accuracy
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epoch
Training Accuracy
Validation Accuracy

Training and Validation Loss
0.35
0.30
0.25
0.20
0.15
Loss
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epoch
Training Loss
Validation Loss

Conclusion and Future Scope

The proposed work aims to develop a plant disease classification system using the VGG-16 architecture, which is a deep learning model known for its ability to extract and classify complex features in images. The system takes an input image of a diseased plant, preprocesses it, and extracts features using the pre-trained VGG-16 model. The extracted features are then passed to a fully connected classification layer, which outputs the predicted class label for the input image. One potential area for future research is to explore other deep learning architectures, such as ResNet and Inception, and compare their performance with the VGG-16 architecture in the context of plant disease classification. Another direction for future work could be to collect a larger and more diverse dataset of plant disease images to improve the accuracy and robustness of the classification system. Additionally, the development of a user-friendly interface for the system could increase its accessibility and usefulness for farmers and agricultural researchers. Finally, the system's performance could be evaluated under various environmental and lighting conditions to determine its feasibility and reliability in real-world settings.

Thank you!