Leaf Disease Detection

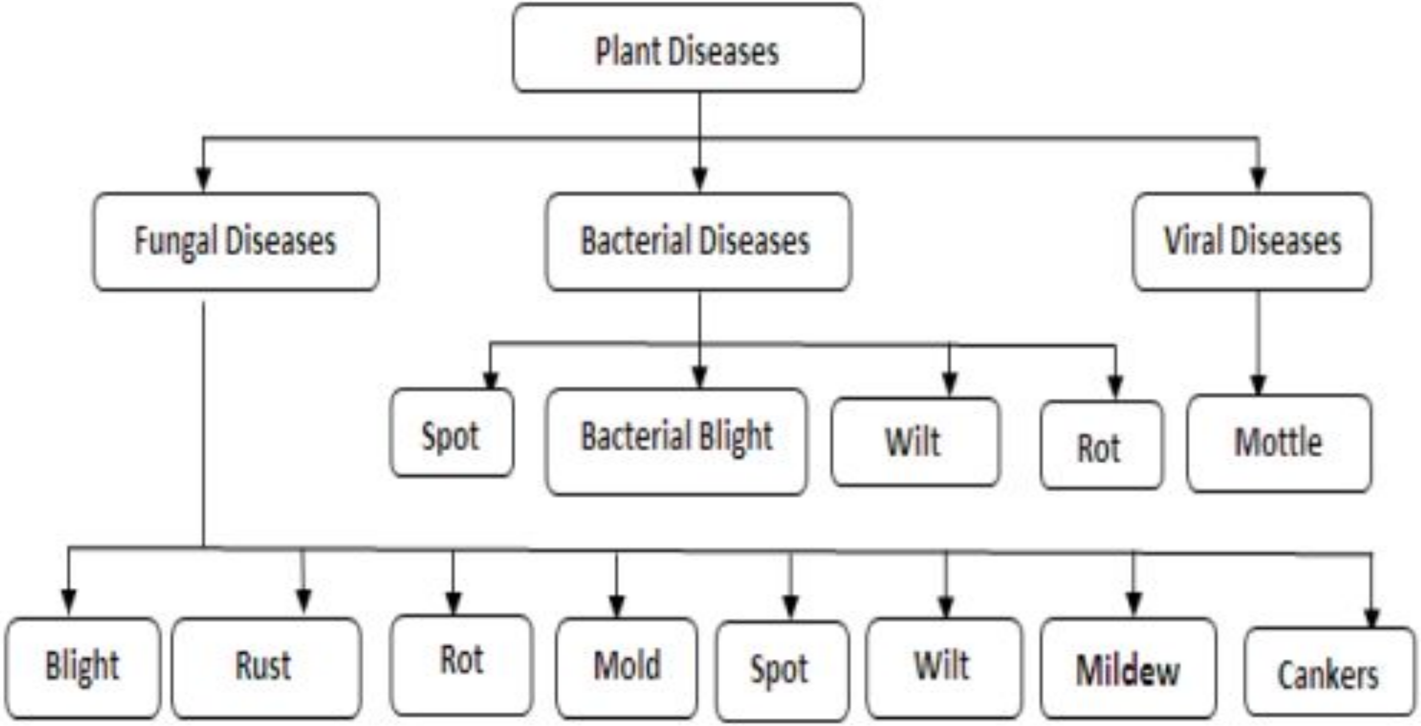
Team Members: Riza Khan,Ketan Dhakate,Krishna Parthani, Abhishek Agnihotri
Guide:Prof.P.S. Agrawal
Session 2022-23

Department of Electronics & Communication Engg.

Introduction

Plant diseases are a major threat to agricultural productivity and food security, causing significant losses in crop yields worldwide. Timely and accurate detection of plant diseases is crucial to prevent their spread and minimize the economic impact on farmers. Traditional methods of disease detection are often time-consuming and rely on the visual inspection of plant samples by experts.With the recent advancements in machine learning and image processing techniques, there is an opportunity to develop a more efficient and automated approach for the detection of plant diseases.

Classification of Diseases



Problem Statement/objectives

- The problem statement is to develop a machine learning-based system that can accurately identify various leaf diseases in crops using image processing techniques. This system should be able to analyze images of plant leaves and provide accurate diagnosis of diseases such as rust, blight, and powdery mildew, among others.
- The objective is to create a system that can assist farmers in early disease detection so that the measures can be taken as early as possible. The system should be user-friendly and accessible to farmers with limited technical expertise, helping them to reduce crop losses and improve their overall yields.

Existing Solutions/Literature Survey

At present the existing solutions of leaf disease detection involve popular approaches like:

1. Image Processing and Feature Extraction
2. Deep Learning-based Approaches
3. Transfer Learning
4. Ensemble Methods

Some examples of existing solutions for leaf disease detection include PlantVillage, Leaf Doctor, CROPSCAN,etc. These are just a few examples of solutions to our problem statement using machine learning and computer vision.

Proposed solution

- The model is developed based on the IP and ML approaches for detection of leaf disease.
- User can take an image their plants' leaf and load it into the APP we are providing(currently local host).
- In this app the image will be reshaped and ALL the required features be extracted. Now and using the CNN algorithm & plant disease dataset ->image scanned -> result immediately(predicts class) .
- Throughout the process user does not need to compromise with plant as there's no chemical or physical damage involved.
- Identifies the most probable disease it belongs to.

Proposed solution

- The model is developed based on the IP and ML approaches for detection of leaf disease.
- User can take an image their plants' leaf and load it into the APP we are providing(currently local host).
- In this app the image will be reshaped and ALL the required features be extracted. Now and using the CNN algorithm & plant disease dataset ->image scanned -> result immediately(predicts class) .
- Throughout the process user does not need to compromise with plant as there's no chemical or physical damage involved.
- Identifies the most probable disease it belongs to.

Block Diagram

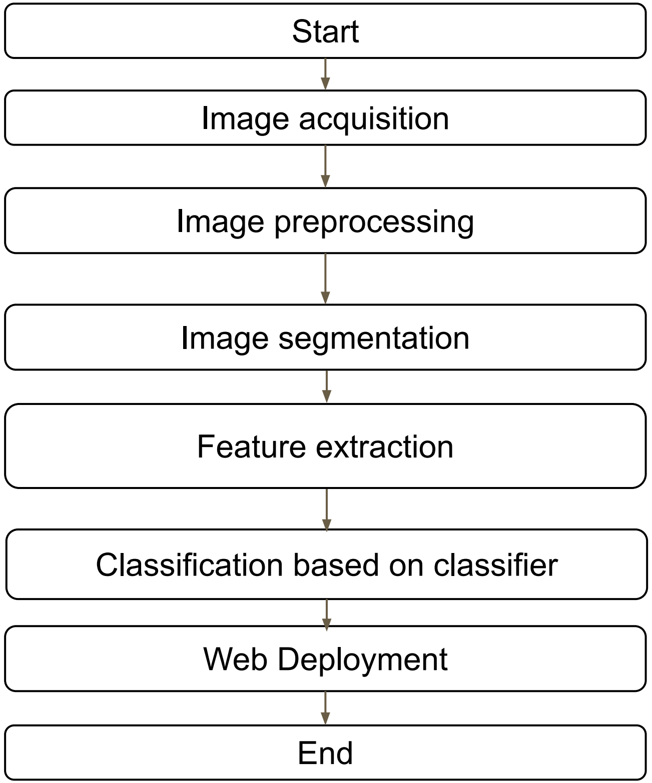


Figure 1: Basic Flowchart of Leaf Disease Detection & Classification

Description of Block Diagram

Fig 1: Basic Flowchart Of Disease Detection And Classification

1. Image Acquisition:Images of the infected leaves are obtained. This database has different types of plant diseases, and the images are stored in JPEG format. These images are then read in MATLAB using the read command.
2. Image Pre-processing: Image pre-processing is used to erase noise from the image or other object exclusion, different pre-processing techniques. Image scaling is used to convert the original image into thumbnails because the pixel size of the original image is large and it requires more time for the overall procedure hence after converting the image into thumbnails the pixel size will get decreases and it will require less time.
3. Image segmentation:Image segmentation is one of the most widely used methods to distinguish pixels of image well in a targeted app. It distributes an image into numerous discrete states such that the pixels have great similarity in each area and high dissimilarity between areas.
4. Feature Extraction:Feature Extraction is an important part of disease detection. It plays an important role in the identification of an object. Feature extraction is utilized in several applications in image processing. Colour, texture edges, morphology are the features, which are utilized in disease detection.
5. Detection and classification of plant diseases:The final stages are the detection of the diseases and with the help of disease classify the plants with the disease matches with the given dataset.
6. Web Deployment: The detection module is deployed on a web-based platform,namely,streamlit making it user friendly.

Working of SVM

Support Vector Machines (SVM) is a supervised machine learning algorithm that can be used for classification and regression tasks. In the context of leaf disease detection, SVM can be employed to classify whether a leaf is healthy or diseased based on various features extracted from the leaf images.

Here's a high-level overview of how SVM works:

- **Data preparation:** Collect a dataset of leaf images, where each image is labeled as healthy or diseased. Extract relevant features from these images, such as color histograms, texture descriptors, or shape characteristics. Also, preprocess the data by normalizing or scaling the features if necessary.
- **Feature representation:** Represent each leaf image using the extracted features. This step aims to transform the raw image data into a format that SVM can understand and process.
- **Training phase:** In this step, the SVM model learns the boundary that separates healthy and diseased leaves in the feature space. SVM finds the best hyperplane that maximally separates the two classes while maximizing the margin (distance) between the hyperplane and the support vectors (data points closest to the hyperplane).
- **Kernel trick:** SVM can employ the kernel trick to transform the input feature space into a higher-dimensional space, allowing for nonlinear separation of data. This helps SVM capture complex relationships between features and improve classification performance.
- **Classification phase:** Once the SVM model is trained, it can be used to classify new, unseen leaf images. The SVM calculates the distance (or similarity) of the new leaf image to the hyperplane learned during training. Based on this distance and the location of the new sample in the feature space, the SVM assigns a label, i.e., healthy or diseased, to the leaf.

SVM is known for its ability to handle high-dimensional feature spaces and its robustness against overfitting. It can effectively classify leaf diseases by finding a decision boundary that separates healthy and diseased leaves based on the extracted features. However, the performance of SVM heavily depends on the quality and relevance of the chosen features, as well as the choice of appropriate parameters and the training dataset.

Working of CNN

Convolutional Neural Networks (CNN) have proven to be very effective in leaf disease detection. The working of CNN in leaf disease detection involves several steps:

- **Data Preprocessing**: The first step is to preprocess the input images. This may involve resizing the images, normalizing the pixel values, and applying other image processing techniques.
- **Convolutional Layers**: The next step is to apply convolutional layers to the preprocessed images. The convolutional layers are responsible for extracting features from the images. These features are learned by the network during the training process.
- **Pooling Layers**: After the convolutional layers, pooling layers are applied. Pooling layers reduce the spatial dimensions of the feature maps by downsampling the images. This helps to reduce the computational cost and also prevents overfitting.
- **Fully Connected Layers**: Once the feature maps have been downsampled, they are flattened and passed through a fully connected layer. This layer performs classification on the extracted features to determine the type of leaf disease present.
- **Output Layer**: Finally, the output layer of the CNN produces the probability distribution over the different classes of leaf diseases.

During training, the weights of the CNN are updated using backpropagation to minimize the loss function. This allows the CNN to learn to identify the patterns and features that are important for distinguishing between different types of leaf diseases.Basically, CNNs use a combination of convolutional layers, pooling layers, and fully connected layers to extract features from preprocessed images, which are then used to classify the type of leaf disease present.

Results (Intermediate results/simulation results)

Accuracy of test data using SVM

```
[ ]  from sklearn import metrics
     metrics.accuracy_score(y_test, y_pred)

     0.7338888888888889
```

Results (Intermediate results/simulation results)

Accuracy of train data using CNN

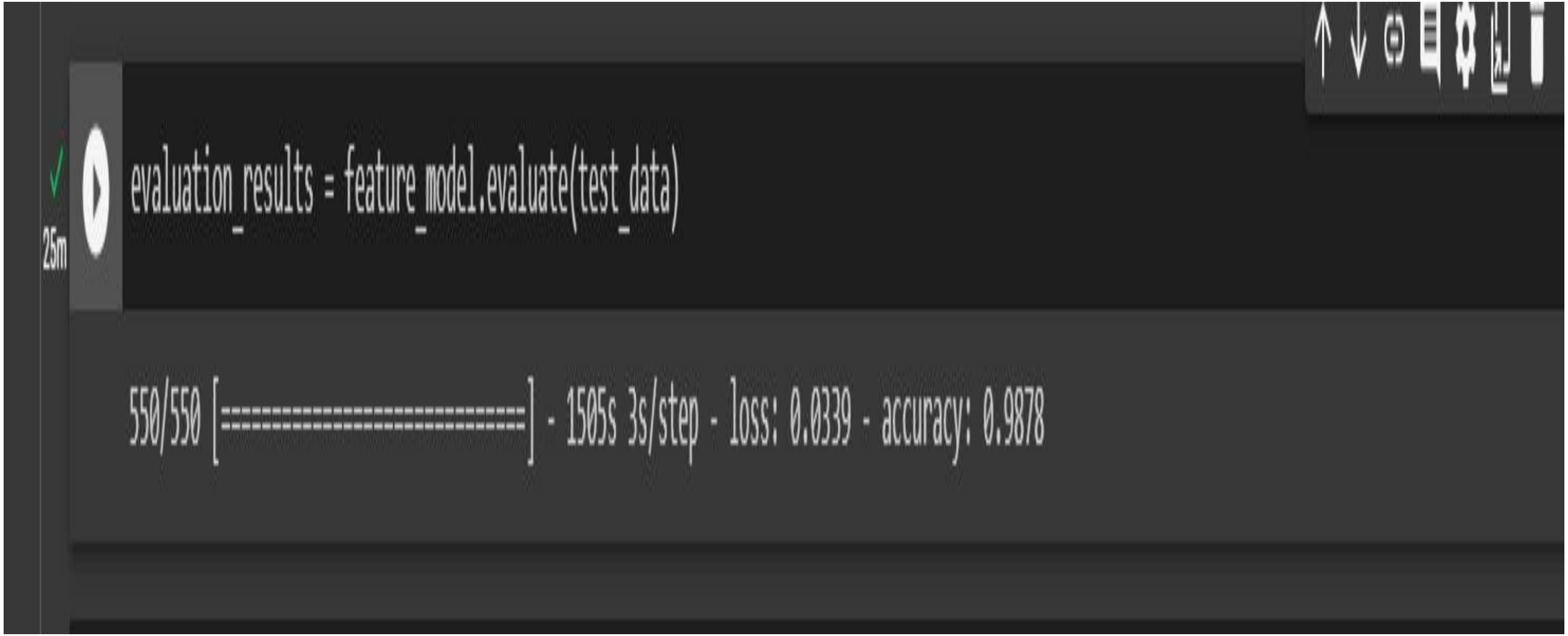
```
initial_epochs = 1

history1 = feature_model.fit(
    train_data,
    epochs=initial_epochs,
    validation_data=test_data,
    callbacks=[
        early_stopping,
        model_checkpoint,
        reduce_lr,
        create_tensorboard_callback('plant_disease_model', 'EfficientNetB010')
    ]
)
```

```
Saving TensorBoard log files to: plant_disease_model/EfficientNetB010/20230507-133547
2197/2197 [==============================] - 8450s 4s/step - loss: 0.1366 - accuracy: 0.9589 - val_loss: 0.0339 - val_accuracy: 0.9878 - lr: 0.0010
```

Results (Intermediate results/simulation results)

Accuracy of test data using CNN



actual:Grape___Leaf_blight_(Isariopsis_Leaf_Spot),
pred:Grape___Leaf_blight_(Isariopsis_Leaf_Spot),
prob:100.00%

actual:Tomato___Target_Spot,
pred:Tomato___Target_Spot,
prob:99.96%

actual:Tomato___Early_blight,
pred:Tomato___Early_blight,
prob:98.87%

actual:Blueberry___healthy,
pred:Blueberry___healthy,
prob:100.00%

actual:Tomato___Tomato_mosaic_virus,
pred:Tomato___Tomato_mosaic_virus,
prob:100.00%

actual:Tomato___Target_Spot,
pred:Tomato___Target_Spot,
prob:83.41%

actual:Grape___Leaf_blight_(Isariopsis_Leaf_Spot),
pred:Grape___Leaf_blight_(Isariopsis_Leaf_Spot),
prob:100.00%

actual:Cherry_(including_sour)___healthy,
pred:Cherry_(including_sour)___healthy,
prob:100.00%

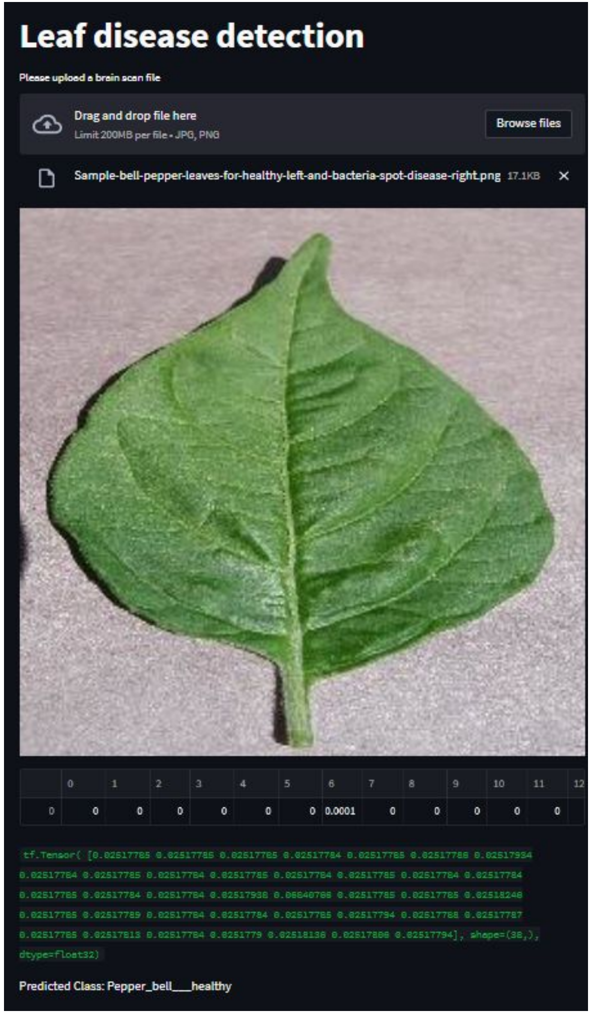
actual:Corn_(maize)___Northern_Leaf_Blight,
pred:Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot,
prob:69.49%

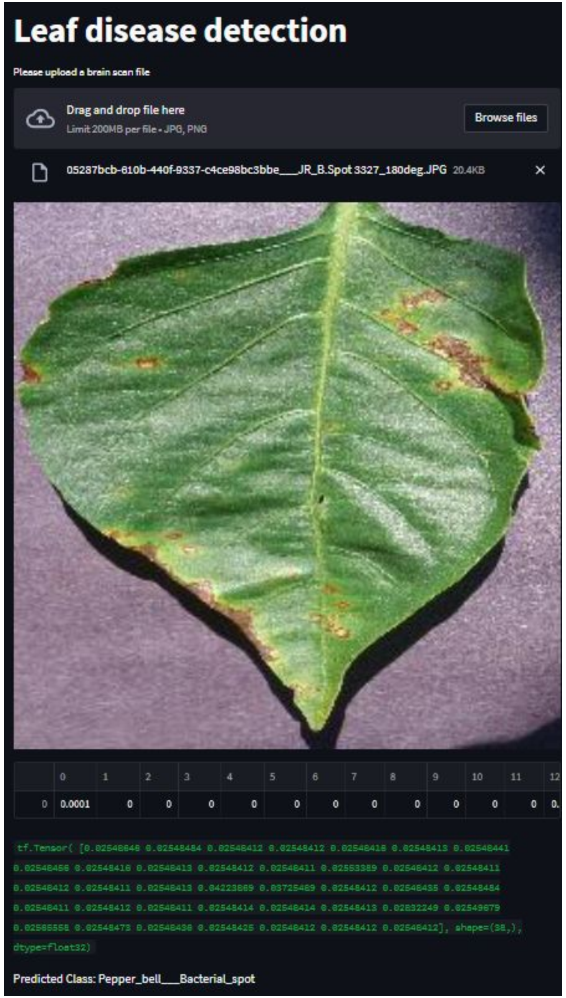
00a3fc0e-64cc-4e35-ac2f-aef04fda9b22___Mary_HL 9177_newPixel25.JPG 15.2KB

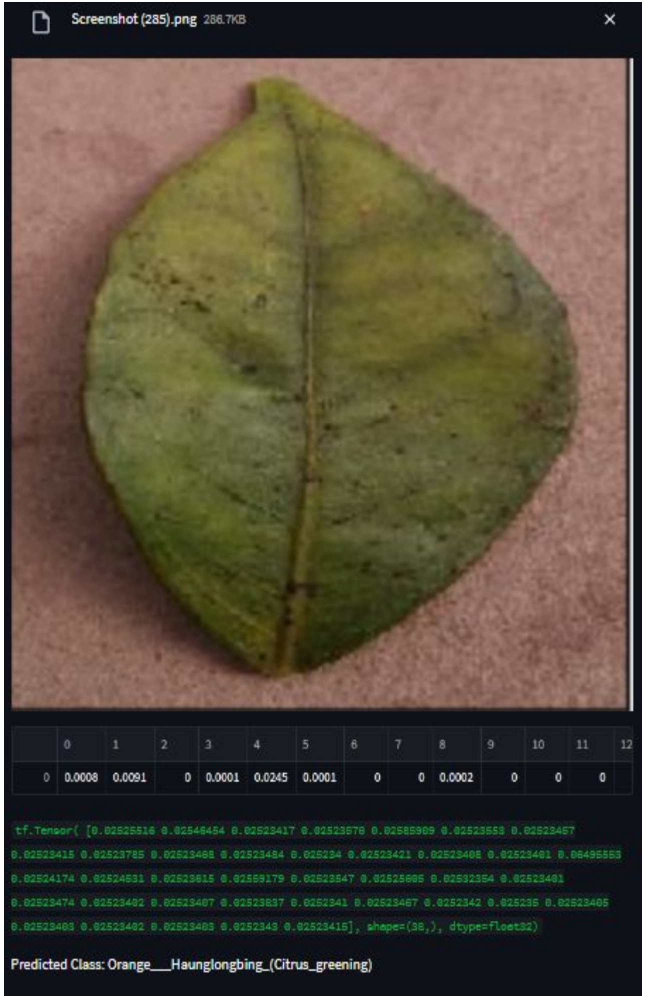
	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0.0003	0.0001	0	0	0	0	0	0	0	0	0.0

```
tf.Tensor( [0.02517934 0.02517925 0.02518617 0.02518189 0.02517942 0.02517929 0.02517923
0.02517923 0.02517927 0.02517923 0.02517923 0.02517924 0.02518874 0.02517924 0.02517923
0.02517923 0.02517923 0.02517924 0.02517924 0.02517924 0.02517924 0.02517924 0.0252083 0.06832416
0.02517924 0.02517923 0.02517945 0.02517988 0.02517924 0.02517925 0.02517953 0.02517924
0.02517923 0.02517937 0.0251887 0.02517924 0.02517944 0.02518809], shape=(38,), dtype=float32)
```

Predicted Class: Raspberry___healthy


Leaf disease detection
Please upload a brain scan file
Drag and drop file here
Limit 200MB per file • JPG, PNG
Browse files
Sample-bell-pepper-leaves-for-healthy-left-and-bacteria-spot-disease-right.png 17.1KB
0 1 2 3 4 5 6 7 8 9 10 11 12
0 0 0 0 0 0 0 0.0001 0 0 0 0 0
tf.Tensor([0.02517785 0.02517785 0.02517785 0.02517784 0.02517785 0.02517788 0.02517934
0.02517784 0.02517785 0.02517784 0.02517785 0.02517784 0.02517785 0.02517784 0.02517784
0.02517785 0.02517784 0.02517784 0.02517938 0.06840788 0.02517785 0.02517785 0.02518248
0.02517785 0.02517789 0.02517784 0.02517784 0.02517785 0.02517794 0.02517788 0.02517787
0.02517785 0.02517813 0.02517784 0.0251779 0.02518138 0.02517888 0.02517794], shape=(38,),
dtype=float32)
Predicted Class: Pepper_bell___healthy


Leaf disease detection
Please upload a brain scan file
Drag and drop file here
Limit 200MB per file • JPG, PNG
Browse files
05287bcb-610b-440f-9337-c4ce98bc3bbe___JR_B.Spot 3327_180deg.JPG 20.4KB
0 1 2 3 4 5 6 7 8 9 10 11 12
0 0.0001 0 0 0 0 0 0 0 0 0 0 0 0.
tf.Tensor([0.02548846 0.02548484 0.02548412 0.02548412 0.02548418 0.02548413 0.02548441
0.02548456 0.02548418 0.02548413 0.02548412 0.02548411 0.02553389 0.02548412 0.02548411
0.02548412 0.02548411 0.02548413 0.04223869 0.03725469 0.02548412 0.02548435 0.02548484
0.02548411 0.02548412 0.02548411 0.02548414 0.02548414 0.02548413 0.02832249 0.02549679
0.02585558 0.02548473 0.02548438 0.02548425 0.02548412 0.02548412 0.02548412], shape=(38,),
dtype=float32)
Predicted Class: Pepper_bell___Bacterial_spot


Screenshot (285).png 286.7KB
0 1 2 3 4 5 6 7 8 9 10 11 12
0 0.0008 0.0091 0 0.0001 0.0245 0.0001 0 0 0.0002 0 0 0
tf.Tensor([0.02525516 0.02546454 0.02523417 0.02523578 0.02585989 0.02523553 0.02523457
0.02523415 0.02523785 0.02523488 0.02523484 0.025234 0.02523421 0.02523408 0.02523481 0.06495553
0.02524174 0.02524531 0.02523615 0.02569179 0.02523547 0.02525685 0.02532354 0.02523481
0.02523474 0.02523482 0.02523487 0.02523837 0.0252341 0.02523467 0.0252342 0.025235 0.02523485
0.02523483 0.02523482 0.02523483 0.0252343 0.02523415], shape=(38,), dtype=float32)
Predicted Class: Orange___Haunglongbing_(Citrus_greening)

Feasibility Analysis

The process involves capturing images of leaves, preprocessing the images to extract relevant features, and then using machine learning algorithms to classify the image into different categories, such as healthy or diseased.

1. Data Availability: A sufficient amount of data is required to train machine learning models for accurate disease detection. The data should be representative of the different types of diseases and healthy leaves. If the data is not available or insufficient, it may impact the accuracy of the model.
2. Image Pre-processing: Pre-processing is crucial in extracting relevant features from images that can help in the classification of healthy and diseased leaves. Some of the pre-processing techniques include normalization, filtering, and segmentation.
3. Feature Extraction: Feature extraction involves identifying the relevant features in the pre-processed images. The features can be extracted using techniques such as texture analysis, color analysis, and shape analysis.
4. Machine Learning Algorithms: Different machine learning algorithms such as convolutional neural networks (CNNs), support vector machines (SVMs), and decision trees can be used for classification. The choice of algorithm will depend on the nature and complexity of the problem.
5. Performance Evaluation: The accuracy of the model needs to be evaluated to assess its performance. Metrics such as precision, recall, and F1 score can be used for performance evaluation.

In conclusion, leaf disease detection using image pre-processing and machine learning is a feasible approach. However, careful consideration needs to be given to data availability, image pre-processing, feature extraction, machine learning algorithms, and performance evaluation for accurate disease detection.

Innovation or Uniqueness of the solution

1. The architecture of the CNN model used in here is EfficientNetB0.The uniqueness of the EfficientNetB0 architecture lies in its efficient scaling method that balances the depth, width, and resolution of the network. The architecture uses a compound scaling method that uniformly scales the dimensions of depth, width, and resolution with a fixed set of scaling coefficients. This ensures that the model architecture is not only efficient but also effective.
2. Data Augmentation such as Random Resized Crop , randomly crops the input image and resizes it to the desired size. In this case, the size is set to 224x224 pixels. In HorizontalFlip the image is projected horizontally with a probability of 0.5.
3. Dataset: The dataset we have chosen covers 38 classes ranging from plants like apple , raspberry,squash and potato to tomatoes,pepper and strawberry.

Impact or Usefulness of the solution

The solution for leaf disease detection using machine learning and image processing can have a significant impact on agriculture and food security. Here are some of the potential benefits and usefulness of this solution:

- Early detection: The system can detect leaf diseases at an early stage, allowing farmers to take necessary actions to prevent the disease from spreading and causing significant crop losses. Early detection can lead to timely treatment, resulting in reduced use of chemicals and increased yield.
- Increased efficiency: Traditional methods of disease detection rely on visual inspection, which can be time-consuming and prone to errors. The machine learning-based system can analyze images of leaves quickly and accurately, allowing farmers to take necessary actions promptly, leading to increased efficiency in disease management.
- Reduced costs: Crop losses due to leaf diseases can result in significant financial losses for farmers. The early detection and timely treatment of diseases can reduce the need for expensive chemicals and increase crop yields, resulting in reduced costs and improved profitability.
- User-friendly: The system can be designed to be user-friendly, allowing farmers with limited technical expertise to use it easily. The user-friendly system can provide recommendations for effective disease management strategies, leading to increased adoption and use by farmers.
- Improved food security: The solution can help to reduce crop losses due to leaf diseases, leading to increased food production and improved food security.

Future Scope

- Integration with IoT: The integration of Internet of Things (IoT) technology can enable real-time monitoring and detection of leaf diseases. IoT sensors can be installed in fields and greenhouses to continuously monitor the health of plants and detect any signs of disease or stress. This technology can provide farmers and growers with timely alerts, enabling them to take immediate action to prevent the spread of disease.

- Integration with Unmanned Aerial Vehicles (UAVs): The use of UAVs or drones equipped with cameras can provide high-resolution images of plants in real-time. These images can be used to detect leaf diseases with high accuracy, and the use of drones can enable large-scale monitoring of crops in a short amount of time. Integration with UAV technology can be particularly useful for monitoring large farms or fields that are difficult to access.

- Improved accuracy using advanced algorithms: Although current algorithms such as CNN and SVM are already effective in detecting leaf diseases, there is still room for improvement in terms of accuracy. Advanced machine learning algorithms such as deep learning, reinforcement learning, and transfer learning can be explored to improve the accuracy of leaf disease detection.

- Multispectral imaging: Multispectral imaging is a technique that uses a range of wavelengths of light to capture images of plants. This technique can provide more detailed and accurate information about the health of plants, including the detection of leaf diseases. Research is ongoing to develop multispectral imaging techniques for leaf disease detection.

- Mobile applications: Mobile applications can be developed for leaf disease detection, allowing farmers and growers to quickly and easily diagnose plant diseases using their smartphones. These applications can use machine learning algorithms to analyze images of plants taken with the smartphone camera and provide real-time diagnoses.

Summary

Leaf disease detection using machine learning and image processing is a growing area of research that aims to help farmers and researchers detect plant diseases at an early stage to prevent crop losses. The process involves capturing images of the plant leaves and using machine learning algorithms to classify the images into healthy or diseased leaves.

The first step in the process is to acquire images of plant leaves, which can be done using cameras or smartphones. The images are then preprocessed to enhance the contrast, remove noise, and normalize the lighting conditions.

Next, features are extracted from the preprocessed images using image processing techniques such as edge detection, color segmentation, and texture analysis. The extracted features are then fed into machine learning algorithms such as support vector machines (SVM), convolutional neural networks (CNN), and decision trees to classify the images as healthy or diseased.

One of the key challenges in leaf disease detection is the large variation in the appearance of the leaves, depending on factors such as lighting, angle, and disease severity. To address this, researchers have developed algorithms that are robust to these variations and can accurately detect diseases across different plant species and environments.

Overall, leaf disease detection using machine learning and image processing has the potential to revolutionize the way plant diseases are diagnosed and managed, leading to improved crop yields and food security.

References

[1] L. S. Pushpa Annabel, T. Annapoorani and P. Deepalakshmi, "Machine Learning for Plant Leaf Disease Detection and Classification – A Review," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0538-0542, doi: 10.1109/ICCSP.2019.8698004.

[2] K. R. Gavhale and U. Gawande, “[PDF] an overview of the research on plant leaves disease detection using image processing techniques: Semantic scholar,” IOSR Journal of Computer Engineering, https://www.semanticscholar.org/paper/An-Overview-of-the-Research-on-Plant-Leaves-Disease-Gavhale-Gawande/7792ab664d5ab93c5e8c45b0c772627d40c962eb (accessed May 9, 2023).

[3] R. M. Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-4, doi: 10.1109/ICIIECS.2017.8275915.

[4] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," 2016 Conference on Advances in Signal Processing (CASP), Pune, India, 2016, pp. 175-179, doi: 10.1109/CASP.2016.7746160.

[5] P. Tm, A. Pranathi, K. Sai Ashritha, N. B. Chittaragi and S. G. Koolagudi, "Tomato Leaf Disease Detection Using Convolutional Neural Networks," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, India, 2018, pp. 1-5, doi: 10.1109/IC3.2018.8530532.

[6] Prajakta Mitkal, Priyanka Pawar, Mira Nagane, Priyanka Bhosale, Mira Padwal and Priti Nagane “Leaf Disease Detection and Prevention Using Image processing using Matlab” International Journal of Recent Trends in Engineering & Research (IJRTER) Volume 02, Issue 02; February– 2016 [ISSN:2455-1457]

References

[7] Dheeb Al Bashish, Malik Braik and Sulieman BaniAhmad "A Framework for Detection and Classification of Plant Leaf and Stem Diseases" 2010 IEEE International Conference on Signal and Image Processing, pp. 978-1- 4244-8594-9/10.

[8] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff and Rohani Binti S Mohamed Farook "Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques" 2012 IEEE Third International Conference on Intelligent Systems Modelling and Simulation, pp. 978-0-7695-4668-1/12.

[9] Sabah Bashir and Navdeep Sharma "Remote Area Plant Disease Detection Using Image Processing" IOSR Journal of Electronics and Communication Engineering (IOSRJECE)ISSN : 2278-2834 Volume 2, Issue 6 (Sep-Oct 2012), PP 31-34.

[10] Usama Mokhtar, Mona A. S. Alit, Aboul Ella Hassenian, Hesham Hefny "Tomato leaves diseases detection approach based on support vector machines" 2015 IEEE pp. 978-1-5090-0275-7/15.

[11] Sachin D. Khirade, A. B. Patil, "Plant Disease Detection Using Image Processing" 2015 IEEE International Conference on Computing Communication Control and Automation, pp. 978-1-4799-6892-3/15.

[12] Ghulam Mustafa Choudhary and Vikrant Gulati "Advance in Image Processing for Detection of Plant Diseases" International Journal of Advanced Research in Computer Science and Software Engineering 5(7), July- 2015, pp. 1090-1093 [ISSN: 2277 128X].