Plant Leaf Disease Detection Using CNN

Chandrapal Chauhan[1], Assistant Professor, Department of Computer Science and Engineering, Government College ofEngineering Chandrapur, Maharashtra, India.

Mohit Bele[2], Ankit Dongare[3], Achal Dongare[4], Sakshi Nikhade[5], Sakshi Salwe[6].

[23456], Final Year Students, Department ofComputer Science and Engineering, Government college of Engineering Chandrapur, Maharashtra, India.

Abstract-

The existence of pests and diseases in plants and crops has a substantial impact on agricultural production within a country. Monitoring plants meticulously to detect and identify diseases is a common practice among farmers and experts. However, this approach is often laborious, costly, and not entirely reliable. To mitigate this issue, we propose a Disease Recognition Model based on leaf image classification. Our objective is to detect plant diseases using image processing techniques, specifically leveraging Convolutional Neural Networks (CNNs). CNNs are a category of artificial neural networks tailored for handling pixel-based inputs, particularly adept at image recognition tasks.

Abbreviation-

CNN (Convolutional Neural Network), SVM (Support Vector Machine), KNN (K Nearest Neighbour), ANN (Artificial Neural Network), GPU (Graphics processing unit)

1. Introduction

Agricultural production has been a fundamental method of obtaining sustenance for centuries, playing a crucial role in providing livelihoods for people worldwide. The significance of plants extends beyond human consumption, as they also serve as essential resources for animals, supplying nourishment, oxygen, and various essentials. Governments and experts are actively engaged in substantial efforts to enhance food production, addressing real-world challenges effectively.

When plants become infected, it impacts the entire ecosystem, affecting various parts like stems, leaves, and branches. The types of diseases affecting plants, such as bacterial or fungal infections, can vary based on factors like climate conditions. Food insecurity affects millions due to insufficient crop yields, exacerbated by significant changes in weather patterns that impact plant growth—a natural disaster that cannot be avoided.

Early detection of plant diseases is crucial to prevent substantial crop losses. Farmers must use appropriate pesticides to protect crops and farmland, seeking expert advice to avoid overuse of chemicals. Researchers focus on plants to assist farmers and stakeholders in agriculture. Visible infections are easy to identify, treatable when detected early through regular crop monitoring. However, this process is effective mainly for rare infections or low crop yields.

The introduction of automated disease detection tools benefits farmers in both small and large-scale rural development projects. These tools deliver accurate results quickly, relying heavily on deep learning and neural networks for operation. The study utilizes a Deep Convolutional Neural Network to distinguish between healthy and infected leaves, aiding in the identification of plant diseases based on leaf images, ensuring precise outcomes aligned with the input data.

2. Related Work

1. CNN based Leaf Disease Identification and Remedy Recommendation System[1].

Publication Year: 2019

Author: Sunku Rohan , Triveni S Pujar ,Suma VR Amog Shetty, Rishabh F Tated [1].

Journal Name: IEEE conference paper

Summary: The agricultural sector plays a significant role in our lives and is crucial for the economy. Farmers face challenges in identifying leaf diseases, leading to reduced crop yields. However, utilizing videos and imagesof leaves provides a better understanding for agricultural researchers, leading to improved solutions foraddressing crop diseases. This approach aims to solve issues related to crop diseases efficiently.

It's important to note that the productivity of crops is directly related to their health; unhealthy crops are less likely to provide good nutrition. With advancements in technology, smart devices can now detect and identify plant diseases faster, allowing for quicker treatment to minimize the negative impacts on crop yields.

This paper focuses on using image processing techniques for plant disease detection. It utilizes an open dataset comprising 5000 images of healthy and infected plant leaves, employing semi-supervised methods to classify different types of crops and identify diseases within four categories.

2. GUI based Detection of Unhealthy Leaves using Image Processing Techniques [2]

Publication Year: 2019

Author: Vellanki Krishna Vamsi, Velamakanni Sahithya, Brahmadevara Saivihari, Parvathreddy Sandeep Reddy and Karthigha Balamurugan [2]

Journal Name: International Conference on Communication and Signal Processing.

Summary:
Enhancing agricultural productivity is key to improving the Indian economy. With this goal in mind, this paper focuses on utilizing image processing techniques to identify unhealthy leaves, particularly in women finger plants. The study aims to detect early signs of various diseases such as yellow mosaic vein, leaf spot, and fine mold.

The process involves capturing leaf images, processing them, segmenting them, extracting features, and classifying them as healthy or unhealthy. Given the challenges posed by diverse climatic conditions and landscapes, noisy image datasets are also considered. K-Means clustering is employed for segmentation, while SVM and ANN are used for classification. PCA is utilizedto reduce the feature set.

The findings indicate that using Support Vector Machine (SVM) and Artificial Neural Network (ANN) resulted in an average detection accuracy of 85% and 97%, respectively. Removing noise improved these accuracies to 92% and 98%, respectively. This study lays the groundwork for full automation in agricultural industries.

3. Materials And Methods

Dataset:
The dataset used in this project is sourced from Kaggle's Plant Village dataset, which is available online. The code required for computation and analysis of training loss and validation was obtained directly from the Kaggle platform, ensuring accurate and thorough assessment of the model's performance.

Image Preprocessing and Labelling:

Image pre-processing is a crucial step that involves various techniques such as eliminating low-frequency background noise, normalizing the brightness of individual particles' images, removing reflections, and masking portions of images. This process aims to improve data quality and enhance the learning process.

During image pre-processing, all images were manually adjusted to include the region of interest (plant leaves) by creating a square around them. Only images with a resolution and size greater than or equal to 500 pixels were considered significant for the dataset. Moreover, images where the region of interest was clearly visible were

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

prioritized for inclusion, ensuring that the dataset contains relevant information for feature learning.

While numerous resources are available online, their relevance can be uncertain. To ensure the accuracy of classes in the dataset, agricultural experts reviewed leaf images and labeled them with appropriate disease abbreviations. It's crucial to use accurately labeled images for both training and validation datasets to build a reliable and robust detection model. In this process, duplicate images that remained after the initial gathering and categorization were removed from the dataset.Neural Network Training :

```
# getting all predictions (actual label vs predicted)
for i, (img, label) in enumerate(test):
    print('Label:', test_images[i], ', Predicted:', predict_image(img, model))
Label: AppleCedarRust1.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust2.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust3.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust4.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleScab1.JPG , Predicted: Apple___Apple_scab
Label: AppleScab2.JPG , Predicted: Apple___Apple_scab
Label: AppleScab3.JPG , Predicted: Apple___Apple_scab
Label: CornCommonRust1.JPG , Predicted: Corn_(maize)___Common_rust_
Label: CornCommonRust2.JPG , Predicted: Corn_(maize)___Common_rust_
Label: CornCommonRust3.JPG , Predicted: Corn_(maize)___Common_rust_
Label: PotatoEarlyBlight1.JPG , Predicted: Potato___Early_blight
Label: PotatoEarlyBlight2.JPG , Predicted: Potato___Early_blight
Label: PotatoEarlyBlight3.JPG , Predicted: Potato___Early_blight
Label: PotatoEarlyBlight4.JPG , Predicted: Potato___Early_blight
Label: PotatoEarlyBlight5.JPG , Predicted: Potato___Early_blight
Label: PotatoHealthy1.JPG , Predicted: Potato___healthy
Label: PotatoHealthy2.JPG , Predicted: Potato___healthy
Label: TomatoEarlyBlight1.JPG , Predicted: Tomato___Early_blight
Label: TomatoEarlyBlight2.JPG , Predicted: Tomato___Early_blight
Label: TomatoEarlyBlight3.JPG , Predicted: Tomato___Early_blight
Label: TomatoEarlyBlight4.JPG , Predicted: Tomato___Early_blight
Label: TomatoEarlyBlight5.JPG , Predicted: Tomato___Early_blight
Label: TomatoEarlyBlight6.JPG , Predicted: Tomato___Early_blight
Label: TomatoHealthy1.JPG , Predicted: Tomato___healthy
Label: TomatoHealthy2.JPG , Predicted: Tomato___healthy
Label: TomatoHealthy3.JPG , Predicted: Tomato___healthy
Label: TomatoHealthy4.JPG , Predicted: Tomato___healthy
```

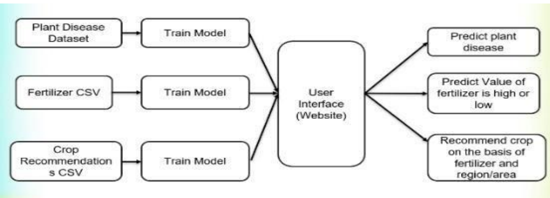

Fig. Block diagram For plant Disease Detection

The proposal involves training a significant convolutional neural network (CNN) using the Tensorflow library, an open-source software tool for mathematical computations using data flow graphs. Tensorflow was initially developed by researchers and engineers at Google Brain toadvance AI and deep neural networks research but has since become widely applicable across various domains.

CNNs, inspired by the structure of the animal visual cortex, utilize convolution operations to process data images efficiently. They are composed of layers of receptive fields, small neuron combinations that process different parts of the image, leading to a higher-resolution representation of the original image. CNNs may also incorporate pooling layers to consolidate the outputs of neuron clusters and improve performance.

One key advantage of CNNs is the use of shared weights in convolutional layers, reducing memory usage and enhancing performance. Rectified Linear Units (ReLU) are employed as activation functions to introduce nonlinearity and improve training accuracy without requiring data normalization. Neurons in a hidden layer of a CNN are organized into feature maps, where each neuronwithin a map shares the same weights and bias, allowing them to detect specific features across different regions ofthe input image.

The convolutional layer, a fundamental component of CNNs, comprises learnable channels or filters that convolve over the input volume, detecting features and generating activation maps. Hyperparameters such as depth, stride, and zero-padding control the output volume of the convolutional layer, influencing the model's performance.

Overall, CNNs are powerful tools for image processing tasks, offering efficient feature extraction, shared weight utilization, and improved training accuracy, making them widely applicable in various domains including image and video recognition, recommender systems, and natural language processing.

Convolutional Neural Network Architecture-

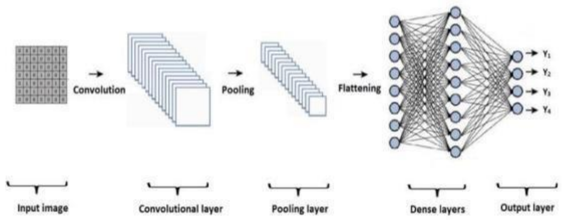

Fig : CNN Architecture

A Convolutional Neural Network (CNN) is composed of three main layers: a convolutional layer, a pooling layer, and a fully connected layer. These layers work together to process and extract features from input data. The image below illustrates the structure of a CNN with all three layers depicted together.

Proposed System-

We are in the process of constructing a neural network model specifically designed for the task of image classification..

Plant Disease Prediction: A CNN model, trained on a plant disease dataset, is used to analyze images uploaded through a user interface and predict the presence of specific diseases.

Fertilizer and Crop Recommendations: A separate model, presumably trained on a dataset of fertilizer data and successful crop yields in various regions, recommends suitable crops based on user input regarding fertilizer type and region. It can also predict whether a high or low amount of fertilizer should be used.

The system appears to function in two stages:

Training: First, a CNN model is trained on a dataset of images labeled with specific plant diseases. Then, a separate model is trained on data that links fertilizer types, regions, and successful crop yields.

Prediction: During use, a user uploads an image of a plant through a user interface. The CNN model analyzes the imageto predict any present diseases. The user also provides data onthe fertilizer type and the region where the plant will be grown. The fertilizer and crop recommendation model then leverages this information to suggest suitable crops and fertilizer

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

application amounts.

In essence, this system combines disease prediction with agricultural data to give users insights into plant health, suitable crops, and fertilizer use. Here are some additional points to consider:

The effectiveness of the plant disease prediction model would depend on the quality and size of the training dataset.

The accuracy of the fertilizer and crop recommendation model would rely on the comprehensiveness of the training data on fertilizer use, crop yields, and regional factors.

4. CNN Model Steps

Conv2D Layer: This layer conducts a 2D convolution operation by applying a convolution kernel to input data, generating an output tensor representing extracted features.

Maxpooling Technique: Max pooling selects the maximum value within each region of the feature map covered by a filter, aiding in reducing spatial dimensions while retaining crucial features.

Flatten Layer: Placed between the convolutional and fully connected layers, Flatten converts the 2D feature matrix into a flattened vector format, fed into the fully connected neural network for classification.

Image Data Generation: This tool automates batch generation of preprocessed image tensors from stored image files, simplifying neural network training with image data.

Training Process: Training involves assessment, motivation, design, delivery, and evaluation stages, ensuring effective learning and model refinement from pre-training to completion.

Epochs: An epoch signifies one complete pass of the training dataset through the machine learning algorithm, often divided into batches for efficient training.

Validation Procedure: Validation evaluates a trained model using a separate testing dataset, assessing its generalization and performance on unseen data.

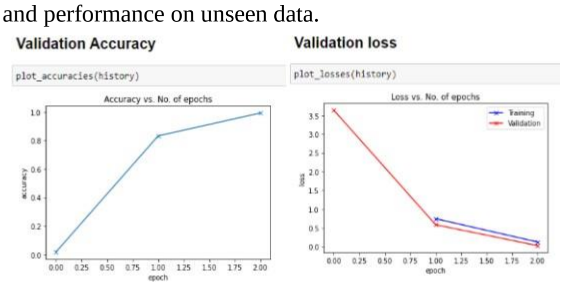

Fig: Validation Accuracy and Loss

Training and Testing Model: The dataset undergoes preprocessing steps like image reshaping, resizing, and conversion to an array format. Similar preprocessing is applied to test images. The dataset contains various plant leaf diseases, and any image from this dataset can be used as a test image to assess the software's performance.

```
%%time
history += fit_OneCycle(epochs, max_lr, model, train_dl, valid_dl,
                        grad_clip=grad_clip,
                        weight_decay=1e-4,
                        opt_func=opt_func)

Epoch [0], last_lr: 0.00812, train_loss: 0.7466, val_loss: 0.5865, val_acc: 0.8319
Epoch [1], last_lr: 0.00000, train_loss: 0.1248, val_loss: 0.0269, val_acc: 0.9923
CPU times: user 11min 16s, sys: 7min 13s, total: 18min 30s
Wall time: 19min 53s
```

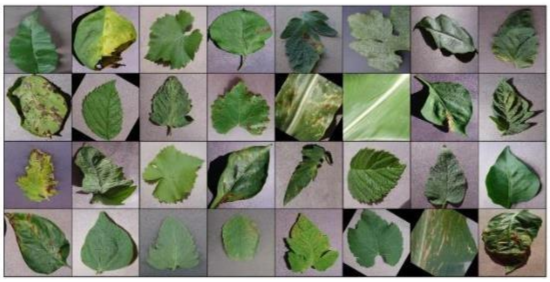

Fig. 2. Pre-processed images - Phase Two augmentation settings = brightness changes (0.4,0.7), warp (0.5), flipping (random), padding mode (reflection)

Conclusion & Future work

We have achieved significant progress in developing a deep learning model tailored for the automatic detection and classification of plant leaf diseases. This model underwent rigorous testing across 13 diverse plant species, including tomato, strawberry, soybean, raspberry, potato, corn, Pepper bell, peach, orange, grape, cherry, blueberry, and apple. The dataset utilized for training and evaluation encompassed 38 distinct classes of plant diseases.

Our methodology entailed leveraging the image data generator API provided by Keras for essential image preprocessing tasks. Additionally, we integrated the VGG-19 model, a sophisticated convolutional neural network architecture, and conducted thorough training using the dataset to facilitate accurate predictions. The model's predictions have exhibited a commendable level of precision.

Moreover, we have successfully integrated these models into a web application for broader accessibility. Our current focus revolves around enhancing the accuracy and performance of both the web application and the underlying deep learning model.

References-

1. A study on leaf disease identification and remedy recommendation using convolutional neural networks (CNNs) was conducted (ResearchGate, 2019). [Link](https://www.researchgate.net/publication/335579139_CNN_based_Leaf_Disease_Identification_and_Remedy_Recommendation_System)

2. Another research project focused on detecting unhealthy leaves through GUI-based image processing techniques (ResearchGate, 2018). [Link](https://www.researchgate.net/publication/332672970_GUI_based_Detection_of_Unhealthy_Leaves_using_Image_Processing_Techniques)

3. Hunter, M., Smith, R., Schipanski, M., Atwood, L., & Mortensen, D. published an article in BioScience in 2017, discussing recalibrated targets for sustainable intensification in agriculture (Hunter et al., 2017). [Link](https://doi.org/10.1093/biosci/bix010)

4. The PlantVillage platform, a valuable resource for agricultural information, is accessible online (PlantVillage, 2020). [Link](https://plantvillage.psu.edu/)

5. Klauser discussed challenges in monitoring and managing plant diseases in developing countries in a 2018 publication (Klauser, 2018). [Link](https://doi.org/10.1007/s41348-018-0145-9)

6. Muimba-Kankolongo's book on food crop production by smallholder farmers in Southern Africa provides insights into agricultural practices (Muimba-Kankolongo, 2018).

7. Boulent, J., Foucher, S., Théau, J., & St-Charles, P. conducted research on automatic plant disease identification using convolutional neural networks in 2019 (Boulent et al., 2019). [Link](https://doi.org/10.3389/fpls.2019.00941)

8. iNaturalist is an online platform for nature observation and documentation (iNaturalist, 2020). [Link](https://www.inaturalist.org/)

9. Imagga Technologies Ltd. shared a success story about training the world's largest plant recognition classifier on PlantSnap (Imagga Ltd., 2020). [Link](https://imagga.com/success-stories/plantsnapcasestudy)

10. Statista provides statistics on smartphone users worldwide as of 2020 (Statista, 2020). [Link](https://www.statista.com/statistics/330695/numberofsmartphone-users-worldwide/)