Emmanuel Adebayo

Professor Mihran Tuceryan

Final Semester Project Report

April 19, 2024

Fruit Disease Detection Using Deep Learning

Fruit disease pose a significant threat to agricultural productivity and food security worldwide. In the paper titled “PLANT DISEASE: A Threat to Global Food Security”, The authors claimed that more than 800 million people are going through food scarcity around the world. The interesting part is ten percent of the world food production is affected because of plant disease. In this project, I explore the use of modern computer vision systems such as Deep learning (Convolutional Neural Network) to see if it can be useful in tackling the issues such as early detection of fruit diseases, quality control, traceability, and monitoring surveillance.

Dataset Description

he dataset, named MangoFruitDDS, comprises 1700 images of mango fruits affected by four common diseases: Alternaria, Anthracnose, Black Mold Rot, and Stem and Rot, along with a category for healthy fruits. Captured in Senegal's orchards using a mobile phone camera, the images are stored in JPG format and are 224x224 pixels in size. Two versions of the dataset are provided: SenMangoFruitDDS_original, containing unprocessed images, and SenMangoFruitDDS_bgremoved, where background elements are eliminated to enhance focus on the fruits. This preprocessing step aims to reduce noise and improve the accuracy of disease classification algorithms. The dataset facilitates research in agricultural science, computer vision, and machine learning for automated mango fruit disease diagnosis and monitoring, contributing to improved mango farming practices.

Methodology

This report presents the architecture, training process, and evaluation of a Convolutional Neural Network (CNN) for classifying mango fruit diseases. The CNN, developed using the Keras deep learning library, aims to categorize images of mango fruits into distinct disease classes based on visual attributes. The primary goal is to assist in early disease detection and facilitate timely intervention strategies to minimize crop losses. The CNN architecture comprises convolutional layers followed by max-pooling layers for feature extraction and down sampling. Two convolutional layers with ReLU activation are utilized, with the first layer applying 32 filters and the second layer applying 64 filters. Max-pooling layers are incorporated to reduce spatial dimensions and capture essential features effectively. Following the convolutional layers, flattened feature maps are fed into dense layers. The model includes two dense layers, with 128 units each and ReLU activation, followed by an output layer with units equal to the number of disease classes and SoftMax activation for multi-class classification. The model is trained using a training dataset consisting of preprocessed images resized to a target size of (224, 224). Training is performed using the Adam optimizer with sparse categorical cross-entropy loss. The training and testing set were split with a eighty to twenty percent ratio.

Experiments

In this study, we aimed to enhance the accuracy of an image classification CNN for identifying mango fruit diseases by employing image augmentation techniques. The original dataset consisted of images depicting various mango fruit diseases, including Alternaria, Anthracnose, Black Mold Rot, Stem and Rot, as well as healthy fruits. Given the limited size of the dataset, we employed image augmentation techniques such as rotation, flipping, and blurring to generate additional training samples. Rotation and flipping variations were applied to the images to simulate different orientations and viewpoints, while blurring was used to introduce variations in image sharpness and texture. Despite the application of image augmentation (generating 100 more data), the fundamental characteristics and features of the original images were preserved to ensure the model's robustness and generalization capability. The CNN model was trained using the augmented dataset with an epoch size of 10, and the training and testing accuracies were evaluated. While the image augmentation techniques successfully increased the dataset's size and diversity, the impact on the model's performance was modest. Although the training accuracy approached 98%, indicating that the model effectively learned the training data's patterns, the testing accuracy plateaued at around 55%. This discrepancy between training and testing accuracies suggests that the model may have overfit to the training data, resulting in reduced generalization performance on unseen test samples. Despite the limited improvement in accuracy achieved through image augmentation, the augmented dataset still provided valuable training samples for the model, enabling it to learn diverse representations of mango fruit diseases and healthy fruits. The depth and the overall model architecture was also increased but this showed no significant improvement. It is now evident that the failure of the model to generalize well is due to not having enough data.

Applications and Future directions

1. **Farmers' Decision Support:** The Crop Disease Detection App empowers farmers to make informed decisions regarding disease management and treatment, helping them mitigate crop losses and improve yields. By providing timely and accurate disease diagnosis, the app enables farmers to take proactive measures to protect their crops and livelihoods.

2. **Precision Agriculture:** The app contributes to the advancement of precision agriculture practices by enabling targeted interventions for disease control and management. By precisely identifying diseased areas within a field, farmers can apply pesticides, fungicides, or other treatments only where necessary, reducing input costs and minimizing environmental impact.
3. **Epidemiological Monitoring:** The app facilitates the monitoring and surveillance of crop diseases at a regional or global scale by aggregating and analyzing data from multiple users. By tracking disease outbreaks and patterns over time, agricultural authorities and researchers gain valuable insights into disease dynamics, contributing to the development of more effective disease management strategies and policies.
4. **Capacity Building and Extension Services:** The app serves as a valuable educational tool for farmers, extension workers, and agricultural practitioners, providing access to information and resources on crop diseases, their symptoms, and management practices. Through interactive tutorials, case studies, and expert advice, users can enhance their knowledge and skills in disease detection and control, ultimately improving agricultural productivity and food security.

Conclusion:

In conclusion, while image augmentation techniques such as rotation, flipping, and blurring contributed to the dataset's diversity and increased the training samples, their impact on improving the model's accuracy was limited. Future research directions may involve exploring alternative augmentation strategies, fine-tuning augmentation parameters, or incorporating other data preprocessing techniques to further enhance the model's performance. Additionally, the integration of transfer learning or ensemble learning approaches could be explored to leverage pre-trained models or combine multiple classifiers for improved classification accuracy. Overall, while image augmentation is a valuable tool for dataset expansion and regularization, its effectiveness may vary depending on the dataset's characteristics and the model's architecture, highlighting the importance of careful experimentation and optimization in CNN-based image classification tasks.

References:

Chirag Chauhan. Mangofruitdds: A dataset of mango fruit diseases. https://www.kaggle.com/datasets/warcoder/mangofruitdds, 2023.

Strange, Richard N., and Peter R. Scott. "Plant disease: a threat to global food security." *Annu. Rev. Phytopathol.* 43 (2005): 83-116.