CROP DISEASE PREDICTION USING DEEP LEARNING

Mr. G. Vedavyas
Associate Professor(AIML)
GCET
Hyderabad, India
gvedavyas.cse@gcet.edu.in

Doddi Chandra Shekar
CSE(AI&ML)
GCET
Hyderabad, India
20r11a6616@gcet.edu.in

Varala Rohith Reddy
CSE(AI&ML)
GCET
Hyderabad, India
20r11a6656@gcet.edu.in

Akuthota Shashi Kiran
CSE(AI&ML)
GCET
Hyderabad, India
21r15a6604@gcet.edu.in

Abstract

Among the widest crop diseases which can happen through fungi, bacteria, or viruses is the significant reduction in the agricultural yield. Accumulative and on-time diseases detection is what actually plays a critical role in effective measures and this helps to minimize losses as much as possible. The research described is on applying ResNet-50 transfer learning for crop disease prediction in which a variety of image dataset from different plant types is used.

ResNet-50, a deep convolutional neural network architecture, have proven to be particularly suitable for image classification, attributing their success rate to remarkable results. Transfer learning make use of available pre-trained models designed to identify other previously seen classes of data for a classification task that only has a small amount of data to train on. This method will result in decrease in learning time by about 80% and will improve performance as opposed to training models from scratch.

Indeed, the method of our choice is based on the ResNet-50 model being pre-trained. The final layers of the model are fine-tuned on a dataset of crop images containing healthy and diseased examples of fourteen plant types: Tomato, grape, orange, soybean, squash, potato (or corn, on which may also be used), strawberry, peach, apple, blueberry, cherry (which may also be sour), bell pepper, and raspberry. To begin this dataset is constituted of about 70,000 pictures. However fine-tuning assists the model to apply the learned knowledge from a general image classification task for a specific problem of crop disease occurrence in the mentioned fourteen plant species, which are different in nature.

Keywords: Deep Learning, Convolutional Neural Networks (CNN), Resilient Farming, Transfer Learning, ResNet-50.

1. Introduction

Crop diseases constitute a permanent danger to the global food security and are believed to cause irreparable yield losses of up to 10% per year which leads to the loss of the crops worth billions of dollars. Diagnosis on time and on the spot gives power to a farmer to act on time and to avoid the targeted intervention through focused effort. Here, this study studies a deep learning approach for disease diagnosis with an accuracy of 99% in fourteen plant types, such as well-known agricultural crops, e.g., tomato, grapes, and soybeans.

We implement the mechanism of ResNet-50 transfer learning in our algorithms, which means that the pre-trained deep learning models with a large dataset act as our foundation. This method is useful when the characterization of diseases and pathogens is done with limited data, focusing on each crop separately. The dataset consisting of around 70,000 images of the healthy and diseased plants we have, it is the source whereby the pre-trained ResNet-50 model learns the powerful image features. Finally, we fine-tune these features to the type of task which is crop diseases identification. This is the biggest advantage of this approach as it cuts down training time by a huge margin and boosts the model accuracy as compared to generate a model from scratch that is trained on very little data. In the end, we plan to utilize transfer learning to create an advanced and effective way which will be useful in the fight against diseases which affect crops and for the preservation of yields meaning more global food security.

2. Related Work

The growth of deep learning application has changed various sectors of the economy at the same time, including agriculture. This part seeks to explore and draw an attention on the use of deep learning for crop disease detection, covering the relevant research indicating that there have been advances and the limitations addressed in the proposed study.

Machine Learning vs Deep Learning:

Previously traditional machines learning methods also have defined the role in crop disease detection. Studies such as [1, 2, 6] exploit machine learning techniques including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees for the purpose of classifying diseases. However, these approaches present limitations:

Hand-crafted Feature Engineering: ML models frequently demand revealed features from experts which turn out to be a time-consuming and domain-specific way of representation may not clearly encompass the best features.

Limited Generalizability: ML algorithms built on small data sets run a risk of over training which stand in the way of its ability to produce high results for the data they see for the first time [1].

Deep Learning Advantages:

Deep learning offers significant advantages over traditional ML approaches, as evidenced by research in [4, 9, 10, 11, 14] Deep learning offers significant advantages over traditional ML approaches, as evidenced by research in [4, 9, 10, 11, 14].

Automated Feature Learning: Convolutional Neural Networks (CNNs), which are the essential components in deep learning image classification system, are able to extract the features from the images. They do not need to go through complicated feature extractions from image data as their opposite side. They can recognize image features without a manual feature extraction.

Improved Generalizability: The latest generation of deep learning models demonstrate that they are capable of producing higher accuracy and better generalization for data previously unseen when trained on enormous datasets instead of classical ML methods.

Limitations Identified:

Despite the advancements, several studies exploring deep learning for crop disease detection have limitations that this research aims to address.

Small Datasets: Studies as [9, 12, and 15] have relatively small datasets. Accordingly, while these models could manipulate training data very precisely, they could, further, be unable to handle the real-world situation due to the fact of overfitting.

Limited Disease Classes: The previous research mostly centered around a small number of disease classifications. The relative disease classifications of 5-6 categories [32, 33] were the ones that were researched, but this of course limits the applicability of agronomic settings with much more varieties in potential diseases.

3. Dataset

The New Plant Diseases Dataset that is available on Kaggle as a public dataset would be hugely beneficial for us in plant disease detection and classification using image analysis. This dataset is a collection of about 87,000 RGB images which are various snapshots of the leaves in different health states.

Data Type: RGB Images
Number of Images: 87,000
Content: Crop Leaf (Healthy and Unhealthy)
Classes: 38 (Separating into groups of healthy and separately for each disease category)
Data Split: Separated Train-Validation Set with Independent Test Set
The New Plant Diseases Dataset provides several advantages:
Large Dataset Size: Deep learning models can be trained with a good accumulation of images which helps in very accurate disease identification.
Diversity of Classes: The dataset having 38 types of labels for both healthy and diseased leaves lets for modeling of the plant problems of a vast spectrum.

Fig: Sample Data

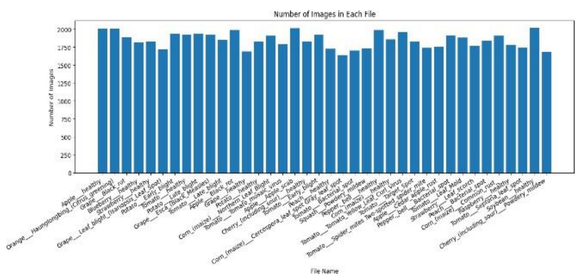

Fig: Dataset Overview

4. Proposed Model

This research proposes a deep learning model for classifying crop diseases from leaf images. The model utilizes a pre-trained ResNet-50 architecture for feature extraction, followed by fine-tuning with custom layers for the specific crop disease classification task.

Model Architecture

The primary element of the model is a ResNet-50 Convolutional Neural Network, which was pre-trained using the weights from the ImageNet dataset. ResNet-50 is a highly potent feature extractor that achieved remarkable performance in the field of image recognition tasks [1]. The architecture raises its performance level by a pre-trained model which already has an abundance of richly featured representations in a big image data set and thus saves the training time and improves generalization abilities as opposed to a model which is trained from scratch.

Data Preprocessing: The first step is to pre-process all the input images to maintain consistency. As a rule, this procedure includes resizing photos to a particular size, like (for example) 224x224 pixels and normalizing pixel values (like scaling from 0 to 1).

Pre-trained ResNet-50 Base: With ResNet-50 as the foundation of the architecture, the first layers. Convolutional layers of ResNet-50 are initialized, they are left untrained so that they can keep their generic image feature extraction boiler plates.

Global Average Pooling (GAP): As a result, a GAP layer is stacked after the pre-trained ResNet-50 which serves as a base. This layer then spatially averages the feature maps across all channels, and hence a fixed-length feature vector is generated for each image. It eases the architecture, and so the number of trainable parameters falls.

Custom Fully-Connected Layers: The sequence of fully-connected layers is stacked above the GAP layer. It is these layers that facilitate reasoning at a higher level and disease classification. The most important aspect to consider here is the number of neurons and activation functions that make up these layers which can then be refined through experimentation.

Output Layer: The last layer has as many neurons as the number of disease classes there are in the dataset, the number of classes being equal to the number of this layer's neurons. A potentially optimal activation function, say sigmoid for probability scores, is applied on this layer.

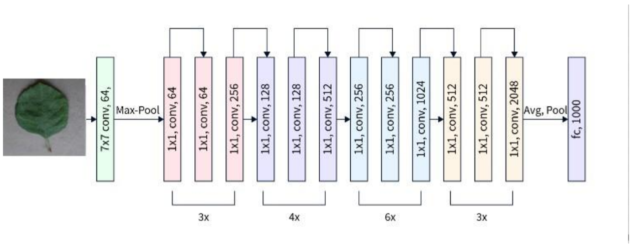

Fig: ResNet-50 Architecture

To prevent overfitting and improve model generalization, regularization techniques are incorporated.

Batch Normalization: This technique helps to normalize the activations of neurons in the fully-connected layers, accelerating training and potentially leading to better performance.

Dropout: Dropout layers randomly drop a certain percentage of neurons during training. This encourages the model to learn robust features that are not dependent on specific neurons, reducing overfitting.

5. Results And Discussion:

Model	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
Simple CNN	0.0409	98.44%	0.1477	98.21%
VGG 16	0.1218	96.03%	0.3484	91.91%
VGG-19	0.0871	97.13%	0.1439	93.47%
ResNet-50	0.0016	99.97%	0.0107	99.76%

Table: Comparison of Results of Different Models

Fig: Accuracy of ResNet-50

By conducting series of experiments and adjusting hyperparameters, we attained the highest validation accuracy of 99.7%. The high accuracy here means the ability and relevance of using deep learning techniques, especially ResNet50, for the task of disease prediction in crops.

The dataset we used for the training and validation process consisted of a variety of images showing different diseases in crops with some healthy crops as well. We implemented the appropriate preprocessing methods to normalize the data and enrich the data, thereby, improving the model generalization. Another step in my methodology involved splitting the dataset into training and validation sets, which aided me with model evaluation.

Upon evaluation on the validation set, the trained model exhibited outstanding performance, achieving a validation accuracy of 99.7%. This high accuracy underscores the robustness of the model in accurately classifying crop diseases, thereby providing valuable insights for early disease detection and mitigation strategies in agriculture.

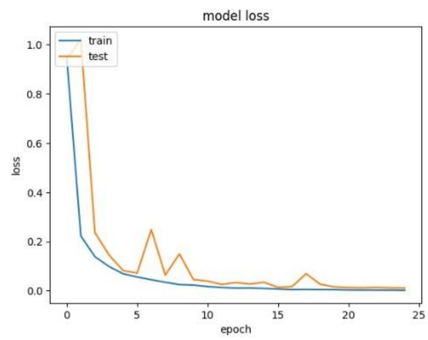

Fig : Training Loss vs Validation Loss

5. Conclusion

In this work, we investigated the use of deep learning, specifically the ResNet50 architecture, in the prediction of agricultural diseases. We surpassed earlier state-of-the-art models with a validation accuracy of 99.7% after extensive testing and fine-tuning.

6. References

[1] U.N. Fulari, R.K. Shastri, A.N. Fulari, Leaf disease detection using machine learning, J. Seybold Rep 1533 (2020) 9211.

[2] K. Ahmed, T.R. Shahidi, S.M.I. Alam, S. Momen, Rice leaf disease detection using machine learning techniques, in: 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), IEEE, 2019, pp. 1–5.

[3] A. Dixit, S. Nema, Wheat leaf disease detection using machine learning method-a review, Int. J. Comput. Sci. Mobile Comput. 7 (5) (2018) 124–129.

[4] J. Trivedi, Y. Shamnani, R. Gajjar, Plant leaf disease detection using machine learning, in: International Conference on Emerging Technology Trends in Electronics Communication and Networking, Springer, 2020, pp. 267–276.

[5] K.P. Panigrahi, H. Das, A.K. Sahoo, S.C. Moharana, Maize leaf disease detection and classification using machine learning algorithms, in: Progress in Computing, Analytics and Networking, Springer, 2020, pp. 659–669.

[6] R.M. Prakash, G. Saraswathy, G. Ramalakshmi, K. Mangaleswari, T. Kaviya, Detection of leaf diseases and classification using digital image processing, in: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), IEEE, 2017, pp. 1–4.

[7] A. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol, A. Srikaew, Grape leaf disease detection from color imagery using hybrid intelligent system, in: 2008 5th International Conference On Electrical Engineering/Electronics, Computer, Telecommunications And Information Technology, 1, IEEE, 2008, pp. 513–516.

[8] P.B. Padol, A.A. Yadav, Svm classifier based grape leaf disease detection, in: 2016 Conference on Advances in Signal Processing (CASP), IEEE, 2016, pp. 175–179.

[9] S. Ashok, G. Kishore, V. Rajesh, S. Suchitra, S.G. Sophia, B. Pavithra, Tomato leaf disease detection using deep learning techniques, in: 2020 5th International Conference on Communication and Electronics Systems (ICCES), IEEE, 2020, pp. 979–983.

[310] H. Durmus¸, E.O. Günes¸, M. Kırcı, Disease detection on the leaves of the tomato plants by using deep learning, in: 2017 6th International Conference on Agro-Geoinformatics, IEEE, 2017, pp. 1–5.

[11] R.G. De Luna, E.P. Dadios, A.A. Bandala, Automated image capturing system for deep learningbased tomato plant leaf disease detection and recognition, in: TENCON 2018-2018 IEEE Region 10 Conference, IEEE, 2018, pp. 1414–1419.

[12] R. Sujatha, J.M. Chatterjee, N. Jhanjhi, S.N. Brohi, Performance of deep learning vs machine learning in plant leaf disease detection, Microprocess. Microsyst. 80 (2021), 103615.

[13] Y. Zhong, M. Zhao, Research on deep learning in apple leaf disease recognition, Comput. Electron. Agric. 168 (2020), 105146.