Advancing Apple Leaf Health: A CNN Approach to Leaf Disease Detection

[1]Dr. Prakasam S, [2] Mr. Mohan T S and [3]Dr. Shanmugapriya P

[1]Associate Professor, Department of Computer Science and Applications,
Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, Tamil Nadu, 631561 – India.

[2] Research Scholar, Department of Computer Science and Applications,
Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, Tamil Nadu, 631561 – India.

[3]Associate Professor, Department of Computer Science and Engineering,
Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, Tamil Nadu, 631561 – India.

Abstract

An identification of agricultural image establishment requires the fundamentals of contemporary farming practices. Convolutional Neural Network (CNN) plays a crucial role in developing a high-performance model for leaf disease detection. It is an advanced machine learning technique most extensively utilized and image classification and recognition. One of the most widely grown plants across worldwide is the apple tree. It is quite detrimental, resulting in lower quality and productivity. Effective control and reduction of the use of chemical fertilizers depends on early identification of these diseases. The implementation of a CNN model for identifying apple leaf diseases was investigated in this research article. The accuracy is attained with image datasets will be split into different classes. The CNN model became more efficient, which enabled high performance even for a limited dataset. This model's great accuracy, recall, and F1-scores across all classes, and the confusion matrix is calculated. This work highlights CNN's promise as a consistent tool for early and accurate apple leaf disease diagnosis, hence improving apple management and environmentally friendly farming methods.

Keywords— Artificial Intelligence, Deep Learning, Apple Leaf Disease Detection, Model Training, CNN Model.

1. INTRODUCTION

The Indian economic survey for indicates that agricultural growth in India has been around 6.9% annually for the past few years. Among the most farmed and eaten fruits across worldwide, apples are the main fruit that offers vital minerals. But apple plants are quite vulnerable for too many illnesses that mostly harm their leaves, which reduces the productivity of the apple and its quality greatly. Considering the number of people in India still depending on agriculture, it is a real concern. It can reshape various general strategies used in agricultural systems [1]. Plant disease detection identifiers highly depend on focusing on the labour demand and are impractical. So, this system reduces the problem of manual identification of detecting diseases in agriculture. This framework uses the Deep Learning method, such as CNN, to enhance the identification of diseases to apply pesticides. This research finds out the accuracy of each class of apple leaf, and it can be compared to pre-trained models for efficient leaf disease detection: a study on custom CNN [2]. In this research paper, the CNN model for the identification of several diseases is explored and enabling boosted accuracy with apple leaf datasets.

2. LITERATURE REVIEW

With an accuracy of 93%, the model uses growing filter counts (32 → 64 → 128) across convolutional layers. This is the best design practice because layers learn simple features (edges, textures) and deeper layers learn complex features (disease spots, colour variations, shape patterns). The proposed CNN model can efficiently extract information and precisely classify many apple leaf diseases. On the dataset of apple leaf diseases, the experimental results of the thorough comparison show that this developed CNN model attains a high accuracy of 93% [4]. An experiment is conducted for five classes: Alternaria leaf spot, Brown spot, gray spot, Healthy leaf, and Rust of the plant village public dataset. The highest accuracy of 93% obtained on layers of the CNN model [5]. This research focused on using Image Processing systems to totally eliminate the negative side effects of fertilizers and chemicals. Five categorizations of apple leaf disease have been found in this study, including one healthy leaf class [6]. Healthy Leaf's average accuracy is 91%. Tomato Guard proved to be chosen as a model for spotting tomato diseases, demonstrating notably superior performance than APP Plantix, a commonly utilized tool for general-purpose plant disease detection [7]. With an average precision and overall good accuracy, the produced IOU score in detection is 93%. Therefore, the work that is shown beats any other stated state-of-the-art [8]. This study uses image processing techniques based on Image division, algorithms, and clustering for apple plant leaf disease, contributing to a reliable, safe, and accurate system of leaf disease with speciality to plants [9]. The CNN's backbone was Inception V3 using the Plant Village dataset, and the best accuracy for tomato leaf diseases is attained [10]. Furthermore, showing more accurate findings than other ensemble models is a trio of models (VGG16, InceptionV3, and Google Net [11].

3. DATASET

The dataset is collected from the open-source platform, such as https://china.scidb.cn/carefully taken to help to identify apple leaf diseases. The set of data used in this paper came from a large collection that was thoroughly split into two sets. So, it is possible to offer a thorough evaluation of the CNN model to build for this work. Ensuring that the model could effectively learn the special characteristics of every disease required the balanced distribution of images across all classes, hence boosting its power to precisely classify and diagnose apple leaf illnesses. Depending on the use of this well-organized and labeled dataset, reliable and practical results follow from which early detection helps to more effectively control apple leaf diseases in agricultural operations.

(a). Alternaria leaf spot

(b). Brown spot

(c). Gray spot

(d). Healthy leaf

(e). Rust

Fig:1. Dataset images for 5different Classes

4. PROPOSED MODEL

Created for image classification, the model described in CNN was built with the Keras Sequential API. It starts with an input layer processing three-channel pictures of shape (image_height, image_width). Following a ReLU activation function, the first convolutional layer down-samples by different layers, such as a Max Pooling layer after 16 3x3 filters. After 32 and 64 filters in the next Max Pooling comes after convolutional layers for dimensionality reduction and is repeated. Following the convolutional blocks, average Pooling layers skip the feature maps' spatial dimensions so transforming them into a single vector. Then this vector passes a Dense layer with 128 ReLU-activated neurons. Then the output layer uses SoftMax activation to display the probability distribution among different classes.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
max_pooling2d (MaxPooling2D)	(None, 63, 63, 32)	0
conv2d_1 (Conv2D)	(None, 61, 61, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 30, 30, 64)	0
conv2d_2 (Conv2D)	(None, 28, 28, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 128)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 128)	3,211,392
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 5)	645

Fig. 2 Proposed CNN Architecture

We evaluate many performances ensures to find that the model's ability to properly categorize the diseases: accuracy, precision, recall, F1-score, and the confusion matrix. Testing the model on images from many sources or with variable quality helps to evaluate its generalization and robustness. This phase guarantees that the model is not only accurate but also dependable in practical uses, therefore offering important information on possible implementation in agricultural environments.

5. METHODOLOGY

PHASE 1. DATASET COLLECTION

This is an initial phase of the project, and it is producing a large collection of Apple leaf pictures. These images are gathered from https://china.scidb.cn/. These sources are well-observed collections to ensure a varied and representative sampling of the various apple leaf diseases under research. All images are tagged depending on the specific disease. Correct training and evaluation of the CNN model relies on a balanced number of images across all categories; Hence, the dataset is carefully selected to include such balance.

PHASE 2. DATSET PRETREATMENT

Before the images are loaded into the CNN model, the gathered dataset is pre-processed in the next phase to improve its consistency and quality. Usually demanded by the CNN model, this stage consists of scaling images to a consistent dimension and using several image augmentation approaches, including rotation, flipping, and zooming. These augmentations increase the variance of the training data, allowing the model to extend to fresh, unprocessed images. Normalization of the pixel values helps to standardize the input data, therefore hurrying model convergence during training.

PHASE 3. MODEL TRAINING AND VALIDATION

The third phase is largely concerned with CNN model validation and training. By use of a CNN model is enhanced using the processed dataset to fit the particular purpose of tomato leaf disease detection. Training and validation sets comprise the dataset; the training set helps to stop overfitting by tracking the model performance and updating the weights of the CNN model. Perfectly set to exact performance, the hyperparameters that are learning rate, batch size, and number of epochs at this point, the algorithm takes up the complex trends connected to many apple leaf diseases such as Alternaria leaf spot, Brown spot, gray spot, Healthy leaf, and Rust.

PHASE 4. MODEL EVALUATION AND PERFORMANCE MATRIX

The evaluation of the model that is trained is utilizing a separate test set not engaged in the training process forms the last phase. We compute several performance measures to figure out the model's ability to appropriately categorize the diseases: accuracy, precision, recall, F1-score, and the confusion matrix. Testing the model on images from many sources or with variable quality helps one also evaluate its generalization and robustness. This phase guarantees that the model is not only accurate but also dependable in practical uses, therefore offering important information on possible implementation in agricultural environments.

6. RESULTS

Identifying Apple leaf diseases, the CNN model performed well with 93% of general accuracy. The confusion matrix was confirming over all 5 classes of its accuracy, recall, and F1-scores. Even in the confusion matrix calculation between the Rust and Gray spot (13.1%) suggests visually similar into their leaf texture and color patterns, which can challenge CNN-based feature extraction. Hence, it is an optimal tool for accurate disease recognition in apple leaf plants.

A. Classification Report Analysis

The classification report determines the plant disease diagnosis capability of a CNN model. Designed especially for image analysis uses, the Neural Network (NN) model is a strong deep learning architecture that is suitable for this purpose as well. With an amazing total accuracy of 93% the model says that it can highly precisely classify plant illnesses. For most classes, precision, recall, and F1-score measures are often above 90%, therefore proving the model's ability to identify and predict many disease types.

Classification Report:

	precision	recall	f1-score	support
Alternaria leaf spot	0.91	0.99	0.95	376
Brown spot	0.97	0.86	0.91	298
Gray spot	0.95	0.91	0.93	563
Healthy leaf	0.88	0.98	0.92	553
Rust	0.97	0.88	0.92	487
accuracy			0.93	2277
macro avg	0.93	0.92	0.93	2277
weighted avg	0.93	0.93	0.93	2277

Fig. 3: ***Classification Report***

Also, this model shows excellent general performance; accuracy varies somewhat among particular courses. This tells that extra training or feature techniques might help the model to improve on specific disease types. Still, given constant and accurate calculations for plant disease detection, the CNN model shows potential as a practical tool.

B. Training and Validation loss Report

During the period prediction, the loss plot indicates a decline in both phases of loss. Especially in the 15 epochs, loss shows a sharp drop denotes fast model learning. In this phase, the loss keeps declining more slowly until the epochs reach low values. Since this validation loss does not rise significantly at later stages, the closeness between the two curves given suggests that the model is well-regularized and shows a minimum of overfitting.

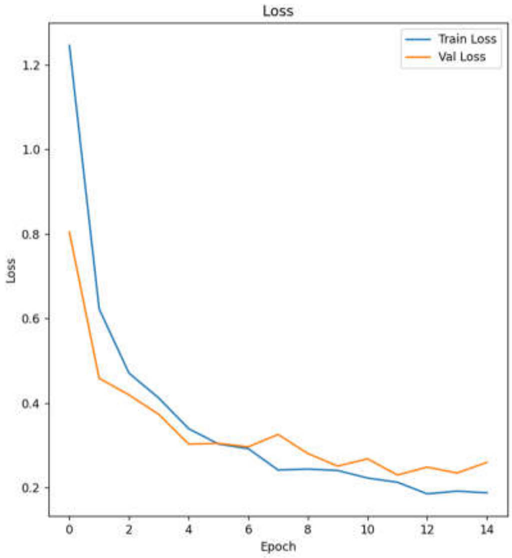


Fig. 4 Training and Validation loss Analysis

C. Analysis of Training and Validation Accuracy

Due to the rise of epochs, the accuracy plot shows a constant improvement in the two given accuracies. The training accuracy rises quickly at first and then starts to settle and often stays around 85% beyond 2 epochs, about 15 epochs. With little variation, the validation accuracy trends similarly to the training accuracy. Given both two accuracies, it indicates that the model is learning efficiently without appreciable overfitting.

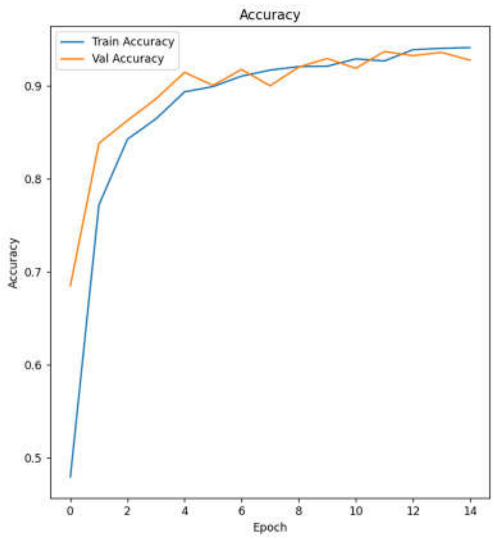
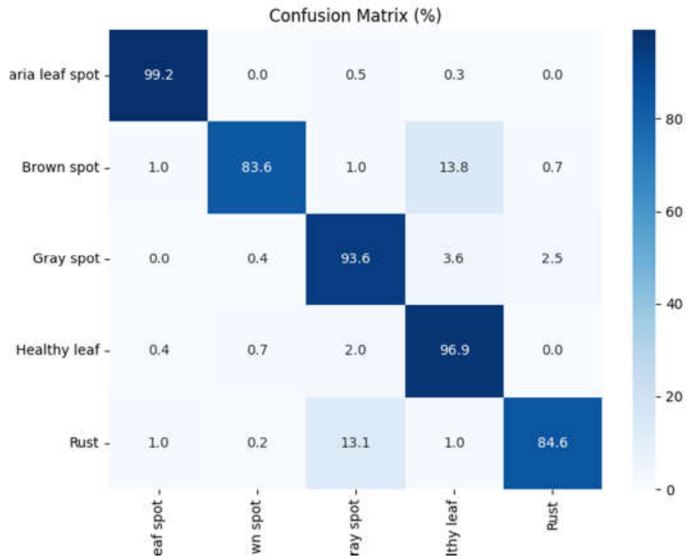


Fig: Training and Validation Accuracy

D. Analysis of Confusion Matrix

The confusion matrix indicates how the CNN model correctly separates several apple leaf diseases. The quite high diagonal values, which return correctly classified occurrences, show good model performance over 5 classes. For instance, healthy leaf was correctly identified 99.2% and Alternaria leaf spot was identified 96.9% time, each with relatively few misclassifications. Reaching 100% accuracy in both groups, this CNN model is well-suited for spotting diseases in Healthy Leaves and Gray Spot. These misclassifications, the general accuracy, recall, and F1-scores are always quite good, so verifying the dependability of the model. For the effective agricultural business, the confusion matrix confirms the CNN model's durability and has an accuracy of 93%. This makes early and reliable apple leaf disease detection possible.



7. CONCLUSION

Early prediction and categorization of apple leaf diseases is a key element of renewable agriculture and requires on a well-tuned Convolutional Neural Network (CNN), which works efficiently. The overall accuracy of 93% was achieved over 5 different classes that include using transfer learning on a pre-trained CNN model. With continuously high accuracy, recall, and F1-scores, the model reinforced its reliability in checking and identifying several apple leaf irregularities with fewer mistakes. Finally, among all diseases like Brown spot and Rust, the confusion matrix analysis revealed even more the resilience of the technique since most classes are accurately identified and very few are minor misclassifications. These predictions confirm if the CNN model can differentiate between minor variations in leaf disease symptoms, therefore providing farmers and other workers with a useful tool. Efficient implementation of this CNN-based approach underlines the changing ability of deep learning technology in modern agriculture. This approach can reduce the need for chemical fertilizers use by enabling disease identification encouraging in advance with sustainable agriculture practices. Early disease diagnosis also minimizes crop losses and maximizes the general output and quality. This work increases the knowledge on the application of artificial intelligence in agriculture by offering a valuable solution to one of the most common issues in apple production. The positive results of this research allow for more research and findings of deep learning models for the classification of plant diseases. Hence, maybe this work can be extended for region-wise leaf classification in agricultural contexts.

8. REFERENCES

[1]. Sandeep Kumar, Rajeev Ratan, J. V. Desai, "Cotton Disease Detection Using TensorFlow Machine Learning Technique", Advances in Multimedia, vol. 2022, Article ID 1812025, 10 pages, 2022.
[2]. Touhidul Seyam Alam, Chandni Barua Jowthi, and Abhijit Pathak, “Comparing pre-trained models for efficient leaf disease detection: a study on custom CNN”, Alam et al. Journal of Electrical Systems and Info Techno (2024) 11:12 Page 2 of 26
[3]. Samal, S. and Verma, V., 2024. Tomato plant leaf diseases detection and classification using an advanced self-developed CNN model. Multimedia Tools and Applications, pp.1-30.
[4].Waleed, J., Albawi, S., Flayyih, H.Q., and Alkhayyat, A., 2021, September. An effective and accurate CNN model for detecting tomato leaf diseases. In 2021 4th International Iraqi Conference on Engineering Technology and Its Applications (IICETA) (pp. 33-37). IEEE.
[5]. Samal, S. and Verma, V., 2024. Tomato plant leaf diseases detection and classification using an advanced self-developed CNN model. Multimedia Tools and Applications, pp.1-30.
[6]. Mim, T.T., Sheikh, M.H., Shampa, R.A., Reza, M.S., and Islam, M.S., 2019, November. Leaf disease detection of tomato using image processing. In 2019 8th International Conference on System Modelling and Advancement in Research Trends (SMART) (pp. 244-249). IEEE.
[7]. Tian, K., Zeng, J., Song, T., Li, Z., Evans, A., and Li, J., 2023. Tomato leaf disease recognition based on deep convolutional neural networks. Journal of Agricultural Engineering, 54(1).
[8]. Roy, K., Chaudhuri, S.S., Frnda, J., Bandopadhyay, S., Ray, I.J., Banerjee, S., and Nedoma, J., 2023. Detection of tomato leaf diseases for agro-based industries using novel PCA DeepNet. IEEE Access, 11, pp.14983-15001.
[9]. Bhoomika1, Goldy Verma “Advancing Tomato Crop Health: A CNN Approach to Leaf Disease Detection”, Proceedings of the 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA 2024), IEEE Xplore.
[10]. Alhaj Ali, A., Chramcov, B., Jasek, R., Katta, R., Krayem, S., and Awwama, E., 2021. Tomato leaf disease detection using deep learning. In Proceedings of the Computational Methods in Systems and Software (pp.199-208). Cham: Springer International Publishing.
[11]. KAUR, P., SINGH, M.P., MISHRA, A.M., SHANKAR, A., SINGH, P., DIWAKAR, M., and NAYAK, S.R., 2023. DELM: deep ensemble learning model for multiclass classification of super-resolution leaf disease images. Turkish Journal of Agriculture and Forestry, 47(5), pp.727-745.