International Journal of Sustainable Studies, Technologies, and Assessments

IJSSTA

Original Article

Prediction and Analysis of Apple Diseases using a Convolutional Neural Network

Deepak Chandra Uprety
Noida Institute of Engineering and Technology (NIET), AKTU University, Greater Noida, UP- 201306, India
deepak.glb@gmail.com
Nitish Kumar
LPU University, Phagwara, Punjab 144001, India
nitishchauhan946@gmail.com,
Dyuti Banarjee,
Koneru Lakshmaiah Education Foundation, Andhra Pradesh-522302, India
dyuti738900@gmail.com,
Abhimanyu Dhiman,
IEC University, Baddi, Himachal Pradesh, India
abhimanyu.dhiman1995@gmail.com

Abstract: Information technology has created important advancements in a range of fields, including education, transportation, fitness, and agriculture. IT utilization has produced outcomes that are affordable, efficient, and successful. IT has been incorporated into agricultural research to address the issues that ranchers, employees, yield production, and the supply chain confront. India, a predominantly agricultural nation, mainly depends on crop production efficiency to power its economy. Almost every industry in the nation is based on agriculture, and for farmers, the main problem is growing good crops. Crop infection detection is essential for avoiding crop damage and crop failure. The detection of diseases like rust, scabs, and others in apples is the main topic of this essay. The study's dataset consists of about 3500 photos of apple leaves that were gathered from Kaggle. The dataset is used to categorize healthy and diseased apple leaves and identify apple diseases using the convolutional neural network. In test photos, the suggested model, which employs a convolutional neural network-based methodology, displays a high accuracy of over 92% for various categories of apple leaves.

Keywords: Convolutional Neural Network, Healthy Leaves, Multiple Diseases, Rust, Scab, Analysis and Results.

1. Introduction

The global economy depends heavily on the agricultural sector. Plant disease misdiagnosis can result in enormous losses for farmers, harm to crops, and harm to the people who consume the goods. The early detection of leaf diseases is one agricultural concern that artificial intelligence (AI) is a useful tool for solving. By using automatic plant leaf diagnostic techniques, farmers may decrease losses and increase output. To investigate plant diseases, researchers have been experimenting with various machine learning, deep learning, and image-processing algorithms.

Our research focuses on the application of an AI subset called deep learning to identify illnesses in apple leaves. To extract characteristics from photos using CNN, such as horizontal edges, straight edges, and green, dark, and red values, we used a deep learning (DL) model. A crucial component of data analysis is machine learning, which enables the model to automatically analyze data, forecast arrangements, and fix outcomes without human intervention. With the help of new technologies, machine learning may apply the skills it has learned and the training it has received to forecast patterns.

For the model to accurately forecast fresh data, researchers must build the model training method. The model's predictions based on actual training can aid in the prediction of fresh data or objects added to the model. In-depth literature, a branch of machine learning, employs a neural web with numerous layers that function as a brain and process a significant portion of the information. All particles in the neurological network function as nerve cells and estimate the information contained therein, whether it be content information or appearance information, and produce predictions. This strategy is not brand new, but it is an important scientific innovation that can help advance agricultural methods.

A. Convolutional Neurological Network

1) The ConvNet is built up so that several slabs are settled, one from the input flake to the output flake. This network, also known as a multi-layer neural net,
2) The type of program that might have been read in it. ConvNet will analyze hierarchic mood using the sequence architecture that you allow.
3) In ConvNet, collection flakes can be pursued by dark flakes, and the same by activating flakes
4) Pre-clarification refers to the analysis that is necessary before ConvNet makes its real prophecy.

B. The slabs in the convolution network are

1) Curlicue Flake.
2) High Integration Flake.
3) An A-Flat Flake.
4) The dense flake.
5) Exit Flake.

2. Literature Survey

In [1], a mathematical model for plant diagnosis and recognition based on in-depth learning has been put forth by Yan Guo et al. to increase precision, consistency, and training effectiveness. A range project network (RPN) is used by the model to recognize and locate leaves in intricate petal structures. The findings of the RPN method are then used to separate the pictures, which are then input into a computer vision (CV) algorithm for feature extraction. Based on a database of infected leaves in a straightforward setting, the model is then trained with other leaves using a transfer learning approach. The model was tested for black breakdown, bacterial slab, and degradation, and it outperformed the conventional approaches with an accuracy of 83.57%. The suggested approach is consistent with ongoing agricultural progress and lessens the negative effects of plant diseases on agricultural production. Deep learning is therefore crucial for environmentally friendly agriculture, agricultural output, and smart agriculture.

In [2], a new approach based on deep transfer learning for grape leaf acquisition has been introduced by Prabhjot Kaur et al. The model employs a variance strategy to eliminate extraneous features from the component extractor course after extracting features using FC layers. Then, with 98.7% accuracy, the following features are designated using a regulation algorithm: According to the statement, relocation training models are more advanced than those that are created from scratch.

In [3], Apple trees are prone to several diseases, according to Amit Gawade et al., including rust, Alternaria leaf spot, brown spot, mosaic, and leaf fall. Due to the excessive humidity in the air during the monsoon season, these diseases are more prevalent in mountainous places. A proposed model has been created to correctly and successfully detect various illnesses in apple leaves. The confusion matrix, accuracy, exactness, memory, and F1 records were used to assess the exemplary performance, and the results revealed a high accuracy of 90%. The exemplary achieves an accuracy of 00.95, a recollection of 00.94, and an F1 record of 00.96 using ResNet50 and Plant Pathology Lab 2020 data.

According to [4] Bin Liu et al., apple trees are prone to infections including Decay, Brown blemish, Mosaic, and Alternaria petal blemish. To stop the spread of these viruses and keep the apple sector healthy, early notification of these infections is essential. Pre-image processing research from the past, however, has limits in its ability to detect apple leaf illnesses with accuracy. In this paper, a novel method for identifying these disorders using deep convolutional neural networks is proposed. To do this, a complex neural convolutional network architecture based on AlexNet was trained using a large dataset of 13,689 photos of damaged apple leaves. Under the dominant set of tests, the anticipated method diagnosed the four most prevalent apple leaf illnesses with a total accuracy of 97.62%.

The authors of [5] D. Jayanarayana Reddy et al. examines and analyze several elements that greatly depend on data accessibility. Although numerous studies have employed common algorithms like CNN, LSTM, and DNN, further advancements for CYP are still required. In order to improve existing models, recent research has taken into account variables including temperature, climatic conditions, and active yield forecast models. Experimental investigations show that improving crop prediction can be facilitated by integrating ML with an agricultural background. However, more work needs to be done to diversify the features available in terms of how they affect agriculture.

[6] By Gianni Fenu et al. In this study, predictors of plant diseases and plant infection are examined and allocated throughout the previous ten years, from 2010 to 2020. The study evaluated approved procedures and pre-processing techniques, examined 46 research activities, and identified the data used while emphasizing the difficulties faced. Due to the interaction of numerous environmental elements and climates in actual conditions, the research shows that predicting plant and plant disease outbreaks is a challenging challenge. It is therefore necessary to put more effort into gathering information and coming up with creative justifications to inhibit and lessen the influence of plant and plant infection shock on meal manufacture, especially for staple foods in less developed countries. Despite the advancements, the field is still young, and there are still many challenges to be solved.

[7] Prakanshu Srivastava and others argue that convolutional networks have a well-established track record of successfully extracting characteristics from huge databases. However, it's crucial to evaluate how well these networks perform using actual photos. The network settings were changed to reach an overall accuracy of 88%. Utilizing photos from the verification collection, the trained model was additionally evaluated for each individual class. According to the principles of good practice, the outcomes were compared to the ones that were obtained. This study's objective was to spot plant illnesses in pictures of leaves. The complete procedure—image collection, pre-processing, deep CNN training, and fine-tuning was explained.

In [8] Prakhar Bansal and colleagues, using a database of photos of apple leaves, a combination of EfficientNetB7, EfficientNet NoisyStudent, and DenseNet121 was suggested and assessed in this study. The database included 3642 pictures of apple leaves that were both healthy and had apple scabs, cedar rust, and other illnesses. A weakness of this study is that it only examined four classes connected to two foliar diseases. When compared to other high-end models trained on the same database, the performance of the proposed models was shown to be superior. An accuracy of 96.25% in the validation was attained using the average forecast of the three models. Farmers may be able to automate the detection of illness in apple trees with the aid of this level of accuracy.

3. Proposed Methodology

A. Comparative Analysis of Existing Solutions

1) The majority of the time, researchers employ machine learning to forecast the type of crop illnesses while taking into account various parameters like crop leaf pictures, etc.
2) The top studies had models built from scratch rather than using an existing tool, such as image processing techniques or algorithms like CNN, etc.
3) To increase the effectiveness of various models, it is important to note that ensemble learning and models built on DL architecture provide the greatest results for prediction.
4) The most frequently utilized architecture and algorithm in these problems is CNN.

Our study's main objective is to use instrument learning and deep learning techniques, such as ConvNet and instrument learning allocation algorithms, to expertly predict apple diseases.

B. The development of a machine learning model for predicting diseases

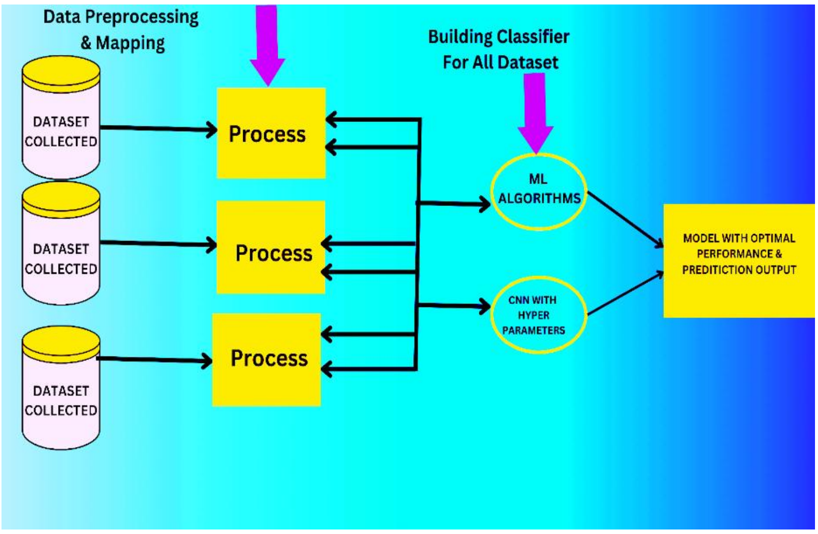


Fig.1 Machine learning model workflow for predicting diseases

4. Apple Leaves Database of plant Pathology

Healthy Leaf

Fig.2 A typical apple leaf

Multiple diseases

Fig.3 Multiple illness leave

Rust

Fig.4 A sick apple leaf with rust

Scab

Fig.5 Leaves with scab disease.

5. Steps Followed in the Proposed Solution

A. Various Steps

Step 1: Data storage and examination

To ensure consistency, most attributes in distant datasets should be the same and belong to the same category. These attributes include types of apple diseases such as rust, scab, multiple diseases, and healthy leaves. The fundamental step in this course is to assemble information from various accessible information platforms, such as Kaggle or data.gov.

Step 2: Data preparation

The data is fragmented into the training data file and the testing data file after the features have been determined. The test data file is used to evaluate the model's performance when applied to unknowable data, whereas the training data file is recycled to train the model. For deep learning architecture, it is advised to use a supplemental set called the validation set.

Step 3: Building Individualized Allocation Models

Using instrumentation literature and a deep literature method, we must frame frequent allocation models for every data file after creating the information sets.

Step 4: Assessing the Performance of the Model

We must fit them into a training set we previously prepared after developing the classification model. The testing set is used to forecast how well the algorithm would perform there. The distraction matrix and inaccurate categorization model predictions are used to measure these achievement metrics.

The same approach can be used to predict the usage of test information, also known as invisible information, once it has gained expertise. Following the completion of this, we may develop a distraction matrix that will explain how skilled our model is. To create a more accurate model, we aim to obtain more codes from true positives and true negatives. The number of divisions determines the size of the distraction matrix in its entirety.

1) True positives: In this case, both the prediction and the expected result are 1.
2) True negatives: Our expected output and the TN prediction are both 0.
3) False positives: The prediction is 1, but the output is 0.

4) False negatives: Although the prediction's value is 0, the outcome is 1.

B. Utilization of Various Performance Metrics

The distraction matrix can also be used to determine the model's degree of certainty.

1) Accurateness = (TP +TN) / (TP+TN+FP+FN)
2) Correctness = TP / (TP + FP)

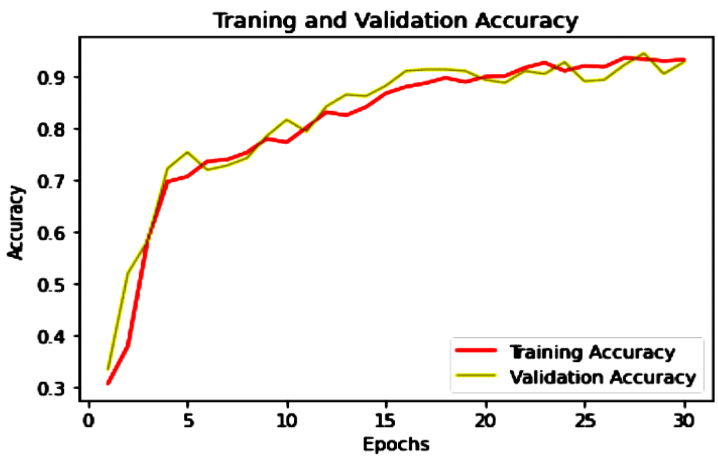
6. Analysis and Results of Experiment

This dataset, which is made available in CSV format by Kaggle, includes several sets of photographs of apples as well as images of apple diseases like scab, rot, and many infections as well as vibrant apple leaves. The data file contains over 3500 photos of apple leaves with various diseases that are common to apples.

Detection of several illnesses in Apple plants using a visual representation of their leaves. Data science plots are created using the Matplot library.

A. Analysis of Validation and Training Accuracy

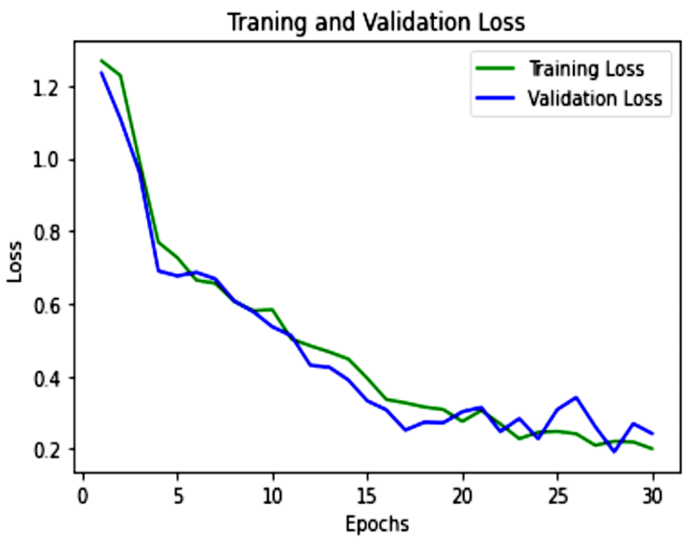
The term "testing accuracy" typically refers to the accuracy of verification or the computation accuracy of a data set that is used throughout training to validate or "test" the replica's ability to be put together or "early stopping" rather than being used for training.



Graph 1 Apple dataset training and validation accuracy broken down by epochs

B. Analyses of Training and Validation Losses

Losses are determined by the exemplar's performance in the training and acceptance data sets, and the definition of losses depends on those results. In the acceptance or training data sets, the total aberration has been identified for each case. After each round of modifications, the exemplar's performance is measured by its disaster rate.



Graph 2 shows the training and validation loss for the dataset for apples by epochs.

C. Algorithm Analysis

In this section, we'll compare and contrast how each categorization performs on various datasets. The performance criteria considered include guidance loss, legalization loss, legalization accuracy, and training accuracy. Each dataset was divided into three parts:

Table 1 shows how the training and test data sets were divided

Utilized Approach	Practice Set	Set Acceptance	Examining Set
Machine Learning	0.75	NA	0.25

We must also point out that the estimated 3500 apple photos, both healthy and diseased, were used as samples for our study.

Before training the data set, we discovered that the following hyperparameter combination is the best one:

1) One Epoch: thirty
2) Adam is an optimizer
3) The batch size is 14.

D. Testing Set Results (Testing and Validation)

Table 2 Results on a test dataset for all photos

Algorithm	Accuracy (%)	Loss (%)	Validation Exactness (%)	Loss of Validation (%)
CNN Model for deep learning	92%	0.22	92%	0.24

.

7. Conclusion

This study has looked at how AI can be used in a variety of domains, including the prediction of plant and crop diseases. Numerous studies contend that DL architecture is more effective for these issues, despite the fact that CNN has been deemed effective in disease prediction. Deep learning is made for issues where a margin on CNN may be required. We also claim that this is the case in our research and application work, where the ConvNet model achieves an accuracy of roughly 98% over the provided dataset. As a result, when working with several complex pieces of information, DL architecture looks to be better than conventional machine learning techniques.

While this is going on, AI tool creators should think of ways to enhance their creations and aid farmers in overcoming obstacles. Additionally, they ought to figure out how to clearly explain how AI can help with practical issues like eliminating manual labor. AI in agriculture has a bright future and is anticipated to produce beneficial outcomes.

Acknowledgement

It would not have been feasible to complete this research project, Title: "Prediction and Analysis of Apple Diseases using a Convolutional Neural Network," without the involvement and support of my teammates. We truly appreciate and heartily recognize your contribution. Nonetheless, the team would like to sincerely thank the following:

Er. Nitish Kumar, Dr. Dyuti Banarjee, Er. Abhimanyu Dhiman, Dr. Mohammad Shahid, and Dr. Ahmad Neyaz Khan for their endless support and understanding of humanity in completing this research work

References

[1] Guo, Y., Zhang, J., Yin, C., Hu, X., Zou, Y., Xue, Z., & Wang, W. (2020). Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society, 2020, 1-11.

[2] Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A., & Alabdali, A. M. (2022). Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction. Sensors, 22(2), 575.

[3] Gawade, A., Deolekar, S., & Patil, V. (2021). Early-stage apple leaf disease prediction using deep learning. Bioscience Biotechnology Research Communications, 14(5), 40-43.

[4] Liu, B., Zhang, Y., He, D., & Li, Y. (2017). Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 10(1), 11.

[5] Adit, S. B. M., Chowdhury, I. A., Alam, W., & Salam, S. M. (2022, November). A Predictive Analysis on Weather Parameters Based on Bangladesh Agriculture Using Machine Learning Approach. In 2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon) (pp. 1-5). IEEE.

[6] Fenu, G., & Malloci, F. M. (2021). Forecasting plant and crop disease: an explorative study on current algorithms. Big Data and Cognitive Computing, 5(1), 2.

[7] Zhong, Y., & Zhao, M. (2020). Research on deep learning in apple leaf disease recognition. Computers and Electronics in Agriculture, 168, 105146.

[8] Akshai, K. P., & Anitha, J. (2021, May). Plant disease classification using deep learning. In 2021 3rd International Conference on Signal Processing and Communication (ICPSC) (pp. 407-411).

[9] Jan, M., & Ahmad, H. (2020). Image features based intelligent apple disease prediction system: Machine learning based apple disease prediction system. International Journal of Agricultural and Environmental Information Systems (IJAEIS), 11(3), 31-47.

[10] Bansal, P., Kumar, R., & Kumar, S. (2021). Disease detection in apple leaves using deep convolutional neural network. Agriculture, 11(7), 617.

[11] Fenu, G., & Malloci, F. M. (2021). Forecasting plant and crop disease: an explorative study on current algorithms. Big Data and Cognitive Computing, 5(1), 2.

[12] Islam, M. T. (2020). Plant disease detection using CNN model and image processing. Int. J. Eng. Res. Technol, 9(10), 291-297.