2022 3rd International Conference for Emerging Technology (INCET)
Belgaum, India. May 27-29, 2022

Plant Disease Detection Using Machine Learning Techniques

Divyanshu Varshney
Department of Computer Application
NIT Kurukshetra,
Haryana
opgmdivyanshu@gmail.com

Burhanuddin Babukhanwala
Department of Computer Application
NIT Kurukshetra,
Haryana
burhanbabukhan98@gmail.com

Javed Khan
Department of Computer Application
NIT Kurukshetra,
Haryana
jkhanjma144@gmail.com

Deepika Saxena
Department of Computer Application
NIT Kurukshetra,
Haryana
13deepikasaxena@gmail.com

Ashutosh kumar singh
Department of Computer Application
NIT Kurukshetra,
Haryana
ashutosh@nitkkr.ac.in

2022 3rd International Conference for Emerging Technology (INCET) | 978-1-6654-9499-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/INCET54531.2022.9824653

***Abstract* - Plant diseases may have a major impact on food safety, also a considerable decline in agricultural product output. The great majority of automated systems developed thus far are based on digital pictures, allowing for the rapid deployment of algorithms. The difficulty of autonomous illness identification in plants has been solved using traditional machine learning approaches such as (SVM) support vector machines, Multilayer Perception Neural Networks, and Decision Trees. The focus of this article was on leaf plant disease. A new plant leaf disease detection technique has been developed that is based on a transfer learning methodology such as deep learning, where CNN is employed as a feature extractor and SVM is used for classification. A benchmark dataset called PlantVillage was used to assess the evaluation of the proposed model. The suggested model was examined and compared to current methodologies, and it outperformed previous work, achieving an 88.77 percent training accuracy.**

Keywords - Transfer learning, Machine learning, CNN, Data mining, Image processing, Plant disease detection

I. INTRODUCTION

Plant disease detection is crucial in agriculture because farmers must constantly assess if the produce they're harvesting is of acceptable quality.

It is critical to treat this seriously because it can cause major difficulties in plants, affecting product quality, quantity, or productivity. Plant diseases have had an influence on society and history around the world. Several countries such as the United States affected with great economic losses due to plant diseases. It can affect any type of plant. Because Plant leaves are the most sensitive and display disease symptoms earliest. From the beginning of their life cycle until They're ready to be picked, the crops must be monitored for illnesses. Plant diseases generate disease outbreaks regularly, resulting in large-scale death and a significant economic impact.

Image-based automatic inspection can be provided using computer vision techniques. Manual identification is time-consuming, inaccurate, and limited to small areas at a time. Plant diseases may be discovered early using this technique, and pest and infection management strategies can be employed to treat pest issues while minimizing dangers to humans and the environment. An automated system that can identify plant illnesses based on the look and visual symptoms of the plant could be extremely useful. This system can be used in agricultural fields to automate the entire pipeline. This would not only enhance efficiency by allowing machines to perform these redundant duties better than humans, but it would also increase farm yield. We use deep learning and computer recognition techniques to solve the problem of autonomous bacterial blight classification.

One of the most significant tasks in agricultural practices is the early diagnosis and classification of plant diseases. Every year, illness infection causes a significant economic loss to farmers. As a result, early, accurate, and prompt detection of the condition avoids product loss while also improving product quality (Asma Akhtar et al., 2013). As a result, it contributes to the country's economic growth. Traditionally, these disorders were diagnosed based on either pathogen-caused ocular symptoms or pathogen detection in the laboratory (Sharda P. Mohanty, 2016). The visual evaluation of the illness lesion is a subjective process that may fail to diagnose the disease correctly. Pathogen identification in the laboratory, on the other hand, becomes a time-consuming process because of pathogen culture, which may fail to deliver results promptly. Furthermore, each of the aforementioned techniques necessitates the use of specialists as well as paper; we have taken into consideration photographs of various common plants such as the Tomato plant, Potato leaves, and Pepper.

A variety of strategies have been tried to produce semiautonomous plant disease detection devices in latest days. These methods have so far shown to be quicker, lighter, and more exact than growers' hand observation, which is the traditional method of observation. As a result, researchers are being urged to develop more intelligent technical systems for detecting plant diseases that do not necessitate the involvement of a human. The purpose of this work is to evaluate and discuss several plant disease detection methods in terms of various factors. For the sake of global health and well- being, it is important to obtain a correct diagnosis of plant diseases. In this ever- changing world, early detection of disease, as well as early prevention, is critical to avoiding issues that may otherwise arise. Any of these issues, such as global food scarcity, may have catastrophic consequences for humanity. It is important to avoid squandering financial capital in the pursuit of a healthy lifestyle, by looking at climate change from an ecological standpoint. A person can't detect all types of plant disease problems with their naked eye. Repeating the procedure is often time-consuming and inefficient. To identify plant diseases with accuracy, a plant pathologist must have outstanding observation skills to understand unique symptoms. The foundation of effective plant disease prevention and control is early warning and forecasting. They are extremely important in agricultural production management and decision-making. Plant diseases are a global harmful to food security, but they can be especially damaging to small-scale farmers who depend on healthy crops for their livelihood. Smallholder farmers account for more than 80% of agricultural production in developing countries.

A. *Our contribution*

Existing approaches have used the classical machine learning models like Support Vector Machine (SVM), Decision Tree, etc. which are not long-lasting because they are static models. On the other hand, we have used transfer learning as a better approach as it advances learning. In this paper, a model has been developed on transfer learning and getting better accuracy than the existing model.

II. RELATED WORK

After reviewing a lot of research papers related to our work, we have decided to add some quality papers and recent papers to our paper. Various studies in this field have been conducted utilizing various datasets, feature reduction approaches, and classification algorithms.

In 2018, S. Sannakkiet et al. proposed the faulty region is the feature, while color and texture are the features, in a categorization method based on the Segment system. The main advantage is that it makes to L*a*b to eliminate picture chromaticity layers and learns 97.30% of classification. The biggest disadvantage is that it can only be utilized for specific crops. To solve diverse classification issues, the BPNN classifier employs an active contour model to limit the vitality inside the infected area. Computer vision technologies and fuzzy logic are used to identify and assess leaf diseases. To assist evaluate the severity of the sick leaf, they used an artificial neural network (ANN) as a classifier. Pre-processing, attribute extraction, classifier training, and classification are all steps in the classification process. Pre- processing reduces the image's overall scale to a more consistent size. The brightness variations in this classification model represent the appearance of the item and the form of the picture. This has no bearing on conversions. The moment's shape descriptor and Haralick characteristics can only be measured across a single channel, therefore this is done. The goal of this algorithm is to detect irregularities in plants that occur in their greenhouses or natural habitat. Start by typing or pasting something into this box, then hit the enter key.160 images of papaya leaves were used to train the model using the Random Forest classifier. Model might be graded with around 70% accuracy. Accuracy may be improved by combining a hueg number of photos and other local characteristics with high - level features like Scale Invariant Feature Transform, Speed Up Robust Features, and DENSE with Bag Of Visual Word.

In 2019, Tejal Chandiwade et al. have devised a method for detecting plant diseases as well as therapies that can be used to avoid disease. The databank accessed from the Internet is correctly segregated and the various databanks are isolated. Plant species are characterized and renamed to create an appropriate database, and then a test database of various plant diseases is created to ensure the project's correctness and level of confidence. Plant disease may be recognized by looking at the diseased leaves of the afflicted plant. The Convolutional Neural Network technique for image processing was utilized to identify plant ailments (CNN). A prototype drone model with a high-resolution camera attached is also intended to be used for live monitoring of huge agricultural regions, as well as to take photographs of plants that might be used as a software input to determine whether the plant is safe. Plant diseases may be detected using two methods: artificial neural networks (ANN) and support vector machines (SVM). The study's main goals were to identify crop illness and treat it once it was discovered. This might aid in reducing illness and increasing potency. The suggested system is Python-based and has a 78 percent accuracy rate. The use of Google's GPU can improve processing accuracy and speed. The gadget might be mounted to drones and used to monitor agriculture fields from above.

In 2020, Yan Guo et al. noted in today's dynamic world, Plant disease identification is critical for accurate and successful disease prevention. This study proposes a mathematical model based on deep learning for detecting and identifying plant diseases that improve accuracy, generality, and training performance. To begin, in a complicated environment, the area proposal network (RPN) is developed to determine and find the leaves. Based on the findings of the RPN approach, the function of symptoms is then merged into pictures segmented using the Chan–Vese algorithm. The segmented leaves are then put into a transfer learning algorithm on a limited sick- leaf dataset. Black rot, bacterial plaque, and dust disease are all evaluated on the model. The accuracy of the method is 83.57 percent, which is higher than the conventional method, decreasing disease's impact on agricultural output and ensuring long-term agricultural continuity. As an outcome, the industry research deep learning approach has important implications for smart agriculture, nature conservation, and agricultural growth. Because novel methodologies were utilized listed in this article, it had a high level of accuracy.

Aliyu M. Abdu et al. [12] presented In 2020, a plant disease detection system that is based on photographs will be used to track the occurrence and severity of crop changes. The support vector machine and deep learning, two well-known machine learning algorithms for diagnosing plant illnesses using leaves image data, are compared in this study. Most of these image analysis approaches used "simplistic" machine learning models very recently then and still do. For image identification and image processing and analysis studies, the Deep Learning network is swiftly becoming the gold standard. To compare the two models, both (SVM and Deep Learning) were attributed to the large dataset of plant leaf disease images using standard settings and allowing for the three primary aspects of design, computing power, and training set. The findings revealed which model is superior in this case, as well as modifications and which model would be picked based on a set of characteristics. The purpose of this study is to learn more about the critical role of machine learning in today and upcoming food and nutrition security. Because of the use of classical detection methods, the reliability was not as great as in the other articles.

III. PERFORMANCE EVALUTION

A. Experimental Setup:

The Experiments were carried out on Intel(R) Core(TM) i7-8565U CPU @ processor of.80GHz 4 GB of RAM and an increasing clock rate As a platform for execution and simulations, we chose Python with Keras and the library.

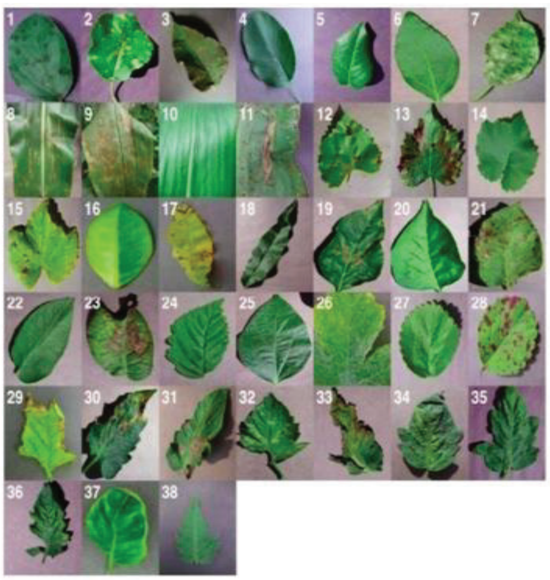

Fig. 1. PlantVillage Dataset

Under controlled settings, a large dataset including 54,305 images of broken and good plant leaf was made accessible. Apple, blueberry, cherry, grape, orange, peach, pepper, potato, raspberry, soy, squash, strawberry, and tomato are among the 14 crop species featured. There are 17 basic diseases, four bacterial diseases, two mold- related diseases, two viral diseases, and one mite-related disease depicted. Every classifier indicates a crop-disease combination, and we aim to predict the crop- disease pair using just the leaves picture. All of the classes in the PlantVillage dataset are shown in Figure2.

B. ALGOROTHM

1. Take dataset as a variable. Say plantVillage.
2. Classifying and labeling of the images basedon size, color, and texture or shape.
3. Preprocessing and augmentation is done in the next step.
4. Passing the dataset to the proposed model i.e. TDM model. So, we will get the predicted output.
5. Testing of output is done, in the model establishment phase.
6. Passing testing data into INC-VGGN model.
7. Updating the sample library.
8. The final output will be the category of detected disease and the accuracy of the proposed model.

IV. PROPOSED APPROACH

The concept of transfer learning for classification has been applied in this paper. Transfer learning is commonly expressed in image classification through the use of pre-trained models. To handle an issue comparable to the one we're dealing with, a pre-trained model was trained on a big benchmark dataset. For our research, we used five pre- trained models: Inception v3, InceptionResNet v2, and ResNet50, MobileNet, andDenseNet169, which are the widely-used image recognition models on the various datasets.

Our model consist of 4 major steps:

A. Plant Disease image collection

From the training stage through the assessment of detection techniques' output, suitable datasets are necessary at all set of object recognition studies. In this part, only the datasets from various sources have been collected and downloaded from the internet. Generally consists of the images which afterwardare being examined for detection for Diseases.

B. Image enhancement and preprocessing of images

Images acquired from the Internet came in a variety of formats, resolutions, and quality levels. Final photographs were normalized to increase uniformity and optimize feature extraction before being used as a dataset for a deep neural network classification. Further, the picture preparation approach included manually cropping all of thephotos, creating a rectangle from around leaves to draw attention to the important area (plant leaves). Photos of a lower resolution and a dimension of fewer than 500 pixels were not deemed legitimate images for the dataset throughout the collection phase.

Furthermore, only photos with a greater resolution of the region of interest were identified as suitable for the dataset. As a result, photos were guaranteed to have all of the necessary information for feature learning. The dataset's images were reducedto 256 pixels to save training time, which wascomputed automatically by a Python scriptusing the OpenCV framework.

By searching the Internet, you can get a wealth of information, however, their accuracy is frequently questioned. To validate the correctness of the classes in the collection, agricultural expert reviewed leaf pictures and tagged all of the photos with applicable disease acronyms, which was first classified using akeywords search. It is well recognized that for the training and validation datasets, it is critical to use reliably identified images. Only in this manner can a suitable and trustworthy detecting model be created. At this point, any duplicate photographs that remained after the first cycle of gathering and categorizing photos were excluded from the analysis.

Image Augmentation

The main goal of enhancement is to enhance the dataset size and introduce slight distortion to the pictures, reducing over-fitting during the training phase. Over-fitting occurs whenever a statistical model depicts background data or mistakes rather than the structural relationships in machine learning and statistics. The transformation techniques employed inappropriate changes included Fourier that transform, perspective evolution, and basic picture rotations. Affine transformations (linear transformations and vector addition, respectively) were used to define translations and rotations were in the produced image, In the original picture, all parallel lines remained parallel. 3 Reference points from the original image, but also their matching places in the output image, are required to find a projection matrix. Prospects for the future necessitated the use of a 3x3 transformation matrix. After the change, horizontal lines would stay straight. Simple picture rotations, as well as rotations on additional axes of varied degrees, were used in the augmentation technique. In the first row, pictures are created by applying a transformation matrix to a single photo, in the second row, images are done by applying the consider finding to the input image, and in the third row, photos are obtained by simply rotating the input image.The augmenting technique was chosen to meet the requirements; leaves in the wild setting may have different visual perspectives.

C. Model Establishment

Labeled Samples has three steps:

1. Model Training: The Testing set is the ultimate standard for assessing the model. It's used only once a model has finished all of the required training.
2. Model Testing: The Testing set is the ultimate standard for assessing the model. It's used only once a model has finished all of the required training.
3. Model Validation: A model is indirectly influenced by a validation collection. The Dev set or Production set is another name for the validation package.

D. Result and Evaluation

It is the final phase of the whole process in which we expect the output of the method. In which firstly, the Sample Library is updated and then the detected diseases are classified based on their category.

Training with Neural Network

A dataset was suggested for training a deep convolutional neural network to build an image classification model. The layer's attributes are a set of trainable cores with a restricted region of interest that spans the input vector's depth. The data collection was used to train the model (In the case of a Neural Network, the weights and biases). This data is seen and learned by the model.

Performed Test

The dataset into two sets: a training phase and a testing data. The training phase is used to train a neural network, while the testing data can be used to test the efficiency of artificial neural networks. Since both the original and expected outcomes for the testing set are known, our prediction's accuracy can be calculated. For the accuracy test, a predictive model was assessed using a ten-fold cross-validation process. The cross-validation approach was used after every 1000 training iterations. The top-1 graphically represents the test's overall estimated result, which is used to see if the top class is the same as the target label.

MODEL

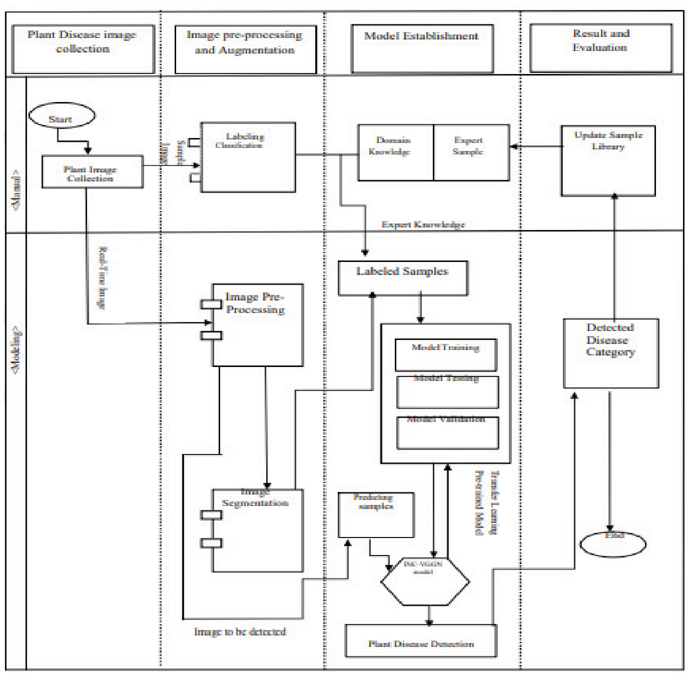

Fig. 2. Proposed Plant Disease Detection Model

V. RESULT

In this part, we provide our findings. We experimented with pre-trained models like Inception v3, InceptionResNet v2 and ResNet 50, MobileNet, and Densenet169 by fine-tuning the network's final layers. On the upper edge of the transfer learning structures, we've added four custom convolutional and max-pooling layers.

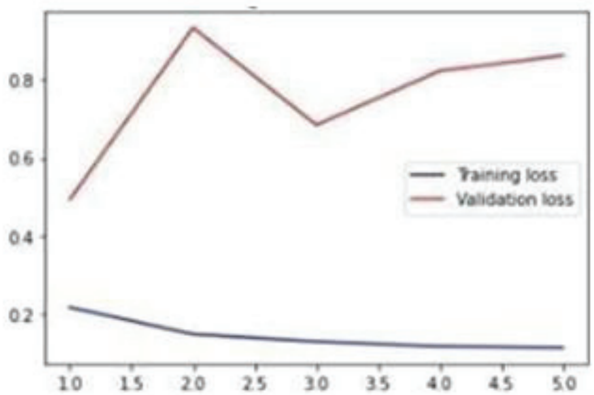

Fig. 3. Training and Validation Accuracy

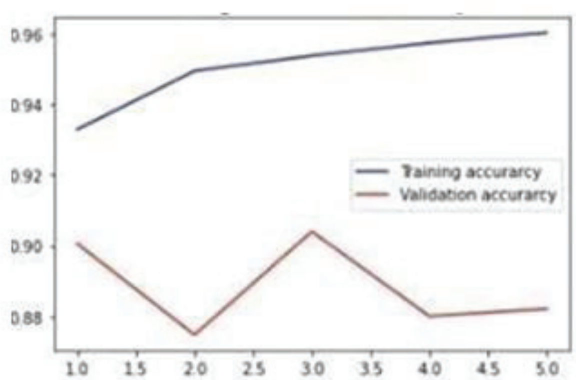

Fig. 4. Training and Validation loss

VI. CONCLUSION

Plant diseases pose a significant threat to global storage supplies. This research shows that deep learning may be used in conjunction with a convolutional neural network to enable autonomous sickness detection through image categorization. A public dataset of 54,306 pictures of harmed and good plant leaves was used to train a deep learning model to recognize different crops and disease severity of 38 distinct classes, comprising 14 crop species and 26 disorders. Using deep learning algorithms, a novel method for automatically identifying and detecting plant illnesses from leaf images was studied in this study. The suggested approach was able to tell the difference between healthy leaves and illnesses that could be seen. The entire technique was described, from obtaining the necessary pictures for training and testing through types of methods and eventually the deep CNN learning and good technique. ntually, the deep CNN learning and good technique.

REFRENCES

[1] M.Sharif, M. A. Khan, Z. Iqbal, M. F.Azam, M. I.U. Lali and M. Y. Javed, "Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection," Computers and Electronics in Agriculture, vol. 150, pp. 220- 234,2018.

[2] S. Dimitriadis, C. Goumopoulos," Applying Machine Learning to Extract New knowledge in Precision Agriculture Applications," Panhellenic Conference on Informatics, pp. 100-104, 2008.

[3] Akhtar, A. Khanum, S. A. Khan, and A. Shaukat, "Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques," International Conference on Frontiers of Information Technology, pp. 60- 65, 2013.

[4] K.P. Ferentinos, "Deep learning models for plant disease detection and diagnosis"Computers and Electronics in Agriculture, vol. 145, pp. 311-318, 2018.

[5] S. P. Mohanty, D. P. Hughes, and M. Salathe, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in plant science, pp.1-10,2016.

[6] S. Sladojevic, M. Arsenovic,A. Andela, Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification, "Computational intelligence and neuroscience, pp. 1- 12,2016.

[7] K. R. Gavhale and U. Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques," IOSR Journal of Computer Engineering, vol. 16, pp. 10-16, 2014.

[8] J. Pujari, R. Yakkundimath and A. Byadgi, "Image Processing Based Detection of Fungal Diseases in Plants," International Conference on Information and Communication Technologies (ICICT), vol. 46, pp. 1802- 1808, 2015.

[9] M. A. Karol, D. Gulhane and T. Chandiwade, "Plant Disease Detection using CNN & Remedy," International Journal of Advanced Research of Electrical, Electronics and Instrumentation Engineering, vol. 8, no. 3, pp. 622-626,2019.

[10] D. Saxena, and A.K. Singh, "A proactive autoscaling and energy-efficient VM allocation framework using online multi-resource neural network f o r cloud data center", Neurocomputing 426 (2021): 248-264.

[11] Kumar, Jitendra, DeepikaSaxena, Ashutosh Kumar Singh, and Anand Mohan., "Biphase adaptive learning- based neural network model for cloud datacenter workload forecasting." Soft Computing (2020): 1-18.

[12] Saxena, Deepika, and Ashutosh Kumar Singh. "Auto- adaptive learning-based workload forecasting in the dynamic cloud environment." International Journal of Computers and Applications (2020):1-11.

[13] M. A. Hussein and A. H. Abbas, "Plant Leaf Disease Detection Using Support Vector Machine," Al- Mustansiriyah Journal of Science, vol. 30, no. 1, pp. 105-110,2019.

[14] N. K. Durga and G. Anuradha, "Plant Disease Identification Using SVM and ANN Algorithm," International Journal of Recent Technology and Engineering (IJRTE), vol. 7, pp. 471-473, 2019.

[15] S. Pavithra, A. Priyadharshini, V. Praveena andT. Monika, "PADDY LEAF DISEASE DETECTION USING SVM CLASSIFIER," International Journal of communication and computer Technologies, vol. 03, no. 01, pp. 16- 20,2015.

[16] E. Hossain, M. F. Hossain, and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease using KNN Classifier," International Conference on Electrical, Computer and Communication Engineering (ECCE), pp.1-7,2019.

[17] S. Ramesh, R. Hebbar, P. Bhat and P. V. Vinod, "Plant Disease Detection Using Machine Learning," International Conference on Design Innovations for 3Cs Compute Communicate Control, pp. 41-45, 2018.

[18] S. Sharma, V. Kaur, and N. Dhillon, "PLANT DISEASE CLASSIFICATION WITH KNN-SVM CLASSIFICATION," International Journal of Advance Engineering and Research Development, vol. 5, no. 05, pp.752-758,2018.