Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

PLANT LEAF DISEASE DETECTION USING MACHINE LEARNING

[1]BELLAMKONDU SINDHUJA, [2]S V SUKEERTHANA PHANISREE, [3]MEESALA JAYANANDITHA, [4]Dr.K.V.UMA MAHESWARI
*[123]B.Tech Student, [4]*Associate Professor
Department of CSE
Dr. K. V. Subba Reddy Institute Of Technology, Kurnool

ABSTRACT:

Rapid human population growth requires corresponding increase in food production. Easily spreadable diseases can have a strong negative impact on leaf yields and even destroy whole crops. That is why early disease diagnosis and prevention are of very high importance. Traditional methods rely on lab analysis and human expertise which are usually expensive and unavailable in a large part of the undeveloped world. Since smartphones are becoming increasingly present even in the most rural areas, in recent years scientists have turned to automated image analysis as a way of identifying crop diseases.

1. INTRODUCTION

Human population steadily continues to grow, and along with it the need for food production increases. According to the UN projections [1], human population is expected to reach 9.7 billion in 2050, 2 billion more than today. Considering that most of the population growth is to occur in the least developed countries (around 80% increase in the next 30 years), where the food scarcity is the main problem, it is easy to conclude that minimizing food loss in those countries is a primary concern. It is estimated that the yield loss worldwide is between 20 and 40 percent [2], with many farms suffering a total loss.

Tomato and potato are a major commercial crop grown all over the world. It is sensitive to various illnesses, which reduces tomato and potato quality and yield while also causing significant economic losses. Tomato and potato grey leaf spot is a common disease that damages and kills the leaves of tomato and potato es, preventing them from growing and producing fruit. The infection that produces grey leaf spots on tomato and potato es is brutal to remove. Contact, invasion, latency, and onset are the four phases of infection for the pathogen that causes tomato and potato grey leaf spot. As a result, early preventative and control methods are suitable before a large-scale pandemic. Early disease detection can also aid in reducing pesticide usage and pollution, as well as the quality, safety, and health of tomato and potato es. Traditional disease detection systems cannot meet large-scale planting demands due to low diagnostic efficiency and fast disease transmission, and plants typically miss the appropriate management time [1, 2]. Manually detecting leaf disease with the naked eye needs a team of professionals and ongoing monitoring. When the farm is large, it is costly. As a result, image processing techniques may be used to automatically detect illnesses in leaves, saving time, money, and effort as compared to traditional methods. The early detection of illnesses in leaves improves crop productivity. Disease-affected leaves may be found at an early stage using image processing techniques like as segmentation, identification, and classification, and crop yield and quality can be improved. Many farmers lack the resources or understanding on how to contact specialists, which makes it more costly, time-consuming, and inaccurate. In this case, the suggested approach proved to be more advantageous in terms of crop observation. The technique is more accessible and less costly when plant illness is detected using leaf symptoms. It takes less time, effort, and precision to use an automated detection technique. Manually detecting leaf disease with the naked eye needs a team of professionals and ongoing monitoring. When the farm is large, it is costly. As a result, image processing techniques may be used to automatically detect illnesses in leaves, saving time, money, and effort as compared to traditional methods. The early detection of illnesses in leaves improves crop productivity. Disease-affected leaves may be found at an early stage using image processing techniques like as segmentation, identification, and classification, and crop yield and quality can be improved. Many farmers lack the resources or understanding on how to contact specialists, which makes it more costly, time-consuming, and inaccurate. In this case, the suggested approach proved to be more advantageous in terms of crop observation. The technique is more accessible and less costly when plant illness is detected using leaf symptoms. It takes less time, effort, and precision to use an automated detection technique. Image processing technology can quickly and accurately diagnose illnesses based on their features. Disease prevention approaches may be applied fast, and efforts to avoid additional illnesses can be performed using this strategy. People used to identify tomato and potato ailments based on their own experiences, but the ability to discern between various diseases is limited, and the process is time-consuming. Machine learning and image processing technologies are fast expanding and more widely employed in various industries, including agriculture. The following are two key contributions of this research: R-CNN framework for classifying leaf diseases ii) comparison of different classifiers.

2. LITERATURE SURVEY

1)**The global burden of pathogens and pests on major food crops**

AUTHORS: Savary, Serge, et al.

Crop pathogens and pests reduce the yield and quality of agricultural production. They cause substantial economic losses and reduce food security at household, national and global levels. Quantitative, standardized information on crop losses is difficult to compile and compare across crops, agroecosystems and regions. Here, we report on an expert-based assessment of crop health, and provide numerical estimates of yield losses on an individual pathogen and pest basis for five major crops globally and in food security hotspots. Our results document losses associated with 137 pathogens and pests associated with wheat, rice, maize, potato and soybean worldwide. Our yield loss (range) estimates at a global level and per hotspot for wheat (21.5% (10.1–28.1%)), rice (30.0% (24.6–40.9%)), maize (22.5% (19.5–41.1%)), potato (17.2% (8.1–21.0%)) and soybean (21.4% (11.0–32.4%)) suggest that the highest losses are associated with food-deficit regions with fast-growing populations, and frequently with emerging or re-emerging pests and diseases. Our assessment highlights differences in impacts among crop pathogens and pests and among food security hotspots. This analysis contributes critical information to prioritize crop health management

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

to improve the sustainability of agroecosystems in delivering services to societies.

2)**Using deep learning for image-based plant disease detection**

AUTHORS: Mohanty, Sharada P., David P. Hughes, and Marcel Salathé

Crop diseases are a major threat to food security, but their rapid identification remains difficult in many parts of the world due to the lack of the necessary infrastructure. The combination of increasing global smartphone penetration and recent advances in computer vision made possible by deep learning has paved the way for smartphone-assisted disease diagnosis. Using a public dataset of 54,306 images of diseased and healthy plant leaves collected under controlled conditions, we train a deep convolutional neural network to identify 14 crop species and 26 diseases (or absence thereof). The trained model achieves an accuracy of 99.35% on a held-out test set, demonstrating the feasibility of this approach. Overall, the approach of training deep learning models on increasingly large and publicly available image datasets presents a clear path toward smartphone-assisted crop disease diagnosis on a massive global scale.

3)**A practical plant diagnosis system for field leaf images and feature visualization**

AUTHORS: Fujita, E., et al.

An accurate, fast and low-cost automated plant diagnosis system has been called for. While several studies utilizing machine learning techniques have been conducted, significant issues remain in most cases where the dataset is not composed of field images and often includes a substantial number of inappropriate labels. In this paper, we propose a practical automated plant diagnosis system. We first build a highly reliable dataset by cultivating plants in a strictly controlled setting. We then develop a robust classifier capable of analyzing a wide variety of field images. We use a total of 9,000 original cucumber field leaf images to identify seven typical viral diseases, Downy mildew and healthy plants including initial symptoms. We also visualize the key regions of diagnostic evidence. Our system attains 93.6% average accuracy, and we confirm that our system captures important features for the diagnosis of Downy mildew.

3. SYSTEM ANALYSIS

SYSTEM ARCHITECTURE:

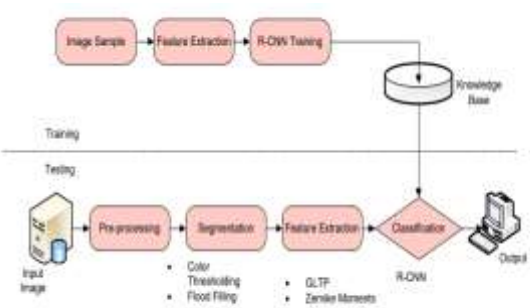

Fig. 1. Architecture of R-CNN-based Plant Leaf Disease Detection.

EXISTING SYSTEM:

Human population steadily continues to grow, and along with it the need for food production increases. According to the UN projections human population is expected to reach 9.7 billion in 2050, 2 billion more than today. Considering that most of the population growth is to occur in the least developed countries (around 80% increase in the next 30 years), where the food scarcity is the main problem, it is easy to conclude that minimizing food loss in those countries is a primary concern. It is estimated that the yield loss worldwide is between 20 and 40 percent with many farms suffering a total loss.Easily spreadable diseases can have a strong negative impact on leaf yields and even destroy whole crops. That is why early disease diagnosis and prevention are of very high importance.

DISADAVANTAGES:

- ❖ Data Collection Problem
- ❖ It searches from a large sampling of the cost surface.

PROPOSED SYSTEM:

Traditional methods for detecting diseases require manual inspection of leafs by experts. This process needs to be continuous, and can be very expensive in large farms, or even completely unavailable to many small farm holders living in rural areas. The Leaf village Dataset is used it consists of images of leaf leaves taken in a controlled environment. In total, there are 54 306 images of 14 different leaf species, distributed in 38 distinct classes given as species/disease pair. Classical methods rely on image pre-processing and the extraction of features which are then fed into one of the ML algorithms. Popular algorithm Connected Neural Networks (CNN. Tomato and Potato leaf are taken into consideration.

ADAVANTAGES:

- ❖ Machine learning algorithm optimizes both variables efficiently, continuous or discrete
- ❖ Gives a number of optimum solutions, not a single solution. So different image segmentation results can be obtained at the same time
- ❖ Large number of variables can be processed at the same time.
- ❖ It can optimize variables with highly complex cost surfaces.

4. MODULE DESCRIPTION

1. Data preprocessing:-

Data is stored in colab. We can download the data and laad the datsets , clean the data then after processes the data

2. Support Vector Machines:-

SVM is a supervised learning algorithm used for classification or regression problems. Classification is done by defining a separating hyperplane in the feature space. In the original form, it performs linear classification on two classes. By using kernels, it can also perform nonlinear classification. Kernels are used for an efficient transformation of the original feature space into high dimensional or infinite dimensional feature space, allowing for highly non-linear hyperplanes. SVM can fit highly complex datasets and at the same time exhibit good generalization propertie.

3.k-Nearest Neighbours:-

k-NN [7] is a very simple algorithm often used for classification problems. It is both non-parametric (doesn't have a fixed number of parameters) and lazy learning (doesn't have a training phase). k-NN works under the assumption that most samples from the same class are close to each other in the feature space. When determining the class of the sample, k-NN will look at its k closest neighbours and decide to which class it belongs by the simple majority rule. Small values of k will allow for higher non-linearity but will be sensitive to outliers. High values of k achieve good generalization but fail to fit complex boundaries. The best value for parameter k is determined experimentally For this dataset, small values of k were shown to give the best results. Varying k from 1 to 9 doesn't change the accuracy much, with best result being 78.06%

much lower than the SVM. We used k=5 in this work

4.Fully Connected Neural Network :-

FCNN is the simplest type of artificial neural networks. It is a supervised learning algorithm able to model highly non-linear functions. As opposed to SVM and k-NN, it does not converge to the global optimum, but when properly configured, it usually gives good enough results. we used an FCNN with four hidden layers with 300, 200, 100 and 50 neurons per layer, respectively. Activation function in hidden layers is a rectified linear unit (ReLU), with a softmax in the output layer [8]. We used L2 regularization with regularization parameter equal to 0.3. Adam optimizer with default parameters was used This configuration gave us the accuracy of 91.46% on the test set.

Our primary objective is to develop a model to categorize input plant leaf pictures as healthy or unhealthy. The disease kind is also determined if a disease is detected on a plant leaf. Our study compares the R-CNN Classifier to previously established tomato leaf disease detection utilizing fuzzy SVM [15] and CNN [16] Classifiers to detect and categorize tomato leaves suffering from common illnesses. Fig. 1 shows the architecture of the R-CNN-based plant disease detection system. The proposed technique includes image capturing, preprocessing, segmentation, feature extraction, classification, and performance assessment.

A. Dataset Description The dataset utilized for this investigation has seven primary classifications. Six leaves classes represent unhealthy, while one represents the healthy leaf class. Each class has 105 examples for a total of 735 leaf images. A classification strategy is required to categorize input photos into one of the classes specified in Fig. 2 for a given image of an apple leaf.

B. Image Pre-processing The visual noise of the tomato leaf is made up of dewdrops, dust, and insect feces on the plants. The input RGB image is transformed to a grayscale image for accurate results to remedy these concerns. The image size in this circumstance is relatively large, needing image resize. The image size is reduced to 256 * 256 pixels.

C. Image Segmentation Plant disease detection and categorization rely heavily on image segmentation. The image is simply divided into various things or sections. It analyses visual data to extract information that may be used for further processing. Our prior work [15] is used to accomplish color thresholding and flood filling segmentations.

D. Classification using R-CNN Rectangular regions are combined with convolutional neural network characteristics in R-CNN (Regions with Convolutional Neural Networks). The R-CNN algorithm employs a two-stage detection procedure. The first stage identifies a set of picture areas that includes a diseased part. In the second stage, each region's object is categorized.

1) R-CNN procedure: The following three approaches are employed to build an R-CNN based algorithm.

a) To find regions in a photograph that could contain a diseased part. Region suggestions are the names given to these locations.

b) Extract CNN characteristics from the region suggestions.

c) To categorize the objects, use the characteristics that were retrieved. The R-CNN generates region recommendations using a mechanism similar to Edge Boxes [10]. The proposed elements have been chopped and scaled out of the image. CNN then classifies the clipped and resized regions. Finally, a support vector machine (SVM) trained on CNN features refine the region proposal bounding boxes. A pre-trained convolution neural network is used to build an R-CNN detector, also known as transfer learning (CNN). As a starting point for learning a new task, we will use a pre-trained image classification network that has already learned to extract robust and informative features from raw photographs. A portion of the ImageNet database [10], which is used in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) [11], is used in the great majority of pretrained networks. These networks have been trained on over a million photographs and can categorize a large number of them. Transfer learning with a pre-trained network is typically much faster and easier than training a network from the ground up.

5. CONCLUSION

Machine learning and image processing technology benefits from traditional manual diagnostic and recognition procedures for crop disease diagnosis. A small number of unhealthy image samples is necessary. Deep learning is one of the most well-known methods of artificial intelligence. Computer vision is one of the many fields where deep learning is used. It is capable of picture categorization, object identification, and semantic segmentation. The disease diagnostic approach based on the deep convolutional network (CNN) minimizes comprehensive picture pre-processing and feature extraction methods compared to standard pattern detection techniques. It employs an end-to-end structure to simplify the detecting procedure. In this study, we compared the best classifier for effectively classifying tomato and potato leaf disease with an accuracy of 96.735 percent.

REFERENCES

[1]United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights (ST/ESA/SER.A/423).

[2] Savary, Serge, et al. "The global burden of pathogens and pests on major food crops." Nature ecology & evolution 3.3 (2019): 430.

[3] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." Frontiers in plant science 7 (2016): 1419.

[4] Fujita, E., et al. "A practical plant diagnosis system for field leaf images and feature visualization." International Journal of Engineering & Technology 7.4.11 (2018): 49-54.

[5] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics 6 (1973): 610-621.

[6] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20.3 (1995): 273-297.

[7] Cunningham, Padraig, and Sarah Jane Delany. "k-Nearest neighbour classifiers." Multiple Classifier Systems 34.8 (2007): 1-17.

[8] Haykin, Simon. Neural networks: a comprehensive foundation. Prentice Hall PTR, 1994.

[9] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[10] Duan, Kai-Bo, and S. Sathiya Keerthi. "Which is the best multiclass SVM method? An empirical study." International workshop on multiple classifier systems. Springer, Berlin, Heidelberg, 2005.

[11] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.