iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XII **Month of publication:** December 2022

DOI: https://doi.org/10.22214/ijraset.2022.48198

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XII Dec 2022- Available at www.ijraset.com

Plant Disease Detection and Cure RecommendationUsing Machine Learning

Gayatri Sanjay Gaikwad[1], Mihir Girish Bhattad[2], Sanyukta Rajendra Holkar[3], Pritamsingh Solanki[4], Dr. V. A. Suryawanshi[5], Dr. M. B. Wagh[6]
[1, 2, 3, 4]*Student, Computer Engineering Trinity Acadaemy of Engineering*Pune
[5]*Guide,* [6]*HOD, Computer Engineering Trinity Academy of Engineering*, Pune, India

Abstract: ***Infection discovery in plants is a significant task thathas to be done in agriculture. To recognize the diseases in leaves, a continuous observation of the plants is required. Program- based identification of diseases in plants makes it easier to detect and reduces human efforts and time-saving. The proposed algorithm distinguishing sickness in plants and classifying them more accurately as compared to existing. Traditional approach for disease detection and classification requires huge amount of time and effort. In the last few years, advancement in technology and researchers' focus makes it impossible to obtain optimized solution for it. A convolutional neural network is a form of artificial neural network specifically intended to process pixel input and is used for image recognition.***
Index Terms: Leaf, Diseases, CNN

I. INTRODUCTION

Agriculture is the spine of our nation. Our country is renowned for agriculture . In India, people are most likely interested in agriculture. It plays a very important role in the Indian economy. Around 70 percentage of rural areas depend on agriculture. A total of 17 percentage is paid to GDP and gives the people employment opportunities for 60 percentage of the total population. Hence plant disease detection playsa significant role in agriculture. Artificial intelligence (AI) can help farmers to recover their farming efficiency and diminish environmental hostile influences. To identify and detect the disease on agriculture product, the AI technique is introduced.In this paper, we are presenting a survey for application of AI in detection of diseases in agriculture. The CNN model is designed to suit both healthy and sick leaves; photos are used to train the model, and the output is determined by the input leaf. The method of processing the image is an appropriate technique used in the application of agriculture. Plants are most affected by fungus and various bacterial diseases. The automatic system helps to avoid the farmers to frequentlyvisit the fields. Automatic detection systems also are helpful in the case of large crops. From the symptoms that appear on the leaves, this system detects the diseases. Due to the increase in population, the changes in climate affect plants . With new technologies, plant disease can easily be predicted. Taking care of leaf needs continuous observation. Especially for the diseases that affect production and post-harvest life. A disease leaf is taken and its image is processed. We aim to detect diseases namely Alternaria alternate, Anthracnose, bacterial blight, and Cercospora leaf spot . The area infected, accuracy, and percentages of infection are measured .

II. LITERATURE SURVEY

1) S. Khirade etAl. dived the problem of factory complaint discovery using digital image processing ways and back propagation neural network(BPNN)[1]. Different ways for the discovery of factory complaint using the images of leaves have been developed. BPNN is used for bracket i.e. to member the infected part in splint. After that they've uprooted the features similar as color, texture, morphology, edgesetc. for bracket.

2) Garima Shrestha etAl. stationed the convolutional neural network to descry the factory complaint[2]. This is the first time that a computer has successfully classified 12 factory conditions with88.80 delicacy. The network has 3blocks of complication and pooling layers which makes itcomputationally precious. Also the F1 score of the modelis0.12 which is veritably low because of advanced numberof false negative prognostications.

3) Loyce Selwyn Pinto et al(and other). In this paper Images arepre-processed using k- means clustering and colorful machine learning algorithms like K- Nearest Neighbors,Multi-Class Support Vector Machine, Naive Bayes and Multinomial Logistic Retrogression to get high delicacy. The point set consists of Differ, Energy,Mean, Homogeneity, Standard divagation and tastelessness. In this paper, the image processing is used to descry and classify sunflower crop complaint[3].

4) Jitesh P Shah et al(and other)[4]. A check of 19 papers covering the work on rice conditions and other different fruits and shops has been carried out by the authors grounded on important criteria like size, no. of classes(conditions), Segmentation and pre-processing ways, clas- sifiers and its delicacyetc.In the study Soybeans, Crop complaint Discovery Using Cnns, Serawork Wallelign, and the others The viability of CNN for crop conditions identification in leaves filmland captured in the natural surroundings is presented in this study.

5) A new method for the construction of a crop diseases recognition model based on plant image classification and deep networks. Caffe, a deep learning framework was used to perform the deep CNN training. The experi- mental results on the developed model achieved precisionbetween 91 and 98 percent for separate class tests.[5]

6) A deep neural network and semi-supervised algorithms were trained to distinguish crop species and disease statusof 57 different classes using a publicly available dataset of 86,147 photos. With a detection rate of 1e-5, it was able to score more than 80 percent in the training phasein less than 5 epochs.[6]

7) The proposed algorithm was. executed on Banana, beans,jackfruit, lemon, tomato, and sapota. The Genetic algo- rithm is used for image. segmentation,which is crucial aspect for disease identification. And then if it is affectedby any. diseases than again NN classifier is used to classify the. particular leaf.[7]

8) Prasanna Mohanty and colleagues developed a deep con- volutional neural network using deep learning to detect14 different crops and 26 illnesses. On a held-out set, the training set model obtained an accuracy of 99.35 percent,compared to 31.4 percent on a collection of photographs taken from the web.[8]

III. METHODOLOGY

To find out whether the leaf is diseased or healthy, certain steps must be followed. i.e., Preprocessing, Feature extraction,Training of classifier and Classification. Preprocessing of image, is bringing all the images size to a reduced uniform size. Then comes extracting features of a preprocessed image which is done with the help of HOG. HoG [9] is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients. One of the advantage of HoG feature extraction is that it operates on the cells created. Any transformations doesn't affect this. Here we made use of three feature descriptors.

1) *Hu moments:* Image moments which have the important characteristics of the image pixels helps in describing the objects. Here Hu moments help in describing the outline of a particular leaf. Hu moments are calculatedover single channel only. The first step involves converting RGB to Gray scale and then the Hu moments are calculated. This step gives an array of shape descriptors.

2) *Haralick Texture:*Usually the healthy leaves and diseased leaves have different textures. Here we use Haralick texture feature to distinguish between the textures of healthy and diseased leaf. It is based on the adjacency matrix which stores the position of (I,J). Texture [10] is calculated based on the frequency of the pixel I occupying the position next to pixel J.To calculate Haralick texture it is required that the image be converted to gray scale.

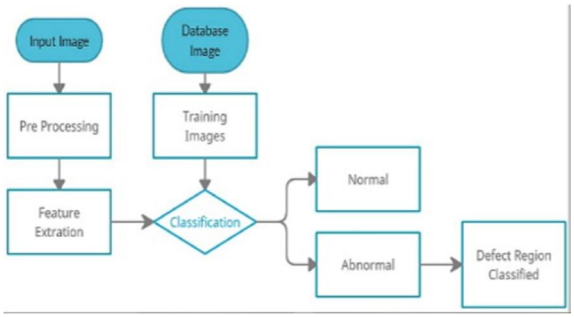


Fig. 1. Proposed Methodology

3) *Color Histogram:* Color histogram gives the representation of the colors in the image. RGB is first converted to HSV colorspace and the histogram is calculated for the same. It is neededto convert the RGB image to HSV since HSV model aligns closely with how human eye discerns the colors in an image. Histogram plot [8] provides the description about the number of pixels available in the given color ranges.The output of this phase is a vector or a matrix representing various gradient descent and various useful features that further get analyzed and classified.

A. Processing of Dataset

Data processing is an important step before visual- ising the information of given features. As we have already collected all the important features required for analysing the attentiveness of a student. The features collected are student emotion, their gaze direction, pos- ture of body and based on this we have given a attention level to each. Each student has assigned with a particular student Id for storing the data and recognizing their features.

B. Classification Algorithm

The Classification algorithm is a Supervised Learning technique that is used to identify the category of new ob- servations on the basis of training data. In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups. Here classification algorithm is used in order to detect the disease. So the result will be disease found or not found.

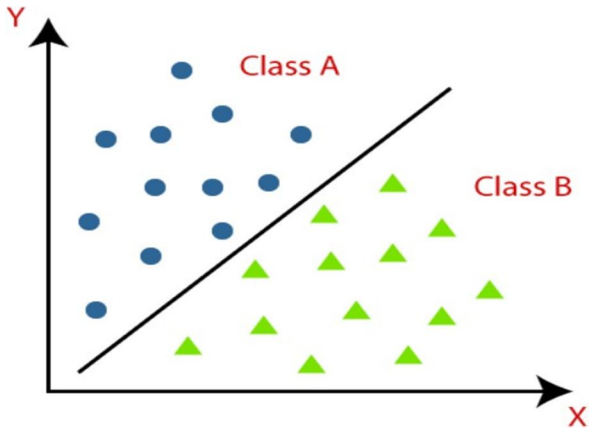


Fig. 2. Classification Algorithm

IV. ALGORITHM

A. K-means Algorithm

1) K-means algorithm is an iterative algorithm that tries to partition the dataset into K pre-defined dis- tinct non-overlapping subgroups (clusters) where each data point belongs to only one group.

2) It tries to make the intra-cluster data points assimilar as possible while also keeping the clusters as different (far) as possible.

3) It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum.

4) The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

5) K-means clustering technique is applied on leaf images to find out the infected region.

6) The K-mean clustering is used to get the data centre of the image and make the clusters of that image and calculates the centre distance from the other cluster.

B. Support Vector Machine(SVM)

Support Vector Machine (SVM) technique is used for classification of images. SVM is the supervised learning method which is usually applied for pattern recognition and classification. It is used to solve both linear and non-linear problems. It basically creates a hyper plane which then separates the data into various classes. For example, it takes a unit of data as an input and then separates those classes if possible, by putting a line as an output. In this algorithm, points that are closest to the line from both classes are found. These points are termed as support vectors. The distance that is computed between the line and support vectors is called as the margin. Then come the optimal hyper plane in which the margin is maximum for the hyper plane. For getting the separation between the two classes as wide as possible, the SVM algorithm creates a decision boundary.

C. Random Forest Algorithm(RF)

Random forest is a supervised learning algorithm. It can be used both for classification and regression. It is said that more the trees it has, more robust a forest is. The main advantage of Random Forest Classifier is to reduce the risk of over fitting and to produce high prediction accuracy value for missing data in the data sets.

The pseudo code for random forest algorithm is -

1) The beginning of random forest algorithm starts with randomly selecting "k" features out of total "m" features.
2) Next, generate the node "d" among the "K" feature,from "m" total feature by using spilt method.
3) After that, again generate daughter node by using split method.
4) Repeat the step I to III until unless get node "1"
5) Repeating the step I to IV for "n" times to create "n" tress in randomly

V. CONCLUSION

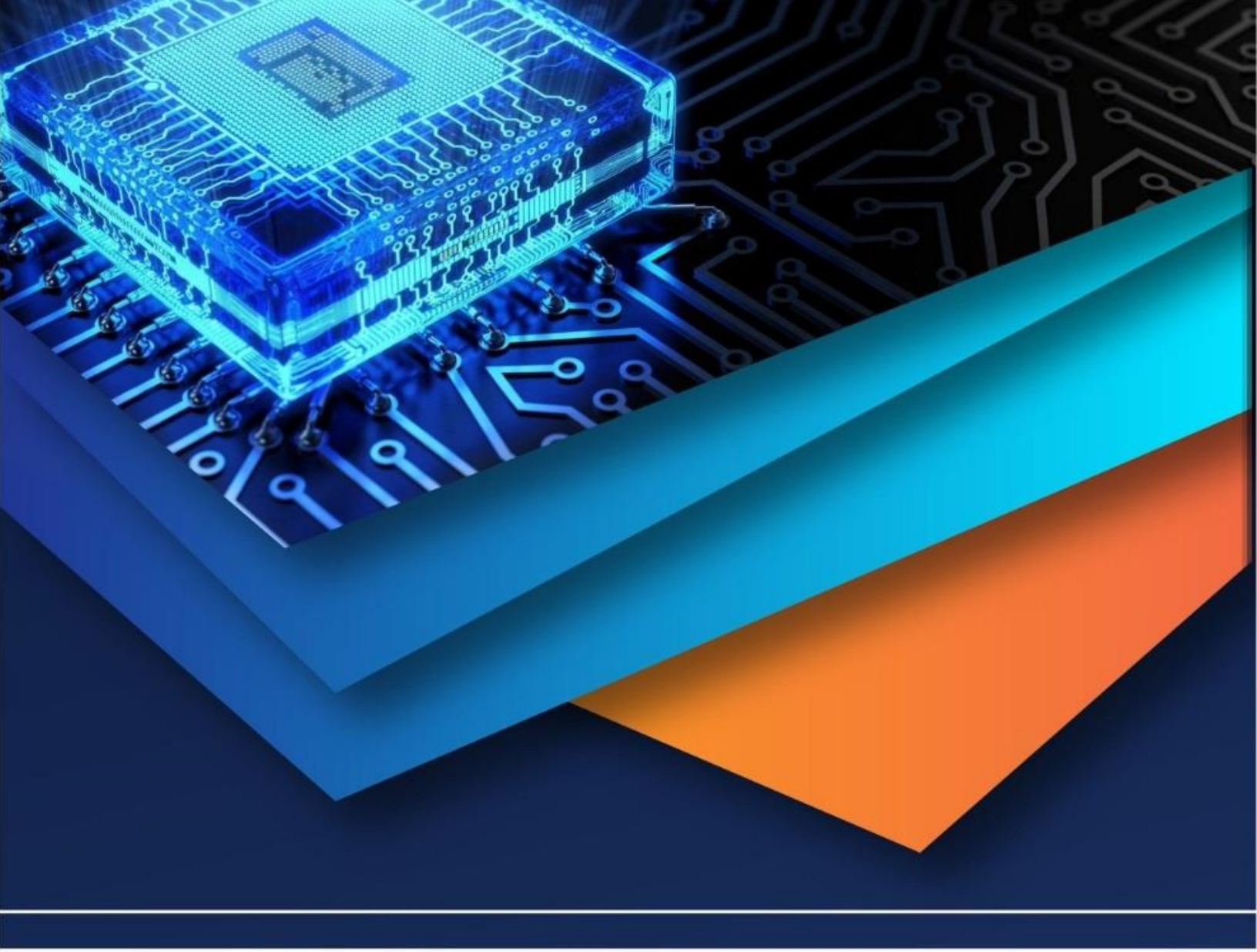
This paper compared the performance results of deep featureextraction and transfer learning for the detection of plant dis- eases and pests. First, we extracted features for fully connectedlayers of these deep models. The obtained deep features had their performances calculated using SVM and KNN classifiers.Deep learning models produced better results compared to the traditional methods. The highest level accuracy scored was 97.86, obtained with the ResNet50 model and SVM classifier. A study has concluded that artificial intelligence is a great toolfor a nation's agronomics, and can be used to improve crop production and yield.

The paper also highlights the different literatures, which reflects various methodologies to detect the diseases in crops.We are successful in creating disease classification techniques used for plant leaf disease detection. A deep learning model that can be used for automatic detection and classification of plant leaf diseases is created. Various types of plants were taken for identification through this work, and the prediction is almost correct.

In our proposed work, we have used various images for detecting leaf diseases. We have used segmentation technique like k-means clustering, for extracting various features Gray Level Co-occurrence Matrix (GLCM) is used and Support Vector Machine (SVM) classifier to classify different types ofdiseases. This process helps us to find the different diseasesin leaves precisely.

REFERENCES

[1] Sachin D. Khirade and A.B. Patil. Plant disease "detection using image processing". In 2015 International Conference on ComputingCommunication Control and Automation, pages 768–771, 2015.
[2] Garima Shrestha, Deepsikha, Majolica Das, and Naiwrita Dey." Plant disease detection using cnn". In 2020 IEEE Applied Signal Processing Conference (ASPCON), pages 109–113, 2020.
[3] Loyce Selwyn Pinto, Argha Ray, M. Udhayeswar Reddy, Pavithra Perumal, and P Aishwarya. Crop disease classification using texture analysis. In 2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT), pages 825–828, 2016
[4] Jitesh P. Shah, Harshadkumar B. Prajapati, and Vipul K. Dabhi. A survey on detection and classification of rice plant diseases. In 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC), pages 1–8,2016.
[5] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culi- brk, and Darko Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience, 2016:3289801, Jun 2016.
[6] Cortes, Emanuel. "Plant disease classification using convolutional net- works and generative adversarial networks." (2017).
[7] M. R. S. M. Ravindra naik, "Plant leaf and disease detection by using hsv features and svm classifier," international journal of engineering science and computing, vol. 6, no. 12, pp. 3794-3797, 2016.
[8] Mohanty, Sharada P., David P. Hughes, and Marcel Salathe. "Using deep learning for image- based plant disease detection." Frontiers in plantscience 7 (2016): 1419.
[9] S. Yun, W. Xianfeng, Z. Shanwen, and Z. Chuanlei, "Pnn based crop disease recognition with leaf image features and meteorological data," International Journal of Agricultural and Biological Engineering, vol. 8, no. 4, p. 60, 2015
[10] J. G. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," Springer Plus, vol. 2, no.660,pp. 1–12, 2013..



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY