RESEARCH
www.commprac.com
ISSN 1462 2815

DISEASE DETECTION ON CASHEW PLANTS USING MACHINE LEARNING AND IMAGE PROCESSING TECHNIQUES

T. Mahendran
Assistant Professor, Department of Computer Applications, Arignar Anna Government Arts College, Villupuram, Tamilnadu, India. Email: maheinfex@gmail.com

Abstract

In India, around 80% of the people depend on agriculture field. Nowadays the agricultural plants and crop have been affected by the disease caused by insects that transfers the infection to other plants in agriculture. During the infection of these diseases, production level has decreased on the farm. Therefore, it is necessary to identify these infections as the earliest to avoid spread one to another plants. A farmer to find the disease and spreading need more knowledge and adequate experience, so in this situation the computational concept that is Image processing and Machine learning is most useful to the farmer. This paper proposed leaf disease detection by feature extraction using GLCM and Deep convoluted neural network (DCNN) based classification. Then the simulation results show the accuracy, recall, precision and f-1 score as the parameters of the proposed method. This technique will detect the disease symptoms of early stage and helps the farmer to encourage pesticides to avoid the spreading of the disease.

Keywords: Cashew Disease, Image Processing, GLCM, DCNN, Feature Extraction and Classification.

1. INTRODUCTION

Commercial cultivation of cashew is taken up in eight states of our country mainly on the west and eastern coast viz., Andhra Pradesh, Goa, Karnataka, Kerala, Maharashtra, Orissa, Tamilnadu and West Bengal. In addition, cashew is also grown in a few pockets of Assam, Chhattisgarh, Gujarat, Meghalaya, Nagaland and Tripura. India has an area of 9.53lakhs ha (2010-11) under cashew with an estimated annual production of about 6.74lakhs tons of raw cashew nuts. India is the third largest producer and exporter of cashew in the world next only to Vietnam and Nigeria. It is the second-largest consumer of cashew and also the biggest processor with the highest acreage under the crop.

The current cashew production of the country accounts for 23.0% of global production. Many small and marginal farmers, especially those living on the coastal belts of India, depend on cashew for their livelihood. Nearly two lakh and more workers, more than 90% of whom are women, are directly employed in cashew processing factories which are concentrated frequently in Kerala, Andhra Pradesh and Maharashtra. Nearly two million people are likely occupied, directly and indirectly in cashew cultivation, processing and marketing.

Cashew farming is taken up in small and marginal holdings and as more than 70% of the cashew area is under this category, cashew plays an important role in the development of small and marginal farmers.

In India in the early 1960s, cashew nut production generates income for many households with an annual production of nuts over a huge amount of tons/year. Nowadays, cashew is one of the most export-oriented horticulture crops in the country. Moreover, the number of processing units is increasing year by year throughout the country. In addition to its socio-economic importance, the cashew tree plays an environmental part. But this cashew plant is affected by various disease factors, this situation; help us with the latest development of IoT, computer vision, machine learning and neural network in many ways and it has created enormous advancement in other disciplines such as medicine, chemical, mechanical and remote sensing domain. In this research work, the cashew plant disease diagnosis system is studied and implemented using a neural network for faster diagnostic purposes [1].

The mainly widespread in various types of diseases that occur in cashew leaves, flowers, stems and flowers are Pets, Viruses, Bacteria and Fungi. Unfortunately, cashew is threatened by many biotic and biotic constraints resulting in significant yield losses. Among biotic constraints, diseases and pests are the most damaging and compromise the cashew nut yield in terms of quality and quantity. More than diseases were reported to infect cashew trees worldwide. Anthracnose foliar blight, fruit rot, and gummosis of twigs and trunks are often considered as the most relevant diseases causing severe damage across cashew-producing countries.

Four diseases including anthracnose, powdery mildew, leaf and nut black rust and bacterial leaf spot disease are responsible for extensive damage. Demonstrate that early action is of the utmost importance in controlling the disease and limiting production losses. The present study aimed to fill this gap by compiling an inventory of the diseases related to different cropping stages of cashew in the main production areas in India.

2. RELATED WORKS

In place, the detection of dangerous viruses and bacteria has been directed to investigate different techniques for mechanized distinguishing proof of plant sicknesses. The ailment can show in different parts of the plant, for example, roots, stem, flower and leaves. Mainly this work focuses especially on controlling diseases. Vijay Singh et al, talked about in paper [2], a strategy for acknowledgement of plant ailments shown on leaves and stems.

The proposed work is using Image Processing and Genetic Algorithm Based Disease Detection. Sachin D. Khirade et al, [3] talked about different procedures to fragment the sick part of the plant. This paper additionally examined some Feature extraction and characterization methods to extricate the highlights of contaminated leaves and the arrangement of plant ailments. From these strategies, we can precisely recognize and group different plant infections utilizing picture symptoms. Yogesh Dandawate and Radha Kokare talked about in paper [4] an approach in light of picture preparation utilized for mechanized plant image characterization given leaf picture handling the examination work is worried about the separation amongst sick and solid soybean leaves utilizing Support Vector Machine classifier.

K. Seetharaman, [5] an early and accurate detection of plant diseases is important for sustainable growth of agricultural productivity. Anthropologists rely on plant defects caused by diseases, pests, poor nutrition or severe weather. Calculating the optimal solution is very costly for large databases, which reduces system performance.

To counter these problems, in this paper, we propose IoT-based real-time automatic detection and classification technique of groundnut leaf disease using hybrid machine learning techniques (GLD-HML). Rama Krishna. and Sahaya Anselin Nisha. As talked about in the paper [6] the work speaks to groundnut leaf sickness extraction and characterization utilizing shading symbolism.

The shading nonexistent change, and shading co-event framework, including extraction, will be done and get a proficiency yield with the neural system, Back spread gives productive ground nut leaf identification with a complex foundation, in this work we grouped just four distinctive illnesses with 97 Al % of effectiveness.

Rajileen Kaur talked about in paper [7] the investigation of recognition of plant ailments and the discovery of a contaminated piece of plants. At first information, pictures are taken and afterwards picture handling is begun.

Foundation and Black pixels are both divided in the initial step. At that point, the Hue and Saturation part of the picture is likewise isolated. In [8], plant leaves have been affected by the disease caused by insects that transfers the infection to other plants in agriculture.

Therefore, it is necessary to identify these infections as the earliest. The banana plant disease detection through image processing becomes more efficient also it is highly essential for farmers in evaluating the plant growth without any manual support economically. This paper proposed leaf disease detection by feature extraction using GLCM and Deep convoluted neural network (DCNN) based classification.

Disease identification in Ujwalla Gawande et al [9], Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques has proven to be more difficult in the field because it is susceptible to a variety of diseases and causes significant losses to farmers. As a result, this research provides improved image processing algorithms for earlier disease identification in cashews leaf, stem and flower.

The images are preprocessed using a histogram pixel localization technique with a median filter and the segmentation is done through region-based edge normalization. Here a novel integrated system is formulated for feature extraction using Gabor based binary patterns with a convolution recurrent neural network. Finally, [10] a region-based convolution neural network is used to identify the disease area by extracting and classifying features to increase disease diagnostic accuracy in the cashews plant.

A manual system has been interchanged in detecting and classifying the disease of cashews plant where the symptoms for consumption of time and the accuracy are low in comparison with the proposed technique.

3. OBJECTIVE OF STUDY

The main purpose of this research is finds what are the types of disease occur in cashew plants, further finds the cashew Blight disease in leaves and recommended appropriate pesticide. The following diseases are mainly affected the cashew plants.

3.1 Healthy Cashew Leaves

The healthy leaves are fresh and attractive manner, in which stage get ready to take buds and flowers.

ISSN 1462 2815

Figure 1: Healthy Cashew Leaf

3.2 Black or Brown Spot Disease

In this disease are arise the black spots that form on the cashew leaves. This disease can be identified by its symptoms like dark of large areas of the leaf, which are brown spots or black spots.

Figure 2: Black Spots on the Leaf

3.3 Anthracnose

Anthracnose is a fungal type of disease, this disease is come up when rainfall coincides with the flowering season and it is destroyed the all parts of the cashew plants.

Figure 3: Anthracnose Affected Nut

3.4 Pink Disease or Die-back

Pink disease is of minor importance because this disease affected level is very low. It is prevalent during South-West monsoon period, the shoots and branches wilt and dry up from the tip downward and hence the name die-back.

Figure 4: Die-Back Affected Fruit And Nut

3.5 Blight

Blight disease is very dangerous one; in starting stage one or more leaves dried and finally whole cashew plant leaves are affected.

ISSN 1462 2815

Figure 5: Blight Lightly Affect

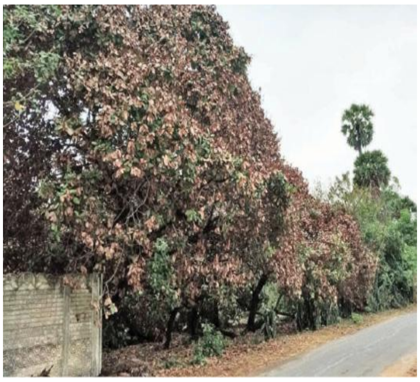

Figure 6: Blight Strongly Affect

4. PROPOSED METHODOLOGY

This proposed technique intends to extract and classify the disease of cashew plants at their starting stage and this can reduce the disease spreading to nearby plants. The dataset of cashew plant leaves has been gathered from pre-historic data from agricultural farms. The cashew plant images have been stored in the database in which 10 samples have been taken into account for disease detection that has been needed in additional processing. The architectural diagram for the proposed system is given in figure-1 and this figure planned initially the image has been preprocessed for image resizing, noise removal and data cleaning. Then the image has been segmented and their feature has been extracted using GLCM and this extracted feature has been classified using DCNN. The classified output will be the disease-detected part of the cashew leaf.

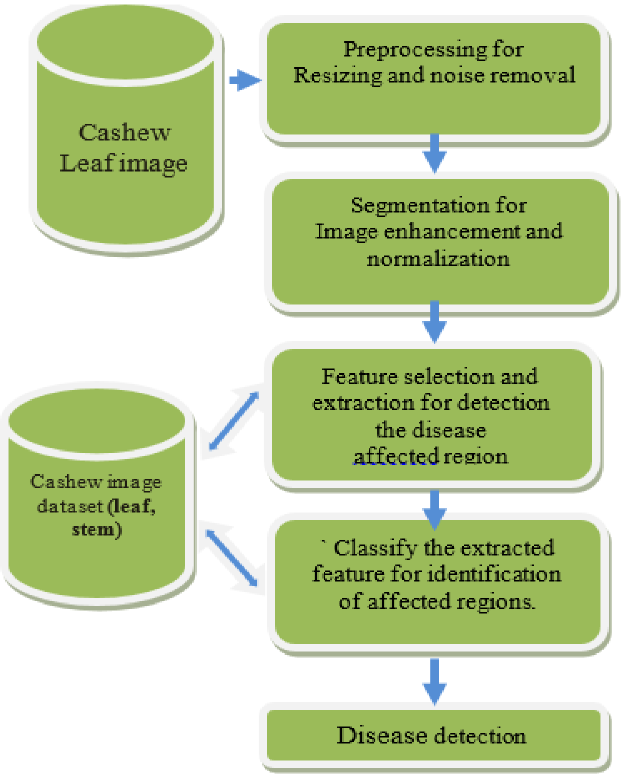

Figure 7: Fundamental Steps in Image Processing To Detect Plant Diseases

4.1 Feature Extraction using Gray Level Co- occurrence Matrix (GLCM)

One of the frequently used textures analyzing techniques is GLCM which is employed in evaluating the characteristics of the image which has been correlated with the statistics of order 2. Matrix design has been done for assessing the relationship among the pixels in the spatial domain. So this technique is on basis of the texture data which has been taken from the output of another correlation. Every input of GLCM is (i, j), according to the pair of occurrences numbers for gray levels as i and j, and this is divided through the d distance of the input image. The calculation of comparison among the various GLCM has been done by the 22 features and a various number of features has been used in existing works this proposed simulation technique differentiates the 22 features of the input image for comparison. Here is the simulation of spatial comparison between the image pixels along with the comparison of pixels which is available in the input image by their direction and distance represented by Θ and d respectively. Every image has been extracted and quantized as 16 grey levels as well as 4 GLCMs (M) for Θ = 0, 45, 90, along with 135 degrees where d = 1 has been acquired. Every GLCM will be extracted with positive features. Hence for every image 0 features have been extracted. The feature normalization has been ranging from 0 to 1 which gets through the classifiers as well as every classifier acquires similar pair of features. Standardized GLCM is computed as shown in Eq.1.

$$G(i,j) = \frac{N(i,j)}{\sum_{m=0}^{l-1} \sum_{n=0}^{l-1} N(m,n)}$$

here i and j represent the grey values in the l-grey image. N (i, j) denotes co-occurrence relative recurrence frequency matrix by:

$$N(i,j) = num\ (\{[(x_1,y_1),(x_2,y_2)]\}|x_2 - x_1 = \ d\ cos\ \theta, y_2 - y_1 = dsin\theta, I(x_1,y_1) = i, I(x_2,y_2) = j)$$

here pixel positions are represented by $(x_1, y_1)$and (x_2, y_2), and *I(·)* denotes the pixel's gray level. Num· means the quantity of pixel matches satisfying the conditions of comparison.

4.2 Classification using Deep Convolution Neural Network (DCNN)

Once the feature vector is obtained from the input image, then the image is represented as a fixed-length vector, and requires a classifier for classifying feature vectors. Generally, a common CNN contains few layers like input, convolution, activation, pool, fully connected, and output layer. The layers of CNN establish the relationship among various nodes and the input information is forwarded through various layers sequentially, and the pool structure of CNN continuously performs decoding, deducing, convergence and mapping the signal features of input information to the feature space of the hidden lay. A fully connected layer performs classification and the outputs are obtained concerning the extracted features. The architecture of DCNN is shown in figure 8.

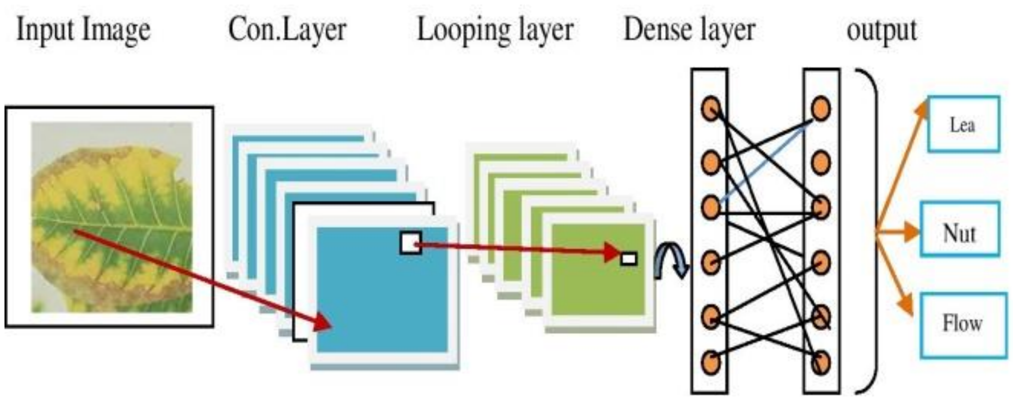

Figure 8: DCNN Architecture

Deep Convolution is a significant analytical mathematical operation. A third function is generated by this operator using two functions namely f and g, where the third function represents the overlapping area between the functions f and g which is either translated or flipped and the calculation is given in Eq.2.

$$z(t)^{def} = f(t) * g(t) = \sum_{\tau=-\infty}^{+\infty} f(\tau)g(t-\tau)$$

The integral form for the above equation is given by,

$$z(t) = f(t) * g(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{+\infty} f(t-\tau)g(\tau)d\tau$$

In classifying image, a digital image is considered as a discrete function f (x, y) of a 2D space. By considering g (x, y), a 2D convolution function, the output image z (x, y) is given by

$$z(x,y) = f(x,y) * g(x,y)$$

Here, the convolution operation is employed for extracting the features of an image. Likewise, with applications involving deep learning approaches, when a colour image is given as input, it is a high-dimensional array with a dimension of 3 × image width × image length; thus, the convolution kernel in CNN is described as accounting in the deep learning algorithm. The computing parameter also is a high-dimensional array. Then, for the given 2D image, the respective convolution operation is given in Eq.3.

$$the\ a\ z(x,y) = f(x,y) * g(x,y)$$

$$= \sum_{t}\sum_{h} f(t,h)g(x-t,y-h)$$

The integral form is the following:

$$z(x,y) = f(x,y) * g(x,y)$$

$$= \iint f(t,h)g(x-t,y-h)dtdh$$

For the given convolution kernel of m × n,

$$the\ z(x,y) = f(x,y) * g(x,y)$$

$$= \sum_{t=0}^{t=m} \sum_{h=0}^{h=n} f(t,h)g(x-t,y-h)$$

here f indicates the input G with the convolution kernel size of m and n. Convolution is generally realized in the computer as a product of a matrix. Consider that the image size is M × M and the convolution kernel size is n × n. While computing, the convolution kernel multiplies the image of size n × n with every image region, which is like obtaining the region of the image with n × n which represents the length of the column vector. While operating zero-zero padding with step 1, totally (M − n + 1) ∗ (M − n + 1) results are possible; when representing the small regions of the image as n x n column vector, the actual image can be given as the matrix [n∗ n∗ (M − n + 1)]. Consider K as the number of convolution kernels, the result of the input image after convolution is k∗ (M − n + 1) ∗ (M − n + 1); i.e., several convolution kernels × image width × image length. Every trained parameter in the proposed model is initialized with a random value ranging from −0.05 to 0.05. The training phase comprises two stages namely forward and back propagation.

5. PERFORMANCE ANALYSIS

This part discusses the simulation consequences for cashew leaf disease detection from the input dataset. Various outputs have been discussed below: initially, the dataset folder has been generated where the different images for cashew leaves are available for the testing process. The indication of leaf disease has been identified with higher accuracy and also trained properly and collected in the dataset. Diseases inventoried several symptoms in leaves, stems, fruit and flower. The first image (Fig.1.) represented a healthy cashew plant and the second image (Fig.2.) represented a disease affected by Black or Brown Spot disease. If the Black Spot disease affected leaf is fully tried and cashew nut yield is also destroyed. Symptoms were characterized on young leaves of cashew seedlings or adult plants by water-soaked spots initially and become orange-brown to light-reddish with the age of the fungus. Generally different diseases (Fig.3.) are Leaves and cashew nuts affected in Anthracnose etc. The below Fig.9. Represent the Comparison of Existing technique with proposed for fruit disease detection.

Table 1: Comparison Between Proposed And Existing Techniques

Parameters	Precision	Recall	Accuracy	F1-score
KNN	86	89	64	90
CNN	72	77	71	74
DNN	89	94	88	94
PRO_DCNN	97	98	97	98

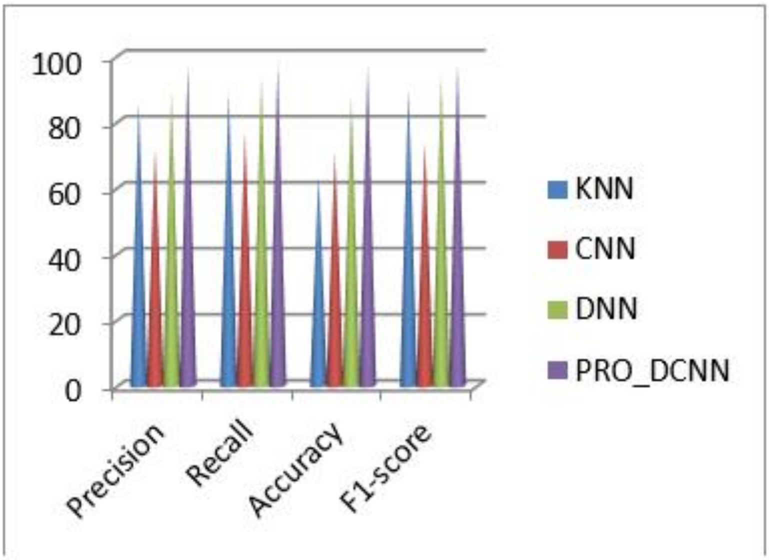

Figure 9: Comparison Of Existing Technique With Proposed For Fruit Disease Detection

The above fig.9 shows on the whole comparison for accuracy, precision, recall, F-1 score for future and existing techniques.

5. CONCLUSION

This disease detection and enhanced accuracy using by image processing techniques in identifying as well as classified the image of cashews plant. Here the simulation of classification and extracting the feature has been done using an image processing application. A manual system has been interchanged in detecting and classifying the disease of cashews plant.

Where the symptoms for consumption of time and the accuracy are low in comparison with the proposed technique. This proposed technique has been used by farmers in determining the disease with higher accuracy and detected earlier before spreading with nearby plants. Then the proposed system gives optimal classification using DCNN along with feature extraction using GLCM which could generate a greater quantity of yield. Hence two cashew leaves have been detected and future work can be carried out for the fruit and stem of the cashew plant with various disease detections.

6. FUTURE SCOPE

The results show that further work is needed in cashew plant disease detection studies, for instance, Convolution Neural networks (CNN), Recurrent Neural Networks (RNNs), and Deep Reinforcement Learning models. In the last few years, cashew disease has been increasing at extraordinary levels and the countermeasures used against this evolving threat have not proven effective despite their constant upgrade and revision. To prevent this threat of disease detection, more advanced disease detection technology is necessary because early-stage detection is better for disease spread avoided. The tools and resources are not sufficient in this research area. Hence, the researchers are in dire need to perform more research efforts to create an applications (APPs) technique in the plant disease detection field.

References

1) Coulibaly, Solemane,et al."Deep neural networks with transfer learning in millet crop images" Computers in Industry 108 (2019): 115-120.

2) Vijay Singh, Varsha, Prof. A K Misra.," Detection of unhealthy region of plant leaves using Image Processing and Genetic Algorithm "2015 ICACEA, IMS Engineering College, Ghaziabad, India.

3) Sachin D.Khirade, A.B.Patil," Plant Disease Detection Using Image Processing"2015 International Conference on Computing Communication Control and Automation.

4) Yogesh Dandawate, Radha K" An Automated Approach for Classification of Plant Diseases towards Development of Futuristic Decision Support System in Indian Perspective" 2015, ICACCI.

5) K. Seetharaman, A Fully Automated Crop Disease Monitoring and Management System Based on IoT: IoT-Based Disease Identification for Banana Leaf, Deep Learning Applications and Intelligent Decision Making in Engineering: IGI Global, Chapter 8, PP: 192-211, 2021.

6) Ramakrishnan. M, Sahaya Anselin Nisha. A," Groundnut Leaf Disease Detection and Classification by using Back Propagation Algorithm" IEEE ICCSP 2015 conference.

7) Rajleen Kaur, Dr. Sandeep Singh Kang," An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection" 2015 IEEE 3rd International Conference on MOOCs, IT Education.

8) T. Mahendran and K. Seetharaman "Banana Leaf Disease Detection Using GLCM Based Feature Extraction and Classification Using Deep Convoluted Neural Networks (DCNN), Journal of Positive School Psychology (2022), Vol. 6, No. 10, pp: 2553-2562.

9) Ms. Kiran R. Gavhale, Prof. Ujwalla Gawande, Mr. Kamal O. Hajari," Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques"2014 IC for Convergence of Technology.

10) Seetharaman, K., Mahendran, T. "Leaf Disease Detection in Banana Plant using Gabor Extraction and Region-Based Convolution Neural Network (RCNN)". *J. Inst Eng. India Ser.A* (2022).

11) https://link.springer.com/article/10.1007/s40030- 022-00-2

12) T. Mahendran., **K. Seetharaman** (2023). Feature Extraction and Classification Based on Pixel in Banana Fruit for Disease Detection Using Neural Networks. **IEEE Xplore ICAECT,** Bhilai, India. **DOI:** 10.1109/ICAECT57570.2023.10117959.