Arecanut Disease Classification using CNN

[1]Pallavi P, [2]Sowmya Nag K and [3]Veenadevi S V
[1,2,3,4]Department of Electronics and Communication, RV College of Engineering, Bengaluru, Karnataka, India.
[1]pallavip.ec19@rvce.edu.in, [2]sowmyanagk@rvce.edu.in, [3]veenadevi@rvce.edu.in

Article Info

Jenitta J and Swetha Rani L (eds.), *International Conference on VLSI, Communications and Computer Communication*, Advances in Intelligent Systems and Technologies,
Doi: https://doi.org/10.53759/aist/978-9914-9946-1-2_4

Abstract - Arecanut is a tropical crop, which is popularly known as betel nut. India ranks second in producing and consuming arecanut in the world. Throughout its life cycle, it is affected by a variety of diseases, from root to fruit. The current approach for detecting diseases is simply observation with the naked eye and farmers have to carefully analyse each and every crop periodically to detect the diseases. In this paper, a new system is proposed which helps in detecting the diseases of arecanut, leaves, and its trunk using Convolutional Neural Networks and suggests remedies for it. To train and test the CNN model, Dataset is created which consists of 200 images of arecanut both healthy and diseased. The train and test data are divided into a ratio of 70:30. For compilation of model categorical cross-entropy is used as loss function with adam as optimizer function and accuracy as metrics. A total of 50 Epochs are used to train the model to achieve high validation and test accuracy with minimum loss. The proposed approach was found to be effective and 81.35 percent accurate in identifying the arecanut disease.

Keywords - Arecanut disease, Machine learning, Convolution Neural Networks.

I. INTRODUCTION

Agriculture is the practice of growing crops and cultivating livestock in a piece of land. It is very important aspect of human life as it is agriculture that enables people to live and work in cities because of food that is produced because of agriculture. Agriculture is an age-old practice that started thousands of years ago and has evolved over time to make the process easier, efficient and productive. In India, agriculture is the largest provider of livelihood in rural areas and is one of the major contributors of the country's Gross Domestic Product (GDP). Areca belongs to the family of Arecaceae and is a genus comprising of 51 different species of palms. It is typically found in hot and humid areas such as in India, Philippines, Malaysia, etc. The name Areca is derived from a name which was locally used in Malabar coast of India. Even though arecanut crops are grown extensively in India, it is not a native crop of India. It is said to be first grown in areas of Philippines and Malaysia and hence is said to found in various varieties in those regions. But arecanut is a tropical nut that is grown in most parts of southern India as a cash crop.

India is traditionally an Areca growing country. There are various co-operatives and scheduled banks set up with an intention to support arecanut cultivation as well as for exporting them to foreign countries. The emergence of new products such as pan masala and gutka has increased the demand for arecanut cultivation. India is the largest producer as well as consumer of arecanut in the world. It constitutes about 88% of the production of arecanut in the world. The crop is mainly grown in the states of Maharashtra, Kerala, Karnataka and Andaman and Nicobar Islands. Arecanut cultivation is economically important and this can be learnt from the fact that roughly 56 million people are employed in this business either directly or indirectly.

The cultivation of arecanut requires various parameters to be considered and its growth is good only if areca plant receives optimal environmental condition. Some parameters to be kept in mind while growing areca plants are seed selection, land preparation, soil and climate requirements, irrigation, manure and fertilizer requirement, altitude, etc. Arecanut palm is said to be affected by several types of diseases. Around twenty types of diseases causing varying degrees of damages to the tree have been identified in India. However, in Uttara Kannada district the major diseases that have been reported are Yellow Leaf Disease, Mahali disease, Bud Rot, Stem Bleeding, etc. This reduces the quantity of crop production and hence the farmers would incur huge loss.

Machine Learning is the process of giving computers the ability of learn various features on data and make accurate decisions on unseen data. It is a branch of computer science and artificial intelligence that focuses on simulating human learning processes using data and algorithms to improve the system's accuracy over time. It primarily is based on creating algorithms that enable a computer to independently learn from data and previous experiences.

Machine Learning is an emerging field in crop yield and health analysis. Disease prediction is a very important issue in agriculture as farmers need to know if a crop is suffering from a disease beforehand so that they can treat it with appropriate measures. In the past, disease prediction was done by farmers who manually had to observe each and every plant and make predictions on if the particular crop is suffering from the disease based on personal knowledge.

Making judgments on the management of agricultural diseases requires accurate knowledge of the history of crop disease. This paper uses a variety of input data to create, build, and deploy the training model. Using data mining and data science techniques, the system will be able to recognize features and anticipate crop diseases. The objectives of the

project are the collection of data set having both healthy and diseased images of arecanut and their leaves, development of an algorithm that can detect arecanut disease and suggestion of solutions for the detected disease.

II. RELATED WORK

In [1] Dhanuja K C et al proposed a system for disease detection in arecanut using image processing technology. For training and testing the model a total of 144 images of arecanut samples were used which included 49 Good, 46 Poor and 49 negative samples. The main aim of this report was focused less on disease detection and more on grading different types of arecanut. K- Nearest Neighbour algorithm is used to detect diseases in arecanut. In [2] Ashish Nage, V. R Raut developed an android application that helps farmers in identifying plant disease by uploading a leaf image to the system which uses convolution neural network algorithm to identify presence of disease in leaf.

In [3] The proposed method consists of detection of diseases related to pomegranate and also suggesting solutions based on diseases. The proposed system consists of image pre-processing, segmentation, extraction of features and classification. In image pre-processing, images are resized. In segmentation, colour segmentation is carried out. Colour, morphology and texture features are used for the feature extraction. Minimum distance classifier is used for classification purposes. In [4], In this paper the authors have used different algorithms (SVM, KNN, Decision tree) for detection of diseases in leaves. This project works on uploading a captured image of sample to the system and algorithm will detect weather the sample is affected by any diseases or not, if it is affected by any disease, it will print the detected disease with an accuracy of 86%.

In [5] a dataset is created for training a neural network for image classification. The dataset was collected manually from the field. To get better feature extraction, image preprocessing was done, which includes noise removal, intensity normalization, removing reflections, and masking portions of the image. Then, using these processed images, a deep convolutional neural network model was trained to classify the images and TensorFlow library was used for numerical computations.

III. METHODOLOGY

The proposed method is shown in **Fig 1**. It consists of three phases: training phase, testing phase, implementation in GUI phase. In the training phase, the data was collected and pre-processed. In this step, the images were resized to a fixed resolution. The output labels were converted into one-hot encoding vector form. The images were further converted into arrays based on the color composition [6].

The pre-processed data was divided into training and testing data. The training data was used to train the convolutional neural network model to predict diseases in the image. In the testing phase, the disease was detected based on the features of the test image [7]. The work starts with the pre-processing step. In this step the collected data was pre-processed to get uniform images so that training the model is made easy [8-12].

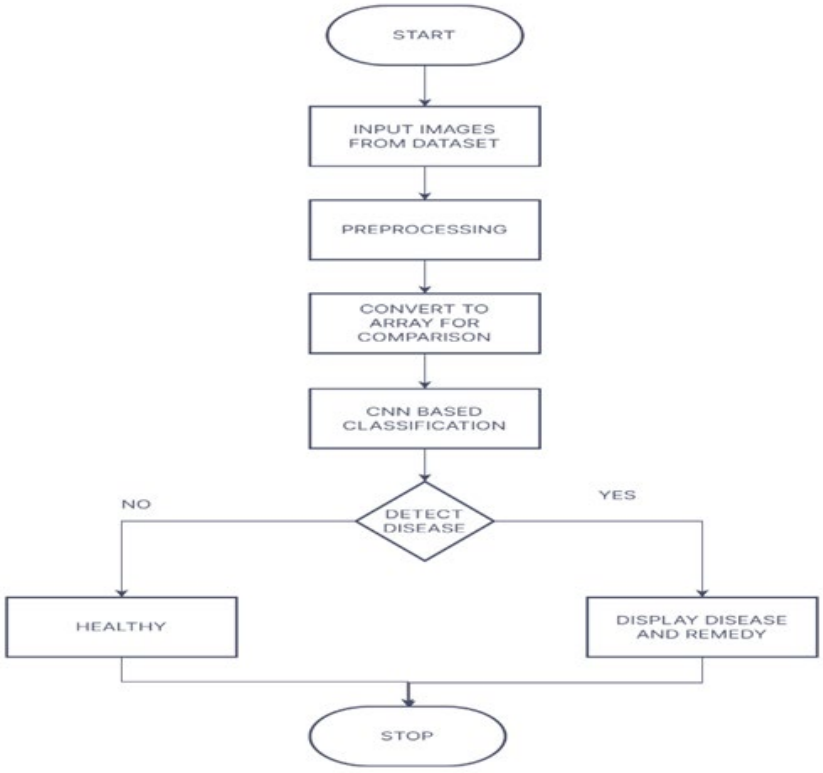

Fig 1. Arecanut Disease Classification

Deep Learning is a part of machine learning process which uses the concept of neural networks present in the brain to mimic the learning process of a human. Deep learning can analyse images in ways that machine learning cannot do easily. Moreover, deep learning could easily learn features of a dataset without much human intervention when compared to machine learning. CNN is a Deep Learning algorithm that can take in an image as input, assign importance (learnable weights and biases) to various aspects/objects in the image then learn from the results and be able to differentiate one from the other. As compared to other classification algorithms, the amount of pre-processing needed by a ConvNet is significantly less. Though filters in primitive methods are hand-engineered, ConvNets can learn these filters/characteristics with enough preparation [13]. In the convolution layer, which is the main layer, the images are divided into partial images based on set step size where a step length is used to estimate the number of pixels that must be taken into consideration before performing calculations. This step is mainly used to reduce the dimensionality of the image.

After the convolutional layer, the next step is to send the data from convolutional layer into the pooling layer. The process that happens in the pooling layer is same as that happens in the convolutional layer. The only difference is that we don't take the whole partial array of pixels as the result. But we take the maximum or average value as the result. This in turn is used to preserve even the smallest of features that are inherent in the pixels of partial images [14].

The last type of layer is the fully connected layer which is similar to the regular neural networks. These are tightly meshed layers which are used to compute and forward every bit of output data from the pooling layer as the images are already greatly reduced in size.

An open-source Python framework called Streamlit is used to create web applications for machine learning and data science. Streamlit makes it simple to launch web apps quickly after they have been developed. It is possible to create apps with Streamlit with good understanding of python code. Working on the interactive cycle of coding and watching outcomes on the web app is made simple by Streamlit [15].

Dataset

To detect various diseases of arecanut dataset was created which consists of healthy and diseased images of arecanut and their leaves. The images shown in **Fig 2** were taken from a digital camera at a half-meter distance from the source. Some of the diseased and healthy arecanut images were collected from Kundapur, Shimoga district of Karnataka. These photographs were taken under the guidance of experienced arecanut researchers and farmers. Others were collected from google. The dataset consists of a total of 200 images which includes leaves, nuts, and trunk of arecanut both healthy and diseased. There are 65 healthy and 135 diseased images of arecanut such as yellow, yellow spot, Mahali/Koleroga, and Stem bleeding disease. The images are resized into 256*256 pixels using open-cv before training the model.

Fig 2. Healthy and Diseased Arecanut and Areca Plant Images.

Preprocessing

The database is pre-processed, which includes image reshaping, resizing, and array conversion. The test image is also subjected to similar processing. Images are resized to 256*256 resolutions and converted to an array before training the

CNN Model

Computers are unable to recognize or evaluate images in the same way that humans do. So, the captured images are converted into array using Numpy. The array contains RGB values of each pixel of an image ranging from 0 to 255.

Model Structure

CNN has several layers, including Dropout, Convolution2D, Activation, Dense, MaxPooling2D and Flatten. For training the model using CNN 1000 neurons were used in the first layer, 500 in the second layer, 250 in third and 5 in the last dense layer. The activation function used in the first 3 layers is relu and for the last layer SoftMax. A total of 123,677,025

parameters are calculated which includes the weights and biases. The last has a SoftMax activation function which gives the probability of detected disease. A total of 200 images which includes healthy and diseased images are used to train and test the Model. The dataset is limited so, Augmentation technique is used which basically performs rotation, shifting, zooming, flipping the image to create new data for training. The train and test data are divided into a ratio of 70:30. For compilation of model categorical cross entropy is used as loss function with adam as optimizer function and accuracy as metrics. A total of 50 Epochs are used to train the model to achieve high validation and test accuracy with minimum loss.

IV. RESULTS

After training the model the test accuracy observed was 81.35 %. The graphs of accuracy versus epochs and loss versus epochs are shown below **Fig 3** and **Fig 4** respectively. The image of the plant was given as input to the model trained using CNN. The trained model detects diseases in arecanut and prints the probability of the detected disease. Also, the remedy for the maximum probability disease is shown for the user reference.

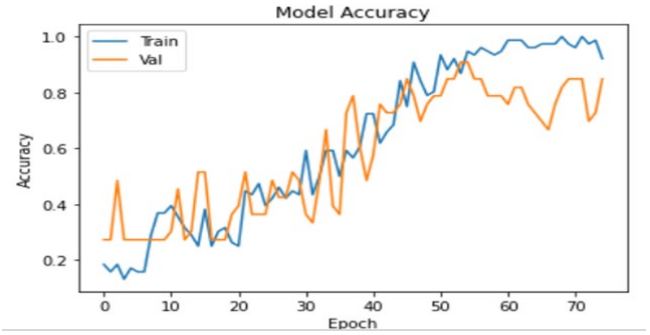

Fig 3. Accuracy Versus Epochs

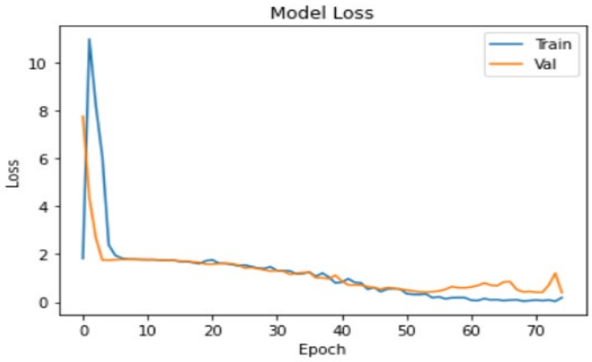

Fig 4. Loss Versus Epochs

Streamlit app allows the user to input an image and see the predicted result in a webpage. The **Fig 5** shows the detection of healthy nut and detection of yellow-leaf-spot disease and even the remedy for the disease is shown in **Fig 6**.

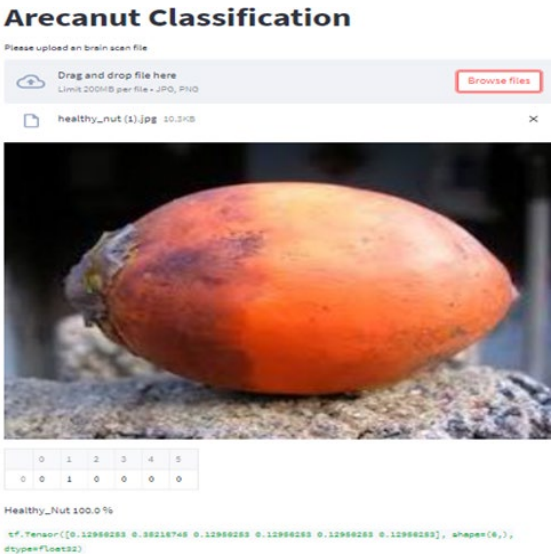

Fig 5. Detection of Healthy Nut.

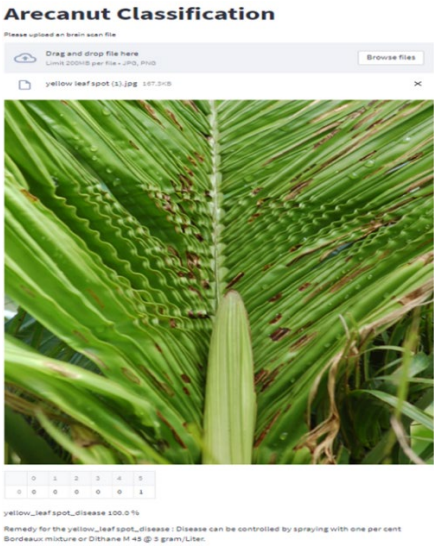

Fig 6. Detection of Yellow-Leaf-Spot Disease and Remedy Suggestion.

Fig 7 and **Fig 8** shows the detection of healthy trunk and the stem bleeding disease. For the stem bleeding disease even, the suitable remedy is suggested.

Fig 7. Detection of Healthy Trunk

Fig 8. Detection of Stem bleeding disease

V. CONCLUSION

This paper focuses on the early detection of diseases in Arecanut, leaves, and trunk using Convolutional Neural Networks. Experimentation is conducted using diseased and healthy arecanut image dataset of 200 images. The input image is first pre-processed, followed by feature extraction, training, and classification. The proposed System detects diseases of arecanut such as Mahali, Stem bleeding, and yellow leaf spot and provides remedies for the same. Depending on the quality of the input image and the stage of the disease, the experimental results show varying levels of disease detection accuracy. The overall accuracy of the system is estimated to be 81.35 percent. As a result, this system takes a step toward encouraging farmers to practice smart farming and allowing them to make better yield decisions by enabling them to take all the necessary preventive and corrective action on their arecanut crop.

ACKNOWLEDGMENT

I would like to acknowledge my mentors Mrs Divya T. L, Mrs. Sowmya Nag K and Dr. Veenadevi S V for providing me an opportunity to carry out this project under Cisco- RVCE Centre of Excellence. The authors thank the management of RV College of Engineering for providing all the support and motivating us to conduct the research.

References

[1]. M. Manjunatha and A. Parkavi, "Estimation of Arecanut Yield in Various Climatic Zones of Karnataka using Data Mining Technique: A Survey," 2018 International Conference on Current Trends towards Converging Technologies,doi: 10.1109/ICCTCT.2018.8551083,pp. 1-4,2018,

[2]. A. S and A. Hegde, "Detection and classification of areca nut diseases," 2021 Second International Conference on Electronics and Sustainable Communication Systems.doi: 10.1109/ICESC51422.2021.9532754,pp. 1092-1097,2021.

[3]. Gupta, Prakash. (2007). Areca nut use in India. Indian journal of medical sciences. 61. 317-9. 10.4103/0019-5359.32674.

[4]. Dhanuja K C. "Areca Nut Disease Detection using Image Processing Technology". International Journal of Engineering Research, 10.17577/IJERTV9IS080352,2020

[5]. Manisha Bhange, H.A. Hingoliwala, "Smart Farming: Pomegranate Disease Detection Using Image Processing", Procedia Computer Science, Volume 58,ISSN 1877-0509, Pp 280-288,2015

[6]. Anandhakrishnan MG Joel Hanson, Annette Joy, Jeri Francis, "Plant Leaf Disease Detection using Deep Learning and Convolutional Neural Network", International Journal of Engineering Science and Computing, Volume 7, Issue No.3, March 2017

[7]. Ashish Nage, Prof. V. R. Raut, "Detection and Identification of Plant Leaf Diseases based on Python",International Journal Of Engineering Research & Technology,Volume 08, Issue 05,2019

[8]. Suresh M, Ajit Danti and S. K Narasimhamurthy, "Classification of diseased areca nut based on texture features," International Journal of Computer Applications (0975 – 8887) Recent Advances in Information, 2014.

[9]. Kuo-Yi Huang, "Detection and classification of areca nuts with machine vision," Computers and Mathematics with Applications 64 739–746, 2012.

[10]. Bhavini J. Samajpati and Sheshang D. Degadwala ,"Approach for apple fruit diseases detection and classification using random forest classifier," International Conference on Communication and Signal P rocessing, April 6-8, 2016, India, 2016.

[11]. Pushparani M.K., Dr. D Vinod Kumar and Dr. Abdulla Gubbi, "Arecanut grade analysis using image processing techniques," International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, 2019.

[12]. Rahamathunnisa U., Nallakaruppan M.K, Anith.A and K.S.Sendhil Kumar, "Vegetable disease detection using k-means clustering and svm," 6 th international conference on advanced computing and communication system, 2020.

[13]. S Siddesha and S K Niranjan, "Detection of Affected egions of Disease Arecanut Using K-Means and Otsu Method," International journal of scientific & technology research volume 9, issue 02, 2020.

[14]. A S M Farhan Al Haque, Rubaiya Hafiz, Md. Azizul Hakim and G. M. Rasiqul Islam, "A computer vision system for guava disease detection and recommend curative solution using deep learning," 22nd International Conference on Computer and Information Technology(ICCIT),2019.

[15]. Swathy Ann Sam, Siya Elizebeth Varghese, Pooja Murali, Sonu Joseph John, Dr. Anju Pratap,"Time saving malady expert system in plant leaf using CNN",Volume 13, Issue No 3,2020.