

Classification of Citrus Leaf Diseases Using Convolutional Neural Networks Compared to Traditional Neural Networks

B. K. Dabhi¹, D. K. Parmar², N. D. Patel³, X. U. Shukla⁴

¹ Laboratory Technician, SMC College of Dairy Science, KAU, Anand - 388110

² Associate Professor, College of Agricultural Information Technology, AAU, Anand - 388110

³ Assistant Professor, B. A. College of Agriculture, AAU, Anand - 388110

⁴ Assistant Professor, College of Agricultural Information Technology, AAU, Anand - 388110

Email: bhumika82aau@gmail.com

ABSTRACT

India is an agrarian country with a diverse agricultural sector, covering approximately 157 million hectares (387.9 million acres) of agricultural land. It is considered the second largest in the world after the United States. The main objective of this research is to compare different algorithms for segmenting disease lesions, collecting shape-related features, and classifying these features for detecting disease spots (*Pseudocercospora angolensis*) and determining the severity of the disease. In k-means clustering, the value of k signifies the total number of clusters, and each cluster centroid defines the center of the region. Centroids are initially selected at various locations, and each point in the dataset is connected to the nearest centroid using the squared Euclidean distance algorithm. Centroids are then rearranged according to the average of their data points. This process is repeated until the centroids no longer change position, thus finding the optimal positions for the centroids. The k-means clustering segmentation technique divides an image into k-means clusters, grouping areas with similar properties together and separating them from other regions of the image. In machine learning, pattern recognition, and image processing, feature extraction starts from an initial set of measured data and builds derived values (features) that are informative and non-redundant. This facilitates subsequent learning and generalization steps and can lead to better human interpretations. Humans can better understand a reduced feature set compared to the whole image. Artificial Neural Networks (ANNs) are models that resemble biological neuronal structures. The starting point for any ANN is an elementary neural element that mimics the behaviour of real neurons. Each computing unit in an ANN is based on the concept of an idealized neuron, which is expected to respond optimally to applied inputs. A neural network is a collective set of such neural units, where individual neurons are connected through complex synaptic associations represented by weight coefficients. Each neuron contributes to the computational properties of the entire system. Convolutional Neural Networks (CNNs) can successfully capture spatial and temporal dependencies in images using appropriate filters. Due to the reduced number of parameters and the reusability of weights, CNNs are well-suited for image datasets. The network can be trained to better recognize the complexities of images. This paper highlights the strength of CNNs over traditional neural networks (NNs). CNNs achieved a multi-class classification accuracy of almost 97%, whereas traditional NNs achieved a binary classification accuracy of about 94% using the trainbr training function. Therefore, it is evident that CNNs and traditional NNs present similar results, with CNNs having a slight edge in performance.

Keywords : pixel identification, processing of image, ANN, CNN, Tensorflow, Keras

INTRODUCTION

Citrus is the most common fruit crop in India and at the world level also. Most of the fruits are citric in nature. Major of citrus growers face many problems to harvest their fruits as they attacked by virus, bacteria, worm, and fungi. Besides, when the citrus crop had been infected or attacked, the other areas had been exposed to be infected. Resulting it will decrease farmer's income and lead to consequence losses to a farmer. Currently, the citrus

farmer determines the type and intensity of disease manually. The errors might occur in order to determine the type of diseases. Automated recognition of plant disease with the help of image processing techniques gives more precise and vigorous guidance for disease management than the traditional technique of visual detection by human raters. Moderately, visual identification is lesser precise and time consuming for disease discovery. It is also similarly significant to accurately measure the volume of

pesticides to be sprayed to hinder the progression of fungal attack. This research assumed to enumerate the diseased area affected by the citrus leaf brown spots using different techniques that employs image processing, which was followed by determination of the rate of development and discovering the bacterial or fungal attack at an early time on citrus yield. [16]

In India Citrus farming faces severe threats from pests and diseases. Major of the diseases are rust, downy mildew, powdery mildew, canker, greasy spot and anthracnose whereas main insect pests are mealy bugs, flea beetle, thrips, hoppers, stem borers and mites. Table 2.3 presents symptoms, effects and pictorial notation of the diseases.

Symptoms	Effects	Image
Disease: Citrus Canker		
Scratches on the leaves, stems, and fruits.	- Causing leaves and fruit to drop prematurely - Fruit is too unsightly to be sold.	
Disease: Citrus Anthracnose		
- Small brown/ gray spot on shoots, flower, berries - Blights on petioles and veins - Short holes in leaves - Fruits may crack	- Reduces the effective area of a leaf - Affected Blossom fails to set fruit	
Disease: Citrus Blight		
- Wilting, die-back, loss of foliage, and the production of smaller/discolored fruits - Dim yellow spots on upper side of young leaves	- Decline in health - Cannot support bunch development - Rootstock develop concurrent infections	

Table 1: Effects of Citrus Diseases [17]

Anthracnose is a fungus that attacks weak branches. The disease is especially common in springs that are prolonged rains and when heavy rains fall later in the season than

usual. Anthracnose spores leak on fruit in wet or foggy conditions, infect the skin and leave it pale, with reddish-green streaks on immature fruit and brownish-black streaks (tear stains) on mature fruit. Blight is a disease that causes citrus trees to fall and die. The disease does not affect non-bearing trees, and trees may show signs of decay by age 8-10. [11] However, it is found on trees as small as four years. Mild bending and graying in the canopy, as well as small leaf loss are early signs of disease. More severe signs such as persistent bending, leaf fall, and twig dieback follow. On damaged plants, root and stem branches are prevalent. Citrus canker is a disease most often caused by scars on the leaves and bark of the fruit, but when there are ideal conditions for infection, it can also cause foliage, twig death and fruit loss. Citrus canker is characterized by brown spots on the leaves, often with an oily or watery appearance. [7] The lesions (or spots) are often surrounded by a yellow halo and may appear on both the upper and lower sides of the leaf. Fruits and stalks may show similar signs.

Artificial Neural Network (or ANN), is a group where we have multiple perceptrons or neurons at each layer and is also known as a Feed-Forward Neural network because inputs are processed only in the forward direction. This type of neural networks are one of the simplest variants of neural networks as They pass information in one direction, through various input nodes, until it makes it to the output node. This type of neural network may or may not have hidden node layers, making their functioning. [3] While Convolutional Neural Network (CNN) are one of the most popular models used today. This type of neural network computational model uses a variation of multilayer perceptrons and contains one or more convolutional layers that can be either entirely connected or pooled. Further, these convolutional layers create feature maps that record a region of the image which is ultimately broken into rectangles and sent out for nonlinear processing. [12] Some of the advantages of convolutional Neural Network is that they offer very high accuracy in image recognition problems and are capable of automatically detecting important features without any human supervision.

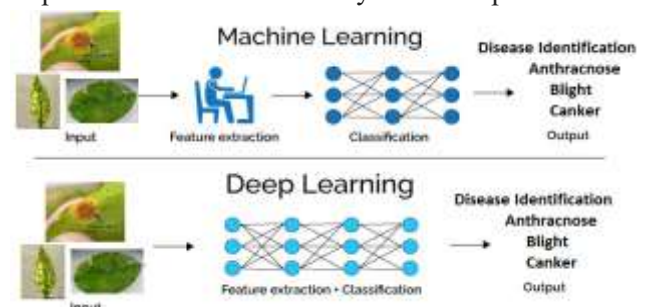


Fig.1: Deep Learning over Machin Learning

MATERIALS AND METHODS

All the images are composed in RGB, 36 bit and are of .jpg format. First of all, the diseased image of leaf taken and set the background as black with sufficient illumination condition. This is to eliminate any reflection in the environment and get consistently distributed light in all places in order to improve the view and brightness of the image to be processed. [4]

Now it is too trivial to calculate the area of a leaf image. For that using the fundamentals of RGB to HSI image conversion the RGB image converted to HIS after that gray level threshold applied and the resulting gray image is finally to be converted into binary. See the image 3.6.

	
Figure 2(a): Leaf Area - Original Image	Figure 2(b): Leaf Area - Binary Image
Total Pixels = 1920000 Background Pixels= 943617 Whole Leaf Pixels= 976383	

Now here in the second image a one rupee coin put as a reference image for calculating the leaf area in centimetre. [6]

Shape related features are used to represent shape, size and boundary of objects of the image as described in previous sections. The feature values are 2870, 202.30, 66.59, 55.93, 0.54, 21.41, 60.44, 0.97, 0.72 for Area, Perimeter, MajorAxisLength, MinorAxisLength, Eccentricity, Orientation, EquivDiameter, Solidity and Extent respectively.

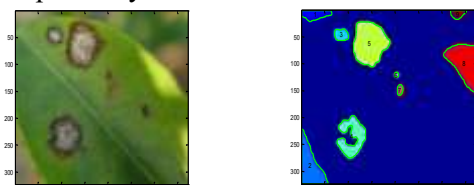


Fig. 3: Disease identified by number to extract features

K--Means Clustering Algorithm

In K - means that the segmentation the motto is to group similar things/matters/data (here pixels) together. A data matrix $M=[m_{(1,)} m_{(2,)} m_{(3,....)} m_{(n)}]$ and k initial segmenting centroids, k-means clustering aims to partition each m_i into k number of clusters, where m_i fits to its adjoining cluster based on the algorithm and process of squared Euclidian distance. For k—means the value k

signifies the total number of clusters and for each cluster centroid defines the central part – center of the region. [16] This function focuses on the identification and recognition of citrus leaf diseases using different classifiers. [13]

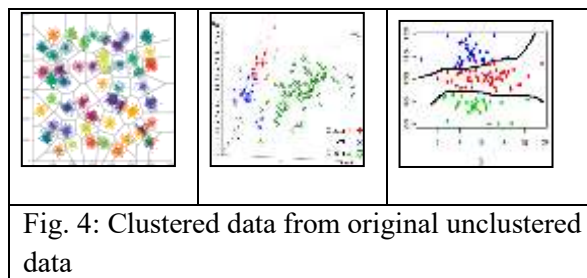


Fig. 4: Clustered data from original unclustered data

K—Means applied for clustering separation of images. This method addresses the early belief of the disease. Farmers can automatically identify disease in early stages of leaves using this method. Here the flow chart in figure 3.19 presents the overall image processing procedure along with image enhancement, image segmentation, feature extraction and completes with image classification. [8]

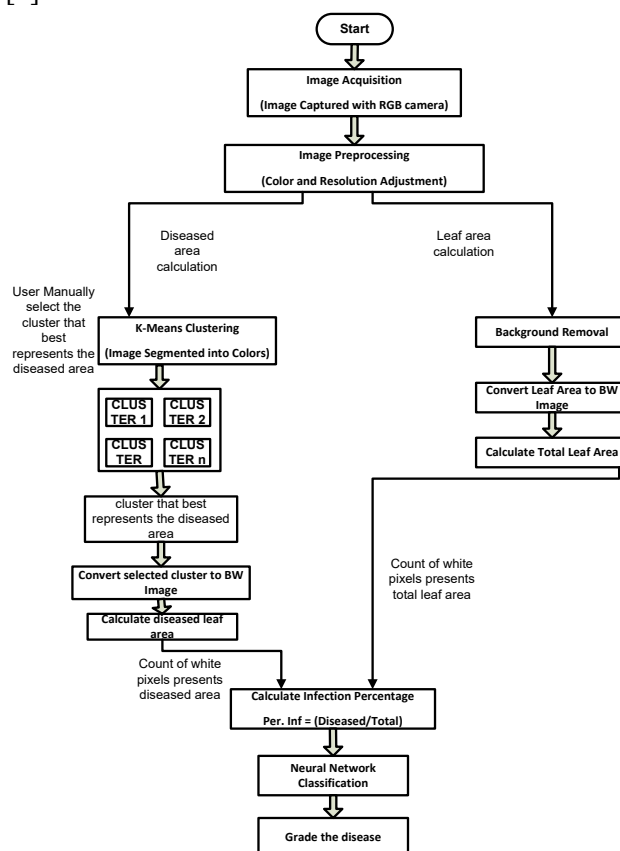


Fig.5: Flow chart for K-Means Algorithm

RESULTS AND DISCUSSION

The present approach steps forward to achieving plant disease identification and prevention. The proposed plant disease segmentation technique is implemented in the working platform MATLAB (version 2011-13) and it is evaluated using almost 100 leaf images, which are

collected from different sources as discussed earlier. [12] The sample image considered in this work and their respective outputs after RGB colour model conversion, colour sharpening, filtering, enhancing and K-means Clustering Segmentation, this segmentation process allows for an image to be more precisely segmented. Here when the clustering algorithm runs, always it presents the different cluster as the diseased cluster because all the time it applies the clustering logic on the leaf image and the algorithm cluster out data into three clusters in this case and then it display accordingly. Now sometimes it may present it as cluster no. 1, then very next moment it may present as cluster no. 2 or 3. But the cluster picture will, be the same that can be seen in the following figure. [11]

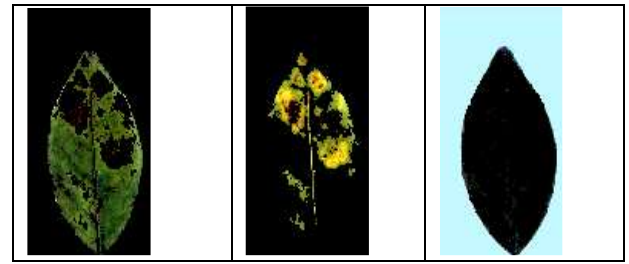


Fig. 6: Different segments of the diseased leaf shape-related features are total in nine so the training matrix will be of 200 samples of 9 elements. Now for testing purpose, a matrix was already prepared that classify two diseases and that is of 200 samples of 2 elements since only two disease classes wanted to be classified. [12] Here the below table 4.11 presents the samples table used for training and testing purpose.

Feature	Sample 1	Sampl e2	Sampl e3	Sampl e4	Sampl e5	Sample 1	Sampl e2	Sampl e3	Sampl e4	Sampl e5
Area	16018.00	4674.00	8882.00	9566.00	8061.00	15302.00	273.00	16563.00	15448.00	15709.00
Perimeter	972.47	369.91	431.35	490.80	391.56	814.17	70.49	577.09	956.12	838.78
MajorAxisLength	172.66	116.97	118.36	121.72	108.01	176.45	24.54	174.34	245.50	161.67
MinorAxisLength	126.98	57.54	99.15	101.54	96.45	120.61	15.39	125.52	105.10	135.59
Eccentricity	0.68	0.87	0.55	0.55	0.45	0.73	0.78	0.69	0.90	0.54
Orientation	57.89	39.34	34.92	9.02	71.86	-12.38	14.95	34.69	27.38	-74.76
EquivDiameter	142.81	77.14	106.34	110.36	101.31	139.58	18.64	145.22	140.25	141.43
Solidity	0.81	0.82	0.91	0.94	0.95	0.87	0.92	0.91	0.67	0.90
Extent	0.65	0.59	0.59	0.76	0.78	0.79	0.76	0.59	0.50	0.90
CLASS	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00
CLASS - 1 = ANTHRACNOSE						CLASS - 2 = BLIGHT				

Table 2: Training and Testing matrix for shape related features

The feature matrix and the characteristic of the data as follows and given below. Many computer vision algorithms are based on local features and their descriptions. Some of their uses are image registration, object detection and classification, tracking, motion estimation and content-based image retrieval (CBIR). Spatial characteristics are used in these methods to better manage scale changes, rotation and obstruction.

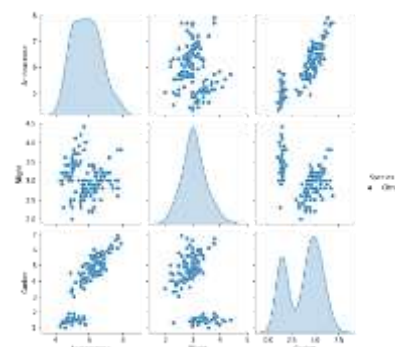


Fig.7: Characteristic and Attributes of the Feature data

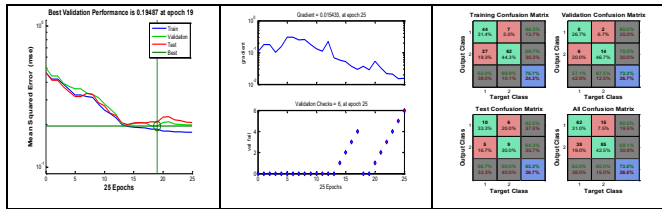


Fig. 8: NN Performance – MSE, Training state and Confusion Matrix

Consider ‘trainbr’ function, it takes very large time that is almost 35-38 seconds but gives very good results that is almost 93-96%. So in nutshell the ‘trainscg’ training function gives higher classification rate i.e. 80% within 2-3 seconds.

For the convolutional neural network, follow the results and given ion the table below. The CNN contains Sequential model with three dense layers with the neurons were 1000, 500 and 300 respectively.

TRAINING FUNCTION = “TRAINSCG” → SCALED CONJUGATE GRADIENT									
Seg. Algo.	Class-I	Class-II	Number of neurons	Train/ Retrain	Class-I Correct (%)	Class-II Correct (%)	Overall Classifi. (%)	Better for	Time
2	Anth.	Blight	10	Train	22.9	32.7	55.6	Blight	2.944
2	Anth.	Blight	10	ReTrain	34.6	35.6	70.2	Blight	2.745
2	Anth.	Blight	20	Train	30.7	39.5	70.2	Blight	3.214
2	Blight	Canker	10	Train	18	29.3	47.3	Canker	2.590
2	Blight	Canker	10	ReTrain	48.8	0.0	48.8	Blight	2.487
2	Blight	Canker	20	Train	30.7	35.1	65.9	Canker	2.635
2	Anth.	Canker	10	Train	11.4	44.1	55.4	Canker	2.610
2	Anth.	Canker	10	ReTrain	20.3	42.1	64	Canker	2.354
2	Anth.	Canker	20	Train	34.7	30.2	64.9	Anth.	2.715
TRAINING FUNCTION = “TRAINGD” → GRADIENT DESCENT									
2	Anth.	Blight	10	Train	16.1	40.0	56.1	Blight	3.60
2	Blight	Canker	10	Train	25.1	34.3	59.4	Canker	4.43
2	Anth.	Canker	10	Train	32.7	20.3	53.0	Anth.	5.95
TRAINING FUNCTION = “TRAINBR” → BAYESIAN REGULARIZATION									
2	Anth.	Blight	10	Train	44.9	47.3	92.2	Blight	35.61
2	Blight	Canker	10	Train	29.0	34.8	63.8	Canker	35.15
2	Anth.	Canker	10	Train	46.5	45.5	92.1	~	29.00

Table 3: Result of Neural Network with different Trainig Fcuntions

Here one dropout function was taken with 0.2. [6] The activation function were changed and shown here.

Model	Layer	Layer	Layer	Activation	Epochs	Accuracy
sequential	Dense_1	Dense_2	Dense_3	Relu, Softmax	30	96-97%
sequential	Dense_1	Dense_2	Dense_3	Relu, Sigmoid	30	97-98%
sequential	Dense_1	Dense_2	Dense_3	Selu, Sigmoid	30	95-96%
sequential	Dense_1	Dense_2	Dense_3	Selu, Sigmoid	30	98-99%

Table 4: Classification Accuracy for CNN

CONCLUSION

Human being totally depends upon nature and plants; therefore, it must be special methods to protect plants from diseases. Reduction in crop production also affects the country's economy. Citrus Industry is currently vulnerable by three major citrus diseases; citrus anthracnose, blight, and canker. By very less computational efforts the most favourable results were obtained, which also shows the effectiveness of the stated algorithm in the identification and organization of the leaf diseases. Citrus leaves were Segmented with K-Means Clustering Algorithm. The most difficult and challenging part was the citrus disease image segmentation which was further divided into two parts: the background portion localization and extraction, and then the diseased portion separation. 9 shape related

features were extracted and evaluated in various experiments. With trainbr training function for the classes of Anthracnose and Blight the classification neural network gave the highest classification rate and that was of almost 92%. The classification rate was good and almost all the leaves were classified properly and correctly. While we have applied the convolutional neural network, we received the classification rate was between 95% to 98% and it was a very great achievement. So here it is immediately known that when a traditional neural network is used it works on whatever features are extracted from the image and on the basis of such features it also generates the result which is going to be around 92%. But when Convolutional Neural Network used it uses the whole image without using features and tries to classify by analysing the whole image and that is why Convolutional Neural Network gives us about 95% to 98% accurate result which this research is proved by.

REFERENCES

- [1] Huete A.R., 1988, A soil adjusted vegetation index (SAVI). Remote Sensing of Environment, Vol.25, pp.121-136.
- [2] J Dheeba, NA Singh, ST Selvi. Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. Journal of biomedical informatics 49, 45-52
- [3] J. Xing, 2005. Detecting bruises on 'Golden Delicious' apples using hyperspectral imaging with multiple wavebands. Biosystems Engineering, vol. 90, (1), pp. 27-36.
- [4] James W. Olmstead and Gregory A., 2001. Lang. Assessment of Severity of Powdery Mildew Infection of Sweet Cherry Leaves by Digital Image Analysis. Hortscience 36(1):107–111.
- [5] Jing Hu, Daoliang Li, Guifen Chen, Qingling Duan & Yeqi Han, 2012. Image Segmentation Method for Crop Nutrient Deficiency Based on Fuzzy C-Means Clustering Algorithm, Intelligent Automation & Soft Computing, 18:8, 1145-1155.
- [6] Muchoney D.M., and Haack B.N., 1994, Change detection for monitoring forest Defoliation, Photogrammetric Engineering and Remote Sensing, Vol.60, pp.1243-1251.
- [7] Muhammad Asraf Hairuddin, Nooritawati Md Tahir and Shah Rizam Shah Baki., 2011. Overview of Image Processing Approach for Nutrient Deficiencies Detection in Elaeis Guineensis. IEEE International Conference on System Engineering and Technology (ICSET).
- [8] Nelder, J.A.1994. The statistics of linear models: back to the basics. Statistics and Computing 4: 221-234.
- [9] Otsu N. 1979. A Threshold Selection Method from Gray-level Histograms. Institution of Electrical and Electronics Engineers (IEEE) Transactions on Systems, Man and Cybernetics, 9(1):62-66.
- [10] Sachin Jagtap, Shailesh Hambarde, 2014. Agricultural Plant Leaf Disease Detection and Diagnosis Using Image Processing Based on Morphological Feature Extraction. IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 4, Issue 5, Ver. I (Sep-Oct. 2014), PP 24-30. e-ISSN: 2319 – 4200
- [11] Sanjay Dhaygude, Nitin Kumbhar, 2013. Agricultural plant Leaf Disease Detection Using Image Processing. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, ISSN: 2278 – 8875, Issue 1.
- [12] Schubert, T. S., Miller, J. W., 2000. Bacterial citrus canker. FDACS, Division of plant industry, Gainesville, Florida. 6 fold.
- [13] Yaming Suny, Lei Liny, Duyu Tangy, Nan Yangz, Zhenzhou Jiy, XiaolongWangy, 2015. Modeling Mention, Context and Entity with Neural Networks for Entity Disambiguation. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence.
- [14] Yeh, Y.H., W.C. Chung, J.Y. Liao, C.L. Chung, Y.F. Kuo, T.T. Lin, 2016. Strawberry foliar anthracnose assessment by hyperspectral imaging. Computers and Electronics in Agriculture 122: 1-9.
- [15] Zaller, J. G. 2004. Ecology and non-chemical control of Rumex crispus and R. obtusifolius (Polygonaceae): a review. Weed Research, 44(6), 414–432.
- [16] Bock CH, Cook AZ, Parker PE, Gottwald TR., 2009. Automated image analysis of the severity of foliar citrus canker symptoms. Plant Dis. 93(6):660–665. doi: 10.1094/PDIS-93-6-0660
- [17] G. S. Reddy, V. D. Murti, 1985. Book of Citrus Diseases and Their Control, Indian Council of Agricultural Research, New Delhi