

TOMATO LEAF DISEASES PREDICTION

Mahesh.S
Dept. of ISE
Vemana Institute of Technology
Bengaluru, India
vi18is053@vemanait.edu.in
Nagesh Sugur K S
Dept. of ISE
Vemana Institute of Technology
Bengaluru, India
vi18is060@vemanait.edu.in
Vishwanath Gowda R
Dept. of ISE
Vemana Institute of Technology
Bengaluru, India
vi18is124@vemanait.edu.in



Publication History

Manuscript Reference No: IRJCS/RS/Vol.09/Issue08/SPAUCS10106

Research Article | Open Access

Peer-review: Double-blind Peer-reviewed

Article ID: IRJCS/RS/Vol.09/Issue08/SPAUCS10106

Volume 2022 | Article ID SPAUCS10106 <http://www.irjcs.com/volumes/Vol09/iss-08/26.SPAUCS10106.pdf>

Received: 25, July 2022 | Accepted: 04, August 2022 | Published: 12, August 2022

doi: <https://doi.org/10.26562/irjcs.2022.v0908.26>

Citation: Mahesh,Nagesh,Vihwanath(2022).Tomato Leaf Diseases Prediction. International Research Journal of Computer Science (IRJCS), Volume IX, Issue VIII, 2022, Pgs 286-293

BibTeX key:

Academic Editor-Chief: Dr.A.Arul Lawrence Selvakumar

Copyright: ©2022 This is an open access article distributed under the terms of the Creative Commons Attribution License; which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Abstract: This paper deals with the optimized real time detection of diseases that affect the plant and the area affected using Convolutional Neural Networks (CNN) algorithms. So that appropriate fertilizers can be used to prevent further damage to plants from pathogenic viruses. The activation function is the core of the CNN model as it incorporates the non – linearity to have an authentic artificial intelligence system for classification. ReLu is one among the best activation functions, but has a disadvantage that the derivative of the function is zero for negative values and leads to neuronal necrosis. New mathematical activation function is developed and compared with existing activation functions to improve the accuracy and performance of the system on a Tensor Flow framework. The experimental results on trained databases show that the developed activation function has improved the CNN model accuracy and performance i.e. 95%. The training speed of the CNN model is improved by 83% when implemented in ARM processor using the proposed optimizer. Further area affected by disease is calculated by using K-means clustering algorithm for optimization of fertilizer usage.

Keywords: Convolutional Neural Networks (CNN), pooling layers, fully connected dense layers, activation function

I. INTRODUCTION

Indian is one of the fastest growing developing nations of the world, where 58% of livelihood depends on agriculture and rural households [1,12]. Even though production is huge India is lagging behind, as the existing literature does not provide any alternative technique which can deal with the exact identification of the crop disease [3,10]. The common ways of crop infections are due to pathogenic organisms and the existing works does not provide the fertilizer optimization to improve the crop quality. This paper implements an image processing approach using Convolutional Neural Networks (CNN) Algorithm implemented in Raspberry Pi kit. The CNN model learns the filters whereas traditional algorithms used various activation functions to train and classify the output. The design of a new feature activation function with apt optimizer is the major advantage of this application. The aim of this project is early prediction of crop disease with greater accuracy and prevention of further damage done to the crops. The area of the disease affected is also found so that fertilizers application can be optimized.

II. LITERATURE REVIEW

In [1,6] the authors proposed a neural network based approach to detect the disease. This paper deals with genetic algorithm the optimization of loss function using genetic algorithm so that the strong parameters can only survive based on natural selection.

In [2,8] the authors proposed a semi-supervised learning in the form of Generative Adversarial Networks to classify images and is used only for training the discriminator. In [3,11] the authors have proposed a system which follows Euclidean and distance clustering technique is used to find the area of leaf, disease affected parts and classify the leaf image. In [4,7] the authors proposed a system for identification and classification of diseases in cotton leaf based on pattern recognition using image segmentation and adaptive neuro – fuzzy inference system. In [5, 9] the authors proposed a system to segment the defects by using clustering algorithm. The image was segmented as number of clusters out of those one or more clusters were found to have only infected areas. In this research, squared Euclidean distance was utilized for the K-means clustering scheme.

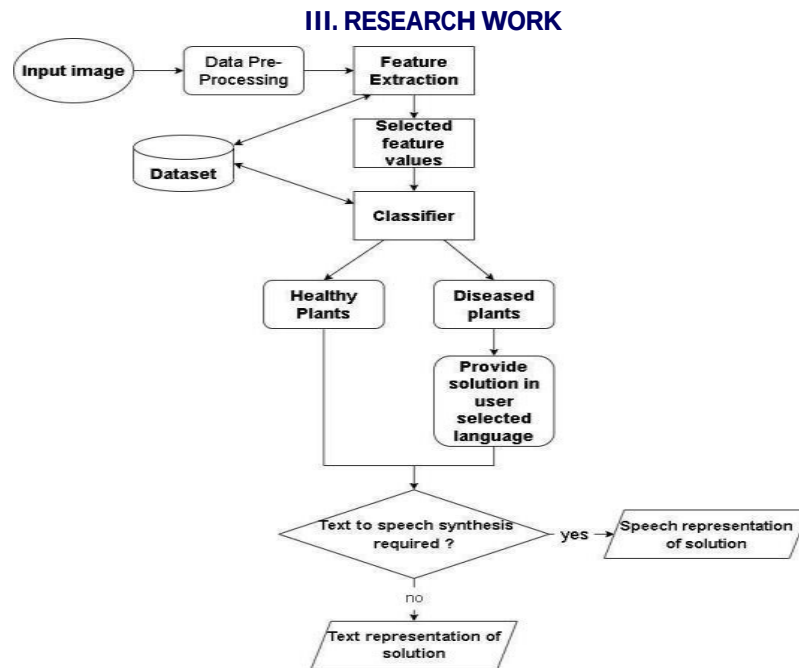


Figure1.Data Flow Diagram

The above mentioned works use different machine learning algorithms for disease classification but the accuracy of prediction was about 70-80%. The proposed model of Convolutional Neural Network consists of two convolutional layers followed by pooling layer each which uses the proposed activation function and the system accuracy is increased up to analysis of the model it is found that the system accuracy and performance has been improved. The new proposed function is seen in Figure 2, and the equation (6) represents the mathematical function proposed, 95%. The proposed work uses unsupervised learning for extracting features and pattern on its own. The above papers use linear classifiers like SVM, ANN whereas CNN uses existing non – linear activation functions for classification. Figure 1 represents the work flow of the proposed system. The input image from the dataset is pre – processed and the features are extracted using convolutional layer with proposed new activation function. Then the model is trained with the dataset. The test is used to find the accuracy and loss of model. The model uses flatten layer, fully connected layer and proposed activation function to classify the images. K-Means clustering algorithm is used to find the area affected for fertilizer optimization. And the same was implemented using Raspberry pi and OpenCV

A. Convolution Layer

Features from the image are extracted using small squares of input data to learn the features and relationship among pixels through this layer. This can be achieved by convolving the input image matrix with the filter matrix.

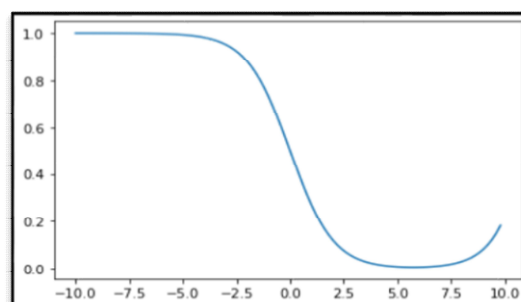


Figure2. Proposed Activation Function

On comparison with existing functions like sigmoid, tanh, ReLu and Softmax, the proposed function has a slope nearer to zero and has larger left half axis and reaches the optimal value. The advantage of the function is the derivative of proposed function curve keeps on changing and converges towards optimal value as shown in Figure2. The proposed function is a combination of various exponential functions so that it can satisfy the design

B. The Proposed Activation Function

Principle being analyzed. To prevent the gradient explosion problem the upward trend of the curve is Slowed and adjusted using log function in the region. As the real world data may have a non-negative linear values, the model must be a non – linear to process it. Hence non – linearity can be introduced only through activation function, The equations of the common activation functions are respectively defined as, Greater than zero.

C. Pooling Layer

Pooling layers reduce the number of parameters when the images are too large. Here max pooling of feature map matrix with stride value of 2 is used.

D. Fully Connected Layer: A flattening layer is used to convert three dimensionality of image to single for finding the probabilistic value, followed by two fully connected dense layer containing proposed activation function for classification.

K-Means Clustering: Image segmentation can be achieved using unsupervised neural algorithms like K-means clustering based on the cluster selection. The data are above equations of activation functions were implemented in CNN model but these functions resulted in poor accuracy and reduces the system performance. Therefore, in this work a new optimized activation function is designed to improve convergence. The model was tested with 100 images per disease with improved mathematical function.

E. Perception without activation Function

Our project analyses the role of mathematical function in CNN model and the optimizations that are achieved by using these functions for classifications. For binary classification of dataset perceptron with single layer [7] can be used as shown in Figure 3.

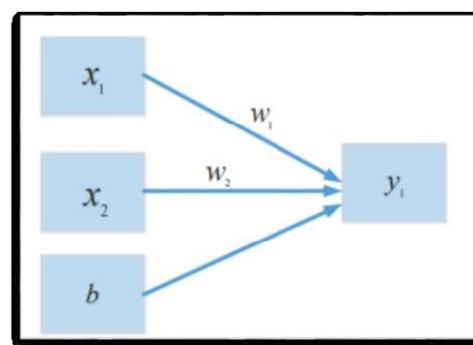


Figure3. Model wit out Activation Function

F. Perceptron with activation Function

But because the function used in the classifier is a linear, it cannot be used for a non-linear classification system. Figure4 shows the perceptron with a mathematical function.

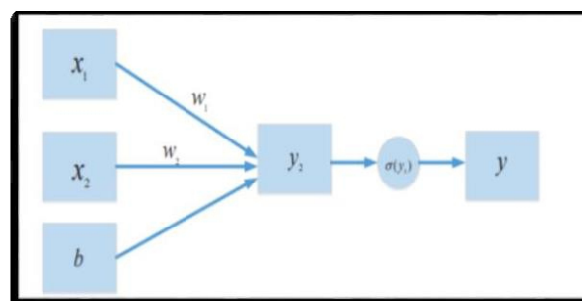


Figure4. Model with an Activation Function.

The perceptron with an activation function can be used to overcome the classification problem in non-linear system by Equations(10) and (11).

EXPERIMENTAL RESULTS AND ANALYSIS

A. Result Analysis

This system detects the presence of disease in leaf at early stages. The camera connected to Raspberry Pi captures the leaf image and is processed in both an aconda navigator and ARM processor for classification of disease using CNN. Further the processed image is clustered using clustering algorithm in MATLAB to find the area affected.



Figure 5. Trained Dataset

Figure 5 shows the trained dataset used for the CNN model. It consists of 100 leaf images infected with disease. Similarly, separate datasets of 100 images each are created for each disease. As the number of iterations increases, the accuracy of parameters for feature extraction increases. The total number of parameters extracted is 813,604 and the same was trained. Parameters in various layers are: 896 parameters were used in convolutional layer 1, 9248 parameters in layer 2, 802,944 and 516 parameters in fully connected dense layers.

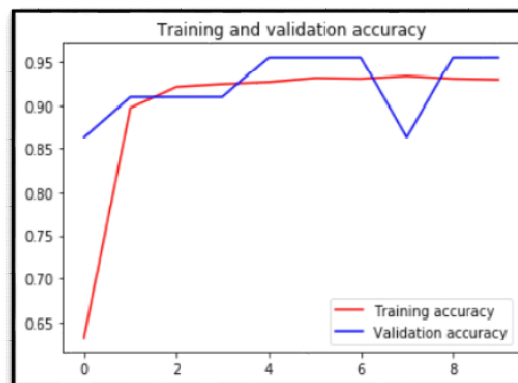


Figure 6. Accuracy Graph



Figure 7. Loss Graph

The Figure 6 and 7 shows the accuracy and loss plot of Anthraconse disease prediction with the proposed activation function. As the number of iteration increases the accuracy of prediction increases. Similarly the model can be tested for Antharas is Bacterial Blight, Cercospora LeafSpot, BlackRot and the accuracy and lossplot can be obtained. The Table1 represents the output of the CNN algorithm implemented in the Anaconda Navigator using Tensor-flow with proposed activation function. It shows the input test image of leaf and the parameters such as accuracy, loss and disease predicted are tabulated. As the training accuracy increases the sensitivity and specificity of the prediction increases.

Table1.Training and Validation Loss

S.NO	LEAF IMAGE	ACCURACY	LOSS	DISEASE PREDICTED
1.		0.9534	0.1091	Black Rot
2.		0.9725	0.0755	Bacterial blight
3.		0.9658	0.0882	Cercospora leaf spot
4.		0.9478	0.1212	Black Rot
5.		0.9667	0.0818	Bacterial blight
6.		0.9600	0.0996	Anthraconse

Accuracy of model is its ability to differentiate the diseased and healthy cases correctly. Mathematically it can stated as,

True positive (TP) = total images classified as diseased
False positive (FP) = total images misclassified as diseased
True negative (TN) = total images classified as not diseased
False negative (FN) = total images misclassified as not diseased

For finding the area affected by the disease, in the K-means clustering algorithm, when K=3 three different areas of clustering regions are obtained and the disease affected cluster is selected and the resultant of area affected is displayed. For the cropdisease Bacterial Blight, the fertilizer suggested is 40kg Murate of Potash / hector. The amount of fertilizer is calculated as:

Amount of fertilizer = 40 x Percentage of area affected x total area of field (Ha)

```

C:\Python\Python374\
File Edit Format Run Options Window Help
Python 2.7.13 (default; Jan 19 2017, 14:48:00)
[GCC 6.3.0 20170124] on linux2
Type "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: /home/pi/Desktop/plant/plant disease/cnn/TrainAndTest1.py =====
ANTHRASIS BACTERIAL BLIGHT

Traceback (most recent call last):
  File "/home/pi/Desktop/plant/plant disease/cnn/TrainAndTest1.py", line 151, in <module>
    main()
  File "/home/pi/Desktop/plant/plant disease/cnn/TrainAndTest1.py", line 132, in main
    retval, npaResults, neigh_resp, dists = kNearest.find_nearest(np480Resized, k = 1)
error: /build/opencv-U0vIN/opencv-3.4.9.1+dfsg1/modules/ml/src/knearest.cpp:370: error:

```

Figure 8. Hardware Execution Output

The figure 8 shows the input and clustered output image with the part of disease affected when the CNN algorithm is implemented in Raspberry Pi. And also the output terminal of disease detection with a message display as "ANTHARASIS BACTERIAL compared to the existing activation functions. From equations (13) and (14) the sensitivity and specificity can be calculated as, Figure 9 shows the comparison of existing activation function with the proposed activation function based on accuracy. And it can be inferred that the new activation function has the maximum accuracy of 95 % compared to the existing activation functions.

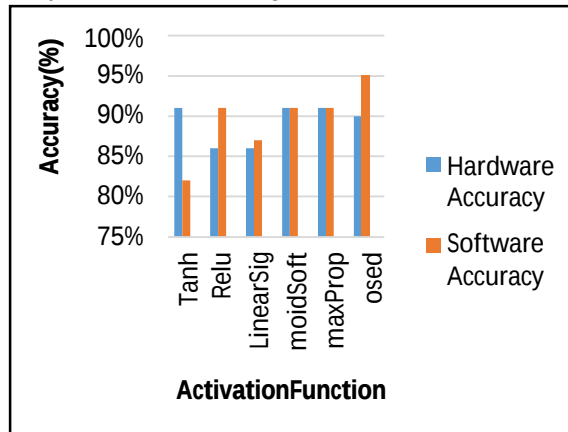


Figure 9. Activation Function vs Accuracy

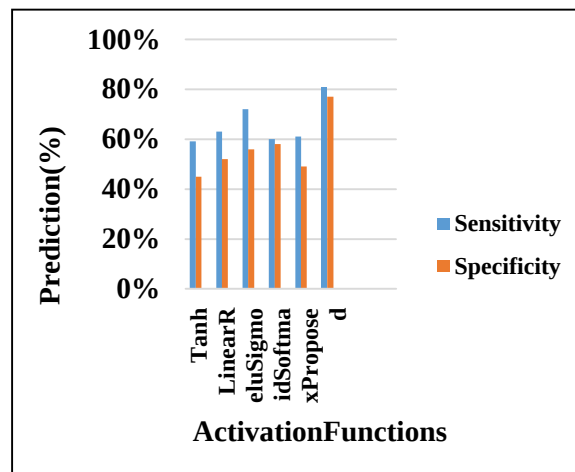


Figure10.Activation Function vs Prediction

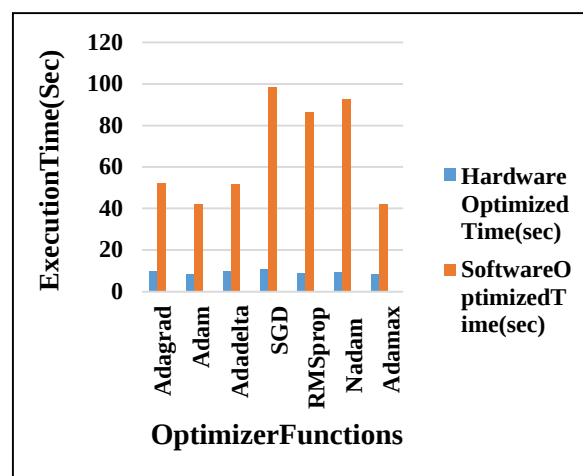


Figure11.Optimizer vs Execution time

Figure 10 shows the comparison of existing activation function with the proposed activation function based on sensitivity and specificity.

And it can be inferred that the new activation function has the maximum sensitivity and specificity of 81%. As the sensitivity and specificity increases the prediction accuracy increases and is found to be 95%. Optimizers are used to reduce the training and validating time of the model. The figure 9 shows the comparison of various optimizers and their execution time. And it can be inferred that 'Adam' and 'Adamax' optimizers are more efficient compared to other optimizers with a minimum execution rate of 12 seconds when implemented without optimizer and 8 seconds when the algorithm is executed with optimizer function.

CONCLUSION

The convolutional Neural Network model was implemented in Tensor Flow back end system to classify the plant diseases. The same was implemented for real time data using Raspberry Pi in Open CV. The most commonly used optimizers and activation functions were analyzed. It can be inferred that the proposed function optimizes the system with greater accuracy of 95%. The parameters like Sensitivity, Specificity were calculated. It is inferred that specificity and sensitivity are more than 80%. The 'Adam' and 'Adamax' optimizers show better optimization compared to other functions. The K-means clustering algorithm was implemented in MATLAB to calculate the affected area percentage and the optimized amount of fertilizer is suggested to improve the crop yield. Further an android app can be developed for this application.

RESULTS

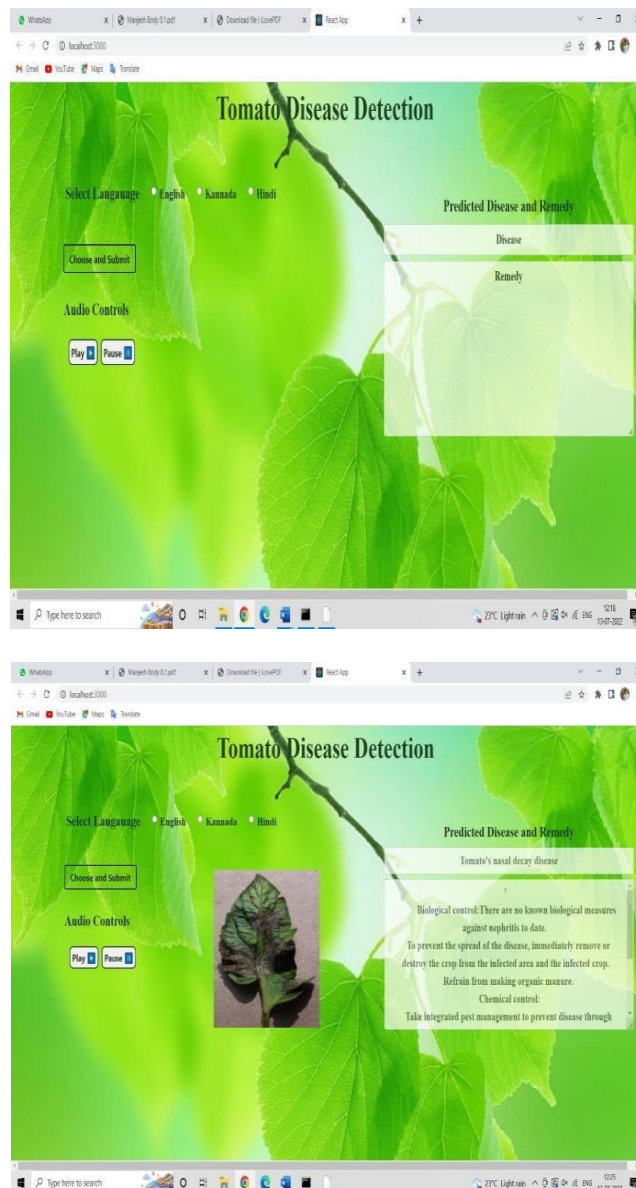


FIGURE 12.
GUI INTERFACE

FUTURE WORK

In India Farmers welfare department wants to automate the detecting of Diseases from eligibility process (real time). Many efforts from both local and state government are done in order to understand and predict leaf diseases of plants. In order to optimize the Artificial Intelligence environment, the prediction system was implemented. With the advancement of IoT infrastructures, big data technologies, and machine learning techniques, real-time monitor and evaluation is desirable for future smart cities. For future work, the commonalities between nearby Rural areas. Which may provide a further boosting for the prediction can be considered. This work can be extended by considering data changes over time for real-time data collection. This can be achieved by building online models that adapt automatically to changes in health care. Also, more data can be seasonality.

REFERENCES

1. <https://en.wikipedia.org/wiki/Leafdiseaseprediction>
2. G.Madhulatha and O.Ramadevi, " Recognition of Plant Diseases using Convolutional Neural Network," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 738-743, doi:10.1109/I-SMAC49090.2020.9243422.
3. S. Y. Yadhav, T. Senthilkumar, S. Jayanthi and J. J. A. Kovilpillai, " Plant Disease Detection and Classification using CNN Model with Optimized Activation Function," 2020 *International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2020, pp.564-569, doi:10.1109/ICESC48915.2020.9155815.
4. D. Samanta et al., "Development of multimodal user interfaces to Internet for common people," 2012th International Conference on Intelligent Human Computer Interaction (IHCI), Kharagpur, 2012, pp.1-8, doi:10.1109/IHCI.2012.6481775.
5. Kamilaris, A. (2018). DEEP LEARNING IN AGRICULTURE: A SURVEY. *Comput. Electron. Agric.*, vol. 147, no. July 2017, pp. 70–90, 2018.
6. Mohanty. S.P and Salathé. M. (2016). USING DEEP LEARNING FOR IMAGE-BASED PLANT DISEASE DETECTION. *Front. Plant Sci.*, vol. 7, no. September, pp.1–10.
7. Muhammad, S. (2018). DETECTION AND CLASSIFICATION OF CITRUS DISEASES IN AGRICULTURE BASED ON OPTIMIZED WEIGHTED SEGMENTATION AND FEATURE SELECTION, Elsevier