International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Chilli Leaf Diseases Detection with Different Features of Original Chilli Using Region Based Convolutional Neural Network

Raja K. [1], Shiny Duela J. [2], M. Gopichandd [3*], Karthik Kannan [4], M. G. Sathish [5]

Submitted: 22/08/2023 **Revised**: 09/10/2023 **Accepted**: 20/10/2023

Abstract: The productivity of the agriculture sector drives the Indian economy. Rural households rely on agriculture to a greater than 70% extent. Almost 60% of the country is employed in agriculture, which generates roughly 17% of the global GDP. Thus, within the domain of agriculture, the detection of crop diseases is crucial. Rice, wheat, groundnuts, and including variety of crops, but not limited to, fruits, vegetables, and other plants. In addition to these crops, Indian farmers also raise potatoes, oilseeds, sugarcanes, and non-food commodities including coffee, cocoa, tea, rubber, and cotton. Most plants primarily grow depending on the energy of their roots and leaves. There are also some more factors that result in various plant-leaf diseases, which ruin harvests and ultimately impact the nation's economy. Chilli production is a skilled labor-intensive operation because plants are constantly under attack from various insects, bacterial diseases, and smaller-scale organisms. Studies of organic products show that leaves and shoots are commonly used to identify attack marks. Currently, chemicals are being applied to plants without paying attention to their needs. This technique will ensure that chemicals such as pesticides are only applied to plants when they are infected with the diseases. Images of the chilli leaf disease were captured using image processing techniques. The leaf image will be used to identify and estimate the state of the plants. There are also several data processing techniques, which are indeed effective as well as efficient for determining plant diseases to support farmers. In this, we have implemented GLCM feature extraction and region-based CNN method. Our Proposed model aims to give an accuracy of above 90%. It will be helpful in various agricultural applications. Firstly, it can be beneficial for assisting non-expert farmers in identifying the right time to apply chemicals to the plant. It also provides the most appropriate time for harvesting crops before they get ruined. Secondly, it can also be used by the researchers of plant research institutes to study crops in detail in different stages of the crop.

Index Terms**:** *Deep learning, CNN, Keras, GLCM, Python, OpenCV-Python, TensorFlow, Anaconda.*

1. Introduction

One way in which deep learning is commonly used in agriculture is through the implementation of convolutional neural networks (CNNs) to identify instances of Chilli leaf disease. The Indian economy heavily relies on agriculture, which accounts for 17% of the GDP and employs over 60% of the population, including 70% of rural households. Chilies are a crucial crop because India produces 43% of the world's total output, making it the top producer. A significant crop that is grown all over the globe is the chilli. Although it is generally believed that chillies originated in central or southern America, reports indicate that they were first cultivated in Mexico. In 2019, there were 38 million tonnes of fresh green chillies, with China producing half of them [1]. Chilli plants are vulnerable to diseases that impacts production and yield.[2,3] which gives the information that which diseases(like leaf spot, leaf curl etc) and crop improvement techniques.

After collecting dataset preprocessing and feature extraction play an important role for the result, choosing correct feature extraction method with respective to classifier will yield us a good accuracy. For detection of disease in chilli leaf we need to do texture analysis so comparing all approaches will result the best approach[4].Finding a best classifier is a difficult task after going through many related works, [5]uses KNN classifier with the help of GLCM feature extraction, [6] using LBP feature extraction and svm and other techniques, [7]in this work we used squeezenet CNN for Tomato disease detection, and some other techniques(like Soft Computing Techniques and squeezenet) [8,9,10].

Keras, a Python deep learning library, can then be used to build the CNN model. This model includes several convolutional layers followed by multiple max pooling layers which helps to reduce dimensionality of feature

[1] *Professor and Head of the Dept. of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Tamil Nadu, India*
drkrajamit@gmail.com
[2] *Associate Professor of the Dept. of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Tamil Nadu, India*
shinyduela@gmail.com
[3] *Student of the Dept. of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Tamil Nadu, India*
mg3445@srmist.edu.in
0009-0002-1520-2847
[4] *Student of the Dept. of Computer Science and Engineering, SRM Institute of Science and Technology, Tamil Nadu, India*
kk6931@srmist.edu.in
[5] *Student of the Dept. of Computer Science and Engineering, SRM Institute of Science and Technology Ramapuram Campus, Tamil Nadu, India*
sm7040@srmist.edu.in

maps. Afterwards, we have the option to include a fully connected layer that condenses to the final number of nodes in the output layer, which should correspond to the total number of classes. (Healthy or diseased).

2. Literature Review

C. Trongtorkid et al. [11] developed an expert system for diagnosing mango diseases through the evaluation of leaf symptoms in 2018. A model based on rules and utilizing a dataset of leaf images has been presented by them. The rule-based model was performed with the images of 129 leaf that were collected as from the mango farms. under the constant monitoring of quality of product, Maejo University, and three answers with regards to the label assigned to a particular specimen, it may fall under one of three categories: Anthracnose, Algal Spot, or normal.has been tested, and the results show that it has an accuracy of 89.92%. The experimental results showed that the rules-based approach might be used for applications such as plant diagnostics, but we observed that climate causes variation as well as accuracy changes. A Neural Network Toolbox-based concept was presented in 2018 by R. G. de Luna et al [12]. Here, we learned that while accuracy was desirable, one downside is difficulty for hyper-parameters.

Table 1 Shows difference between Existing system/Paper.

Paper No	Year	Algorithm	Accuracy	Drawbacks
[1]	2018	Decision tree algorithm	89.92%	Here, climate causes the variations in the leaf images and changes the overall accuracy of the classification.
[2]	2018	Neural Network Toolbox Algorithm	91.67%	There was difficulty for hyper-parameters as well the accuracy was not good.
[3]	2022	Adam optimization Algorithm	95%	leaf images from other geographic areas with varying image quality changes accuracy.
[4]	2021	Hybrid intelligent system Algorithm	92.10%	The process of isolating color pixels that were unclear due to being situated against a similarly colored background was found to be limiting, which resulted in a prolonged classification time.
[5]	2022	2 Level SFAM (2L-SFAM)	90%	It spends long time to finish the classification of grape leaf diseases.

In 2021, Vibhor K Vishnoi, Krishnan Kumar, Brajesh Kumar, Shashank, and Arfat Khan [13] introduced Identification of diseases affecting apple plants through the use of leaf images via Convolutional neural network CNN, which was using Adam's optimization Algorithm. Although this method was well-performing, the accuracy was altered by the use of Leaf images from different geographic areas with different image quality. The thoughts from Ahmad Loti, Nurul Nabilah, Mohamad Roff

Mohd Noor, and Siow-Wee [14] were shared. "Integrated analysis of deep learning and machine learning in identifying chile pests and diseases" was proposed in 2021. it was based on a hybrid intelligent system. In this paper, characteristics of chile diseases and pests which were extracted using a conventional method and were compared to those extracted through deep-learning method were contrasted. A collection of 974 pics of chilli leaves with five different diseases were obtained. Six traditional feature-based methods, as well as deep learning feature-based methods, were employed to extract significant traits related to pests and diseases from images of chili leaves. The SVM classifier had the highest accuracy (92.10%), but there were some limitations of extracting ambiguous colour pixel from the background image.

The Research was conducted on a Grape Leaf Disease System that utilized both color imagery and gray level co-occurrence matrix, specifically with regards to the implementation of a 2-Level Simple Fuzzy ARTMAP was proposed by K. Phookronghin et al. [15] in 2018. It makes use of the 2L-SFAM algorithm. The main advantage of 2LSF AM lies in its ability to incorporate entirely new categories of data without necessitating the complete retraining of the network. Furthermore, it can adapt and modify its data categories on a continuous basis, thereby reducing the time required for learning and classification. The execution of the suggested technique indicates a desired level of accuracy, demonstrating how well the 2L-SF AM will classify and identify grape leaf diseases as well as their various stages. One thing we noticed, however, was that the system was taking a long time to finish, also the performance was not up to the mark. [16,17,18] papers were having one or other drawbacks such as the model would find difficulty in identifying rare diseases also the accuracy of the proposed system can be enhanced. [19] To diagnose diseases affecting plant leaves, they have employed a method that involves using an image feature selection technique based on the slime mold optimization algorithm while there could have better improvement in classification speed of proposed system. [20] This paper discusses the utilization of unsupervised learning for the identification of plant diseases, with a specific focus on the use of image restoration techniques and in it they also use Deep learning correlation model for identification of the diseases in the region but the time for overall accuracy for classification of diseases is long.

3. Existing Work

The current method for identifying chilli leaf infections often comprises a qualified professional manually inspecting the plants. This approach is time-consuming, subjective, and susceptible to the inspectioner's level of experience. Moreover, certain automated approaches for spotting plant diseases, such those utilising machine learning and image processing, have been developed. For example, numerous studies have utilized leaf photographs and conventional machine learning techniques such as support vector machines (SVMs) and decision trees to classify instances of Chilli leaf disease based on certain characteristics and the effectiveness of these techniques, however, may be constrained by the quality of the extracted features, and the diversity of the leaf pictures may have an impact on classification accuracy.

They have used Principal component Analysis, Region based segmentation, KNN Classifier. Dependence on manual inspection, lack of efficiency was some of the drawbacks from the existing model.

4. Proposed Work

A. Input Image

An input image refers to the image that is being analyzed by the model. In the case of detecting chilli leaf disease, the input image would typically be an image of a chilli leaf.

The input image is a critical component of the analysis, as it contains the information that the model will use to make predictions or classifications. The quality of the input image can have a considerable influence on the accuracy of the predicted model. Lighting conditions, picture resolution, and the presence of noise or artefacts are all factors that might alter the quality of the input image.

Healthy Leaf Diseased Leaf

Fig 1. displays leaf imagery of both diseased and healthy plants.

To ensure that the input image is suitable for analysis, it may be necessary to preprocess the image before feeding it into the model. Depending on the model's specific requirements and the features of the input picture, this might entail techniques such as image normalization, image cropping, or image filtering.

Overall, input image is a critical component of any computer vision analysis, and its quality and suitability for analysis must be carefully considered in order to obtain accurate and reliable results.

B. Pre-processing

Pre-processing is a critical step in many computer vision applications, including detecting chilli leaf disease. It involves applying a series of techniques to the input image in order to prepare it for analysis by the model. Pre-processing aims to improve the quality of the input image, eliminate noise or artefacts, and extract valuable characteristics that will help in illness identification.

Pre-processing techniques commonly utilised in computer vision applications include:

1. Image resizing: Scaling the input image to a lower or bigger size improves the efficiency of the analysis or makes it more suited for the model's unique requirements.

2. Image normalization: This involves adjusting the brightness, contrast, or colour balance of the input image in order to standardize its appearance and make it easier to analyse.

3. Image filtering: Here, we are using images from test dataset as input which is already noise free. Therefore, we won't be performing noise removal in our work, but for images directly from source we need to perform image filtering for better performance of our model.

4. Feature extraction: This involves identifying and extracting important features from the input image that can be used to identify the disease. This might involve techniques such as edge detection, texture analysis, or colour histograms.

C. GLCM Feature Extraction

GLCM (Gray-Level Co-occurrence Matrix) is a texture analysis approach used in image processing applications such as identifying chilli leaf illness. It entails creating a matrix based on gray-level values that defines the connection between pairs of pixels in a picture.

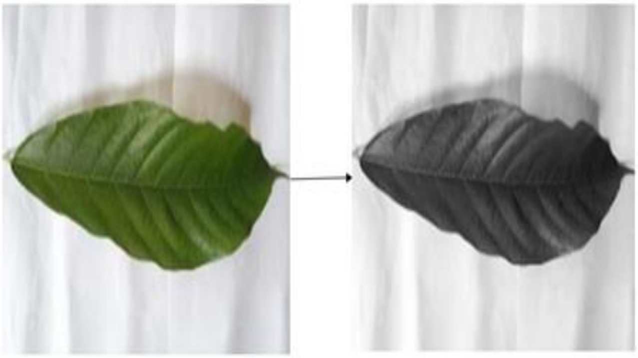

Fig 2. shows Leaf Image converted to Grayscale Image.

The GLCM matrix is calculated by computing the co-occurrence of pairs of pixels in the image at a specified offset or distance. For each pair of pixels, the matrix stores the number of times that the pixel with a given gray-level value occurs in a specified relationship to another pixel with a different gray-level value.

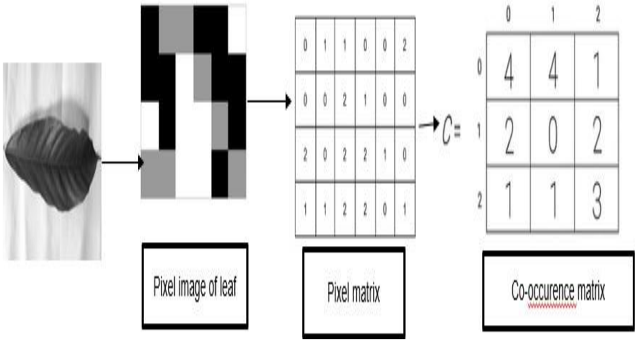

Fig 3. shows image pixel converted to co-occurrence matrix.

Once the GLCM matrix has been calculated, a range of statistical measures can be computed from it to extract texture features that can be used for classification.

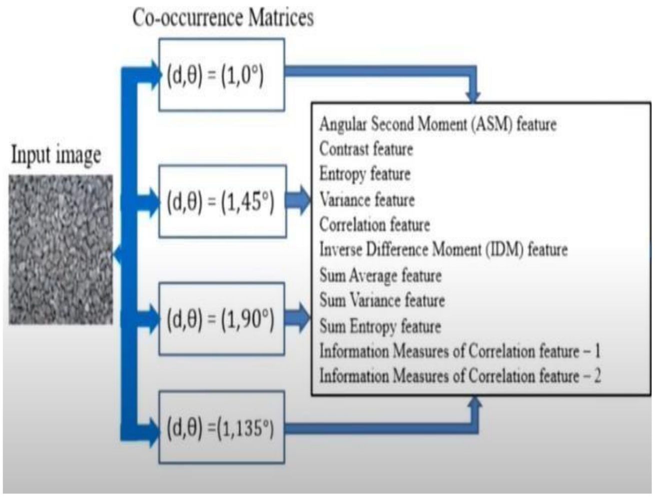

Fig 4. shows input image converted to different features.

Some common features extracted from the GLCM matrix include:

• **Contrast:** The variation in gray-level values among two pixels is measured.

• **Energy:** The broad commonality of gray-level frequencies within the representation can be assessed.

• **Homogeneity:** The proximity of the distribution of gray-level levels in the representation is measured.

• **Entropy:** The fluctuation of gray-level readings in the imagery is measured.

• **Correlation:** Over the whole image, this function provides a measure of how linked a pixel is to its neighbour.

Texture feature	Equation
Contrast	$\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2P(i,j)$
Entropy	$-\sum_{i=1}^{N}\sum_{j=1}^{N}P(i,j)\,lgP(i,j)$
Correlation	$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(i-\bar{x})(j-\bar{y})P(i,j)}{\sigma_x\sigma_y}$
Energy	$\sum_{i=1}^{N}\sum_{j=1}^{N}P(i,j)^2$

Fig 5. Equations for calculating different Features.

Here, D = distance in between two elements of a matrix,

i & j = no of rows and the columns, P = Probability.

Based on the texture of the leaf, these characteristics may be utilised to train a machine learning model to identify chilli leaf disease. Beyond detecting chilli leaf disease, GLCM feature extraction is a powerful approach for analysing picture texture that may be utilised in a range of other applications such as finding faults in industrial materials or analysing medical images.

D. Architecture Diagram

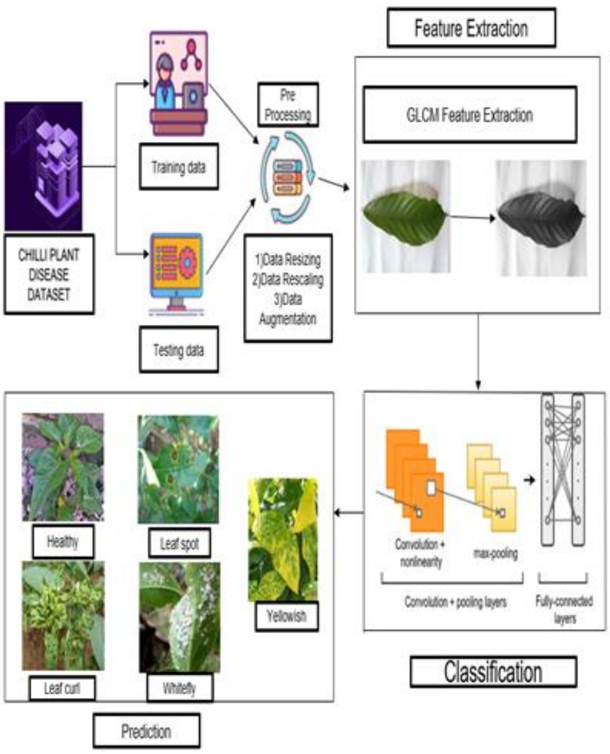

Fig 6. Architecture Diagram of our model.

E. Convolutional Neural Network

Convolutional neural networks (CNN) are a form of neural network that is widely utilised in computer vision applications, such as identifying chilli leaf illness. CNNs are particularly intended to analyse image data and can learn features important for image classification tasks automatically. CNNs are made up of layers such as convolutional, pooling, and fully connected.

Convolutional layers extract features such as corner and edge detail from an input picture through the use of a sequence of filters or kernels. The feature maps generated through the convolutional layers are then down sampled through pooling layers, minimising the amount of parameters in the network and aiding in the prevention of overfitting. Finally, the outcome of the pooling layers is fed through one or more fully connected layers, which predict or classify using the collected characteristics.

F. Benchmark of our Model

Table 2 Shows difference between different algorithms in classification.

Sl.No	Algorithm	Accuracy
[1]	Decision tree algorithm	89.92%
[2]	Neural Network Toolbox Algorithm	91.67%
[3]	Adam optimization Algorithm	95%
[4]	Hybrid intelligent system Algorithm	92.10%
[5]	2 Level SFAM (2L-SFAM)	90%
[6]	Region Based Convolutional Neural Network	96.75%

G. Flowchart of our Model

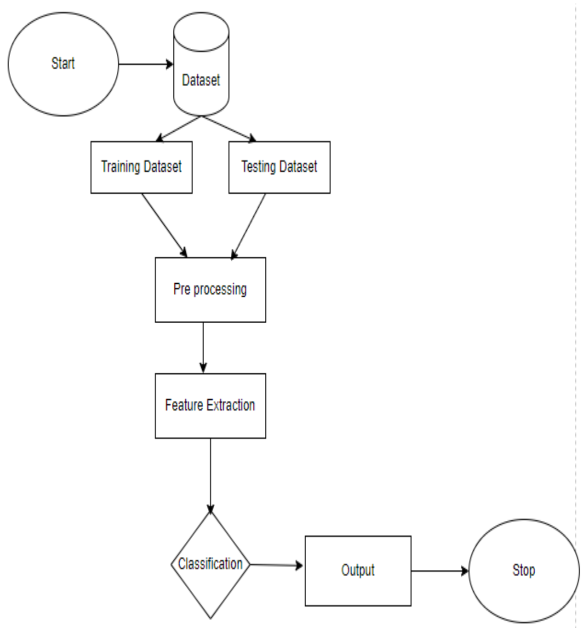

Fig 7. Flowchart of our Model.

5. Algorithm used

The algorithm we are using is Region Based Convolutional Neural Network. The steps involved in creating our proposed model are listed down.

1. We Use a Region Based CNN to detect multiple objects (in this case, chili leaves) in an image.
2. Extract regions of interest (ROIs) corresponding to the detected chili leaves.
3. To obtain texture characteristics such as contrast, entropy, homogeneity, and correlation, apply Gray-level Co-occurrence Matrix (GLCM) feature extraction to each ROI.
4. Use the extracted texture features as inputs to a classification model to classify each ROI as healthy or diseased.
5. Finally, aggregate the classification results for all ROIs in the image to determine the overall disease status of the chili plant.
6. In Result, it identifies the specific diseases of the crop such as whitefly, yellowish, leaf spot, leaf rot and shows if the plant is healthy.

By using Region Based CNN with GLCM feature extraction, we can efficiently detect and classify chili leaf disease while minimizing computational complexity.

6. Results

The proposed system gives an accuracy of 96.75% for our dataset. It helps in identifying different type of diseases such as leaf spot, whitefly, healthy, yellowish. This system will be helpful in classification of our diseases that the crops are affected.

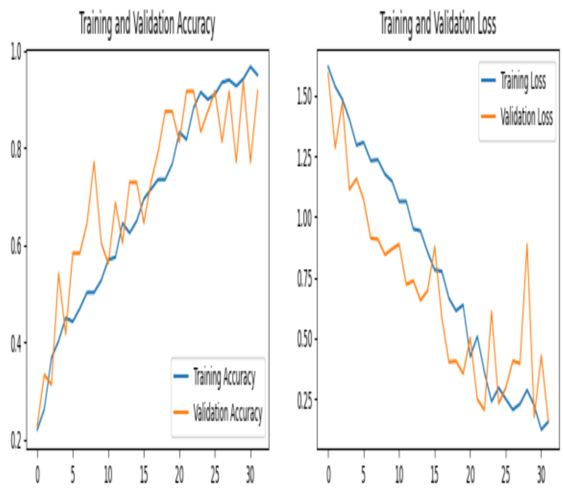

Fig 8. Shows Graph Between Training and Validation Accuracy/Loss.

Fig 9. Shows Classification of Leaf images.

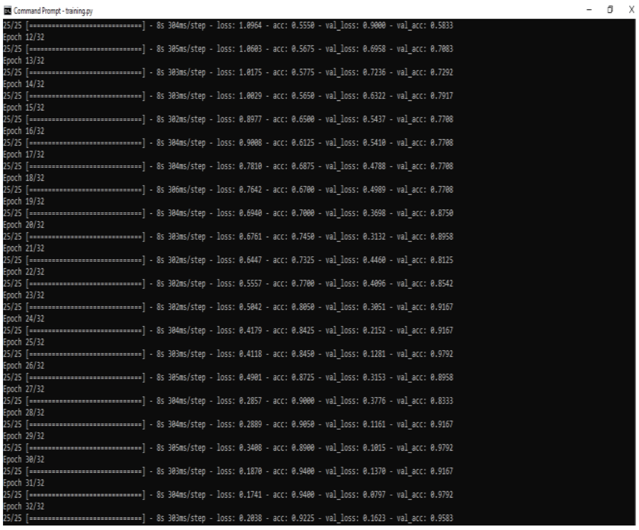

Fig 10. Shows training accuracy output of the model

7. Conclusion

The comprehensive analysis reveals that the suggested model outperforms several pre-trained CNN models. Based on different characteristics such as accuracy and memory needs, the approach was also determined to be superior to certain other current methods. For several disorders, the model assures and achieves high accuracy of 90% or more. The model properly balanced precision and accuracy. Collecting leaf photos from various locations may be a future project. The model's performance is then fed and analysed.

References:

[1] https://en.wikipedia.org/wiki/Chili_pepper.

[2] Manas Kumar Pandit, Ritoban Pandit, Sanjay Bairagi, “Chili”, In book: Ethnopharmacological Investigation

of Indian Spices (pp.253- 268),January 2020, DOI:10.4018/978-1-7998-2524-1.ch018.

[3] N. Nisar, A. Awasthi, M. Chhabra and A. I. Abidi, "Image based Recognition of Plant Leaf Diseases: A Review," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 373-378, doi: 10.1109/I-SMAC49090.2020.9243434.

[4] ANNE HUMEAU-HEURTIER, “Texture Feature Extraction Methods: A Survey” ,IEEE, Electronic ISSN: 2169-3536, 03 January 2019 , DOI: 10.1109/ACCESS.2018.2890743.

[5] Eftekhar Hossain, Md. Farhad Hossainand, Mohammad Anisur Rahaman, “A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier”, 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), DOI:10.1109/ ECACE.2019.8679247.

[6] Matta Bharathi Devi, Dr Amarendra,“Machine Learning based Application to detect Pepper Leaf Diseases using HistGradientBoosting Classifier with fused HOG and LBP Features”, Volume 8. No. 10, October 2020,DOI: https://doi.org/ 10.30534/ijeter/2020/1148102020.

[7] Akbar Hidayatuloh , M. Nursalman , Eki Nugraha , Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model,IEEE, Electronic ISBN:978-1-5386-5693-8, 25 April 2019, DOI: 10.1109/ICITSI.2018.8696087.

[8] J. Francis, Anto Sahaya Dhas D and Anoop B K, "Identification of leaf diseases in pepper plants using soft computing techniques," 2016 Conference on Emerging Devices and Smart Systems (ICEDSS), Namakkal, India, 2016, pp. 168-173, doi: 10.1109/ ICEDSS.2016.7587787.

[9] A. Hidayatuloh, M. Nursalman and E. Nugraha, "Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model," 2018 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia, 2018, pp. 199-204, doi: 10.1109/ICITSI.2018.8696087.

[10] V. A. Metre and S. D. Sawarkar, "Reviewing Important Aspects of Plant Leaf Disease Detection and Classification," 2022 International Conference for Advancement in Technology (ICONAT), Goa, India, 2022, pp. 1-8, doi: 10.1109/ICONAT 53423.2022.9725870.

[11] C. Trongtorkid and P. Pramokchon, "Expert system for diagnosis mango diseases using leaf symptoms analysis," 2018 International Conference on Digital Arts, Media and Technology (ICDAMT), Phayao, Thailand, 2018, pp. 59-64, doi: 10.1109/ICDAMT. 2018.8376496.

[12] R. G. de Luna, E. P. Dadios and A. A. Bandala, "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition," TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018, pp. 1414-1419, doi: 10.1109/TENCON.2018.8650088.

[13] V. K. Vishnoi, K. Kumar, B. Kumar, S. Mohan and A. A. Khan, "Detection of Apple Plant Diseases Using Leaf Images Through Convolutional Neural Network," in IEEE Access, vol. 11, pp. 6594- 6609, 2023, doi: 10.1109/ACCESS.2022.3232917.K. Elissa, “Title of paper if known,” unpublished.

[14] Ahmad Loti NN, Mohd Noor MR, Chang SW. Integrated analysis of machine learning and deep learning in chili pest and disease identification. J Sci Food Agric. 2021 Jul;101(9):3582-3594. doi: 10.1002/jsfa.10987. Epub 2020 Dec 24. PMID: 33275806.

[15] K. Phookronghin, A. Srikaew, K. Attakitmongcol and P. Kumsawat, "2 Level Simplified Fuzzy ARTMAP for Grape Leaf Disease System Using Color Imagery and Gray Level Co-Occurrence Matrix," 2018 International Electrical Engineering Congress (iEECON), Krabi, Thailand, 2018, pp. 1-4, doi: 10.1109/IEECON.2018.8712183.

[16] S. Kaur, S. Pandey, and S. Goel, “Plants Disease Identification and Classification Through Leaf Images : A Survey,” Archives of Computational Methods in Engineering, vol. 6, 2018. [Online]. Available: https://doi.org/10.1007/s11831-018-9255-6.

[17] V. K. Vishnoi, K. Kumar, and B. Kumar, “A comprehensive study of feature extraction techniques for plant leaf disease detection,” Multimedia Tools and Applications, vol. 81, no. 1, pp. 367–419, Jan 2022.

[18] Z. B. Husin, A. Y. B. M. Shakaff, A. H. B. A. Aziz and R. B. S. M. Farook, "Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques," 2012 Third International Conference on Intelligent Systems Modelling and Simulation, Kota Kinabalu, Malaysia, 2012, pp. 291-296, doi: 10.1109/ ISMS.2012.33.

[19] S. M. Javidan, A. Banakar, K. A. Vakilian and Y. Ampatzidis, "A feature selection method using slime mould optimization algorithm in order to diagnose plant leaf diseases," 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS),

Behshahr, Iran, Islamic Republic of, 2022, pp. 1-5, doi: 10.1109/ICSPIS56952.2022.10043928.

[20] M. Pei, M. Kong, M. Fu, X. Zhou, Z. Li and J. Xu, "Application research of plant leaf pests and diseases base on unsupervised learning," 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), Changchun, China, 2022, pp. 1-4, doi: 10.1109/CVIDLICCEA 56201.2022.9824321.

[21] Bhat , A. H. ., & H V, B. A. . (2023). E2BNAR: Energy Efficient Backup Node Assisted Routing for Wireless Sensor Networks . International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 193–204. https://doi.org/10.17762/ijritcc.v11i3s.6181

[22] Sahoo, D. K. . (2022). A Novel Method to Improve the Detection of Glaucoma Disease Using Machine Learning. Research Journal of Computer Systems and Engineering, 3(1), 67–72. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/44