Detection of Apple and Maize Plant Diseases using Self Organized Maps and Deep Boltzmann Machine

Rahul Chiranjeevi V   
*Computer Science and Engineering*  
*Rajalakshmi Engineering College*Chennai, India  
rahulchiranjeevi777@gmail.com

Janani V  
*Computer Science and Engineering*  
*Rajalakshmi Engineering College*Chennai, India  
jananiraghavan1507@gmail.com

Janani Priya N  
*Computer Science and Engineering*  
*Rajalakshmi Engineering College*Chennai, India  
jenipriya118@gmail.com

*Abstract*— The food supply is the absolute effective management across the world where the technology proposes optimizing outcomes that improves the yield and quality day by day. The past detection and classification of plant diseases had a great influence on the agricultural input for the improvement of agricultural technology. The disease can have a destructive effect on food production thus leading to a shortage of food. This emphasizes the automatic identification and detection of diseases of the apple and maize plants that involves many Machine Learning algorithms, technical solutions, and experiments that are based on various datasets. The detection and identification of the disease have a very huge impact on the well-being of farmers as it improves the quality of the crops and helps in the production of the plants. The machine learning models enhance precision agriculture by targeting disease prediction and management and ensuring efficiency in the field. This leads to reduced usage of pesticides and chemicals that are used to overcome the more vigorous disease. This prevents the further spread of the diseases and intimates the farmers to make strategic and managemental decisions or the necessary actions that are to be taken. These reductions in the usage of these chemicals can also minimize environmental pollution and promote sustainable agricultural practices. This incorporates various ensemble models that have the combined architecture comprised of the strengths of both the self-organized maps and the restricted Boltzmann Machine with Autoencoders employed. The study emphasizes the analysis of the plant diseases that affect the apple and maize plants, the utilization of this proposed technique yields an optimizing accuracy.

*Keywords—SOM, DBM, Apple and maize Plant disease.*

# Introduction

Disease of a plant is an abnormal condition that affects the health and life of a plant. Like humans and animals, plants are also affected by various factors that leads to malfunctioning of the plant and its production is reduced. The plants are affected by various factors like pathogens, environmental factors, deficiencies in nutrition, and further more. There are abiotic and biotic. Abiotic diseases are the ones that are caused by the external factors of a plant. They do not include the living agents. The abiotic factors include nutritional deficiencies, soil densification, salt injury, chillness, dryness, scorching sun etc,. The pathogens are nothing but the living organisms that causes the biotic diseases. These pathogens are the ones that cause the diseases in plants. Pathogens can spread from one plant to other and affects the plants that are spread to. These pathogens can affect the leaves, stems, roots, fruits, seeds and may more parts of the plant by affecting the whole functioning of the plant. Agricultural Business is one of the essential source that provides nourishment to the huge population in and around the world. How do Plant disease affect the agriculture? One of the direct impacts of plant diseases in the agriculture is the massive decrease in harvest and crop yield. Diseased plants may affect the quality of the crops, vegetables, fruits, thus by producing smaller, misshapen, discoloured products which in return reduces the market value. Rigorous outbreak of plant disease can result in significant crop losses, crop failures or entire loss of harvest. This in turn causes economic impact on farmers, agricultural industries, etc,. Plant disease can affect the production of good nutritious food thus by threatening the food industry.

Meeting the present-day population’s food requirements has become huge task as the population is increasing in number and also due to the climatic changes that is been happening. Infact the people nowadays are particular in the quality of the food ensuring fresh and uninfected vegetables , fruits and greens. Disease can have destructive effect on the food production thus by leading to shortage of the food. Predicting the plant diseases is an important job for several reasons which includes the early detection and intervention of loss of crop and ensures to provide heathy fruits and vegetables. This prevents the further more spread of the diseases and intimates the farmers to make strategical and managemental decisions or the necessary actions that are to be taken. The Machine Learning models are been used rather than manual detection for detecting the diseases at early stages, even before the symptoms are visible to the naked eye. By accurately predicting and managing the plant diseases the farmers can easily maintain the health and vitality of the plant. The machine learning models enhances the precision agriculture by targeting disease prediction and management and ensures efficiency in the field. This leads to reduced usage of pesticides and chemicals that are used to overcome the disease that is more vigorous. These reduction in usage of these chemicals can also minimize the environmental pollution and promotes sustainable agricultural practices.

Apples are one of the most popular fruits that is been consumed by everyone around the world starting from a child to the grown-ups. As the proverb suggests “An apple a day keeps the doctor away”, Apple has become an essential fruit to be consumed as it has enormous nutritional values. Similarly Maize is also a staple widely consumed food. Particularly its consumption is more in the regions such as North America, Africa and parts of Asia. Maize also plays a major role in the Animal feed for livestock such as poultry, pigs, and cattle. Hence the Apple plant and Maize plant diseases are particularly been concentrated , predicted and managed. In India fungal diseases are the major cause of affecting the Apple’s quality and production quantity. Apples are most widely produced in Himachal Pradesh. The biotic diseases are caused by the pathogens such as the viruses, bacteria and fungus. Scab, Cedar rust, leaf blotch, Powdery mildew, blight, Black rot are some of the visible apple plant diseases. Diplodia seriata fungus causes the black rot fungal infection. This infection is visible as frog eye spots along with reddish or purplish edges on top of the leaf. Venturia inaequalis fungus causes Scab infection. The symptoms include pale yellow or olive green spots on top of the leaf. It leads to curling of the leaves. The severity in the disease leads to damage of the tree. Cedar rust is caused by Gymnosporangium juniperi-virginianae fungus. Maize plant disease include Maize Rust caused by Puccinia sorghi. The symptoms include rust coloured lesions on top of the leaves. Maize Leaf Blight caused by Bipolaris maydis Fungi and Xanthomonas vasicola bacteria. Sypmtoms include dark irregular lesions on the leaves. Maize stalk Rot caused by Fusarium spp and Gibberella spp which leads to falling of all the leaves. Maize Ear Rots caused by Aspergillus spp and the final disease is Maize Downy Mildew caused by Peronosclerospora.

The impotence of the detection and prediction of the plant diseases have resulted in the destruction of the crop plants. There are many methods to manually identify plant diseases that can involve the processes like observation, symptom recognition, monitoring the periodical growth of a plant and these can be compared with other plants based on the similarities. There are many field tests are done to observe how fast the disease spreads throughout the plant in a particular period of time. The plant's environment is also monitored to identify certain environmental conditions like temperature and humidity changes. For further more crucial analysis, various laboratory tests are done where the diseased plant or the leaf is sent to a lab. These plants are monitored for various DNA analysis or more like microscopic evaluations to identify the specific disease-causing agent. Even though manual analysis on plant is done for the identification for the disease- causing specimen, still we need improve the accuracy and efficiency of the plant disease diagnosis that can prevent the disease and help towards the growth of the plant. These manual techniques only provide limited accuracy that only identifies the certain deficiencies, environmental causes or the pesticides involvement in the damage of the plant. This can also be very time consuming in times where in the meantime the plant can literally die before identifying the effective treatment for the disease. The laboratory tests or the manual consultation with an expert in this field can file a comparatively higher costs than developing models to identify the same.

The idea of using Machine learning models is that it provides various processing techniques and ensemble models that improves the accuracy and efficiency of the analysis of the diseased plant. This employs the Convolutional Neural Network (CNN) architecture in a hybrid approach. The Local Binary Pattern (LBP) mainly works with CNN where the key features and identified and merged. The ensemble models involved are the three main meta-heuristic algorithms which are Binary Dragonfly algorithm (BDA), Ant Colony Optimization (ACO) and Moth Flame Optimization algorithm (MFO). The image of the diseased plant is recognised by DenseNet and this DenseNet model is trained on various training images of the leaf diseases and then used to be tested on fresh leaf images for a real time evaluation. The next step is to focus more on the damage on the leaf and this is identified the RGB pixel value combinations which can be easily extracted from the images of the diseased leaf and now these pixel values and CNN is used for the supervised training on these RGB pixel values which provides the accuracy approximately 97% which is the evaluation that is done at pixel-level granularity of the leaf. So many diseases among the plants can have similarities and thus the diseases can be easily identified and proper treatments are provided that takes the consideration of both the plant's disease and the damage it caused on the plant. On of the important pre-trained architecture is the AlexNet that employ machine learning algorithms for identifying and classification of the images of the plants. We use a hybrid approach that improves the precision and robustness of the plant disease classification. AlexNet is mainly used to identify the key features thus it needs to be trained efficiently on the ImageNet dataset to get effective results on the identification of the fresh diseased images of the plant.

The proposed machine learning architectures like CNN can employ the need with high performance Graphical Processing Units (GPU) which is very expensive and may or may not be available by the time we need it and the DenseNet and AlexNet can require large models and thus it can be difficult when we try to deploy in the production environment. When an image is processed that has extensive lighting or exposed to light can easily cause the model to overfit. This can lead to an unexpected performance of the model when it is tested with the fresh leaf images that differs from the attributes that were identified during the training phase. As there are no limitations in the types of diseases on a plant, the training and testing of the model requires huge amount of labelled datasets that provides the diverse information of the disease and thus it can also be time consuming and expensive.

# II.LITERATURE SURVEY

In [[1]] it mainly focuses on addressing the challenges of identifying plant diseases without requiring many images to detect and it uses a Convolutional Neural Network(CNN) to reduce computational complexity. By training the models on the public datasets, it can easily identify the Scab, Black rot, and Cedar Rust diseases in apple plants with an accuracy of about 98%. Here the experimental results demonstrate that this model only requires very little storage and computational resources when compared to the existing CNN models.

The idea proposed in [[2]] focuses on using machine learning techniques and deep learning techniques for accurate detection ,classification and prediction of plant diseases. Here the model trains on the available datasets for plant disease detection and classification. It proposes eighteen classifications on the PlantDoc dataset. The results reveal that YOLOv5 achieves the highest accuracy for object detection, while the ResNet50 and MobileNetv2 supported an optimal trade-off in accuracy and training time for image detection.

This study [[3]] addresses the challenges of accurately predicting plant diseases using the deep learning techniques. It proposes an optimized framework called YR2S (YOLO-Enhanced Rat Swarm Optimizer), based on YOLOv7 that incorporates pre-processing and optimization techniques. This framework also utilizes ShuffleNet with ERSO for classification and PCFAN for map generation optimization. For the disease-prone area segmentation, FCN-RFO is employed. It develops a model that trains and evaluates a customized dataset that contains images of various plant leaves with diseases with an accuracy of 99.69%.

The review [[4]] emphasizes the importance of agriculture as a primary source of income and it also highlights the significant importance of plant disease detection. Many researchers employ computer vision, deep learning, few-shot learning, and soft computing techniques to automatically identify the diseases of the plants to ensure crop security in terms of quality and quantity. By leveraging all these techniques, researchers can be able to overcome the challenges related to feature selection and extraction and also enhance the speed and accuracy of the detection of disease.

An Ensemble hybrid framework [[5]] addresses the vital role of effective agriculture management by providing improvements in agricultural technology to improve production and quality. For the well-being of farmers and agriculture, early detection and classifications of plant diseases affecting the plant has been implemented. It employs ensemble deep learning deep learning-based techniques incorporating a hybrid approach with preprocessing algorithms, texture features, and feature extraction. With 99.8% accuracy, this research contributes to enhancing agricultural security and sustainability by providing an effective solution for the quick detection of plant diseases.

In [[6]] the challenge of more early detection of plant diseases using high accuracy techniques that focus on near-infrared spectroscopy and hyperspectral imaging. In the case of early blight disease in potato cultivars, practical corrections are used to address the subclass effect, kinetic effect, and measurement effect of the disease. It also employs the application of EPO-PLSDA (Extended Orthogonalized Partial Least Squares Discriminant Analysis) that improves the orthogonalizing of the model regarding temporal variation to produce invariant models. With this approach, early blight disease can be detected as early as 36 hours after the inoculation for six of seven testes cultivars.

The study[[7]] addresses the challenges of detecting plant disease with different deep-learning architectures and tests for each analysis. It utilizes the Plant Village dataset that consist of 20,640 images representing more classes and species which are pepper, potato, and tomato. The dataset is divided for training, testing, and validation. The performance of the Adam optimizer is compared with Sgdm optimizer and Hyperparameters such as several epochs and learning rate are optimized. The accuracy of eight deep learning architectures is evaluated at different minibatch sizes like 16 and 32.

In [[8]] it emphasizes the significant factor contributing to low crop yields is diseases caused by bacteria, fungi, and viruses. This can be prevented and managed through various machine learning approaches. Overall, leveraging effective deep-learning techniques for leaf disease detection from the captured images can help prevent major crop losses and improve agricultural productivity. It mainly focuses on deep learning techniques that offer promising methods for disease identification in plants by leveraging data-driven approaches.

Automatic blight disease detection [[9]] concentrates on the early detection of blight diseases in potato and tomato crops, which often lead to significant losses for farmers. After using the “Plant Village Dataset”, the evaluation on the test set containing enormous amount images, the model achieves impressive performance metrics with an accuracy of 99.25%. It also provides detailed explanations of the model’s prediction through saliency maps, shedding light on the reasoning behind the decisions. It also emphasizes the use of ResNet-9 model that considers various factors like leaf shape, diseased areas and general green areas to make predictions that provides the understanding of the model’s behaviour.

The study [[10]] emphasizes the importance of remote sensing tools and machine learning(ML) approaches in crop monitoring and disease surveillance, particularly in applications like crop type classifications and the early detection of diseases that warn the farmers. It monitors using large landscapes with Unmanned Aerial Vehicles(UAV) presents challenges but it also combines high-resolution satellite imagery with advanced ML models through mobile applications that improve the banana plant disease detection and classification and provide comprehensive health status information.

In [[11]] the author highlights the significant economic losses caused by the maize diseased plants in the agricultural field. The author finds out that the convolutional diagnostic methods are time consuming processes and are significantly costly. Hence introduces classification models that are used for identifying various maize leaf diseases like blight, gray leaf spot, common rust, etc. The model makes use of the features that are extracted using Densenet201, CNN, Support vector machine. The SVM is fine tuned using the Bayesian optimization Techniques.

This paper [[12]] addresses the challenges that are gone through while predicting the plant diseases in the Tomato plant. The Tomato plant diseases are detected using Machine learning techniques and image processing techniques. As a preprocessing technique the images of the leaves are subjected to Histogram Equalization for enhancing the quality of the pictures. In the next step the K-means Clustering technique is used for segmenting the leaf samples. The important features are separated for the leaf samples using the Discrete Wavelet transform (DWT), Principal Component Analysis (PCA), and grey level Co-occurrence Matrix (GLCM). For classification, the model makes use of the SVM, K-Nearest Neighbour and Convolutional Neural Network (CNN).

The central thought of this paper [[13]] revolves around the detection of the plant disease in tomatoes and grapes. The detection of these diseases is done mainly using the machine learning techniques. The model uses Convolutional Neural Network (CNN) for image recognition Visual Geometry Group (VGG) for improved performance measures. This model increases the efficiency where it consists of several convolutional layers which are fully connected. The YOLOv3 is used for real-time-detection to improve speed and accuracy.

The research in [[14]] deploys a combination of transfer learning and Gravitational Search Algorithms for detection identification and classification of plant diseases. In transfer learning it uses CNN- Convolutional Neural Network specially MobileNetV2 and ResNet50V2. The models are pre-trained using large image datasets. The Gravitational search Algorithm is used to optimize the features that are extracted from the CNN models. It helps in extracting the most important and desired features that are required for classification. The selected features are now passed on to the Multinomial Logistic Regression (MLR). This is used for creating a relationship between multiple categories. This is mainly used for classification.

This paper [[15]] focuses on the detection of mango plant leaf disease .This includes the development and evaluation of the LeafNet which a type convolutional neural network for detecting diseases in the mango leaves. This includes the detection of diseases such as Anthrocnose, Powdery mildew, Bacterial Canker, Cutting Weevil and sooty Mould. The AlexNet is a deep convolutional neural network which has a high efficiency in image classification. VGG16 is characterized by its depth consisting of 16 layers and it is known for effectiveness and robustness.

It is very important to use the accurate models to provide high accuracy. [[16]] highlights the application of machine learning and deep learning models for detecting the sugarcane plant diseases. This research studies on convolutional diagnostic methods, Immunologic methods, Molecular Detection methods. The convolutional diagnostic methods include naked-eye visible symptoms isolation, culture of pathogens, morphological approaches, etc. The immunological methods include ELISA test for quantifying plant pathogens. The molecular detection methods includes real-time polymerase chain reaction, microarray, LAMP assays.

Mango has an important place amongst all the fruits in India. The paper [[17]] discusses about the various techniques that are used for detecting and evaluating the plant diseases. The Machine learning models are specifically used for automatic detection of the grape and mango plant diseases. The model utilizes AlexNet proposed by Krizhevsky et al., This is a pretrained model and has multiple layers such a convolutional layer, pooling layer, activation layer, connected layer, and Soft max layer. Support vector machine is also used, comparing with the Alexnet in some of the areas. InceptioV3, ResNet50, MobileNet, VGG16, InceptionResNetV2 are all other models that had been utilized in training the plant disease detection model. Difference in these models are the accuracies that they provide.

The significance or importance or need for plant disease detection has risen up. [[18]] This paper provides an insight about the novel approach for plant disease detection. It also integrates with the internet of things to provide a smart monitoring system. This model also combines the IOT with the machine learning techniques such as CNN. This approach does not use a pretrained model for classification and detection. As the name suggest, the smart monitoring system comprises of various components like LED-grow light fixtures to optimize the model with environmental parameters ang controlling lights. The use of light weight model leverages high predictive performance, handles complex backgrounds, variability in sizes and shapes.

The plant diseases started posing threats in the field of Food security. Hence the study [[19]] depicts the traditional method’s insignificance and challenging limitations faced and necessitates more efficient detection techniques. For inducing speedy and accurate decision making, the model that this paper highlights, utilizes advanced technologies like deep learning integrated with filtering methods. The filtering methods emphasize and complement the deep learning techniques by leveraging personalized treatment recommendations. Deep learning techniques that are particularly used are CNN, RNN which provides more accurate performances.

In this study [[20]] the researches investigates four deep learning models or frameworks for fine grained disease classification. It seamlessly handles the subtle and nuanced disease patterns. This model mainly concentrates on the maize plants. The four models include VGGNET, ResNet50 which achieves highest accuracy, Inception V3, InceptionResNetV2. These models are also enhanced using Deep transfer learning techniques. The ML and DL models used include the CNN, K-means clustering, SVM. Yang et al. Zhang S.and Zhang c. applied different methods for dimensionality reduction for SVM classification.

# III. Proposed model

## Methodology

The detection, classification and prediction of plant diseases include a customized machine-learning architecture comprised of advanced self-organized maps and a restricted Boltzmann Machine. The architecture concentrates on various levels such as image preprocessing that includes data resizing, grayscale conversion and data augmentation, model building, feature extraction, and testing phases on a given input. The architecture leverages the combined strengths of self-organized maps, an Artificial Neural Network for clustering and visualizing, and a restricted Boltzmann Machine, a generative Neural Network for unsupervised learning tasks. This fusion results in a robust and reliable plant disease classification across the apple and maize plants. A detailed description of each block of the architecture is explained below.

## Autoencoder and Restricted Boltzmann Machine

The Autoencoder is a method of Artificial Neural Network which used in unsupervised learning of the datasets. It consists of three main components, an encoder that converts the data into a latent space representation, a bottleneck layer, and a decoder that rebuilds the input data from the lower dimensional latent space representation including fully connected layers.

The noising techniques such as adding masking noise to the input data are incorporated to enhance the model’s robustness, where the percentage of each and every input’s elements is dropped to zero in every epoch. The noise levels such as 0%, 30%, 50%, and 70% are used. The loss function is minimized based on the similarity between the input data and the representation learned by the autoencoders.

The Learning rate, regularisation parameters, and optimized algorithms suitable for autoencoders are used. The training of the encoders includes unsupervised learning techniques where the aim is to reduce the error between the input data and the reconstructed data. The mean square error is a common loss function used in Autoencoders training. The weights of the encoder and the decoder network are optimized using algorithms such as stochastic gradient descent.

The restricted Boltzmann Machine (DBM), a type of generative neural network is used to capture complex hierarchical representations of plant leaf images. The DBM is designed with multiple layers of stochastic binary units and the connections are made between the layers. The DBM is trained using unsupervised learning algorithms contrastive divergence to adjust the weights between the layers. Contractive divergence involves approximating the gradient of the log-likelihood function to update the weights iteratively using Gibbs sampling. The training involves two main steps

1. Defining the visible node likelihood which uses an exponential energy function.

This results in the combined probability of visible and hidden nodes.

1. The negative log-likelihood is minimized using gradient descent including contrastive divergence and Gibbs sampling.

The energy function utilizes Gaussian noise, weights, and biases which are supposed to be learned during the training period to capture meaningful information from the input data. The Gaussian noise is incorporated into the model to learn more generalized features. The standard deviation of the Gaussian noise is fixed at a suitable value, say 1, such that the magnitude of the noise added to the input is controlled. The weight parameters are updated based on the difference between the expected values of the visible-hidden nodes and the model’s initial distribution. The presence and absence of the features associated with different diseases are captured by updating the bias parameters. Epochs are utilized during the training period where each epoch consists of forward and backward passes through the network. These are done to update the model parameters.

The DBMs are efficient in learning rich hierarchical representations of the unlabeled input data. It allows the model to capture complex patterns and structures present in the input data by increasingly abstracting features in each layer.

To enhance the prediction of plant leaf disease are dimension reduction process based on self-organized maps is put forward. The process od dimension reduction aims to convert the high-dimensional image data into a lower-dimensional space. This ensures the retaining the important features that are important for disease prediction. The main goal to reduce the data is to utilize it for the classification process. This classification process generates a classification map that accurately labels different types of plant leaf diseases present in the input dataset.

Identify applicable funding agency here. If none, delete this text box.

## Self organized maps**:**

The self-organized maps work based on the theory of the human brain’s ability to extract appropriate features from the environment. The specialty of SOM is that it consists of one input layer, one output layer, and no hidden layers. The number of neurons in the layers defines the structure of SOM. The neurons in the layer are connected by synaptic weights. These synaptic are adjusted during training to automatically detect relationships within the datasets. The neurons in the output layer compete to win the opportunity to interact with input, hence SOM training can be called as competitive learning. The weights of the winning neurons are only adjusted while other neuron's weights remain unchanged. The self-organized is mainly used to project the spectral bands of each pixel of the dimensionally reduced dataset. The architecture of the SOM is defined by the number of input neurons which equals to the number of features and the number of output neurons which is not preferably specified. The self-organized maps generate a synaptic weight matrix where each input vector is multiplied to generate output vectors. Hence the proportionality between the dimensionality detection of the self-organized maps is achieved.

## Algorithm

1. Load the input image X(N\*r) with N samples, each having r features.
2. Define the SOM grid with size as M, where M<r.
3. Initialize the weight matrix K, with M2 neurons, each neuron containing r weights.

K ij(M2\*r) =[K1,K2,K3,…..,Kr…….,KM2]

1. Define structure parameters σinitial, σfinal, σinitial, σfinal, fmax
2. Training loop:

f<fmax

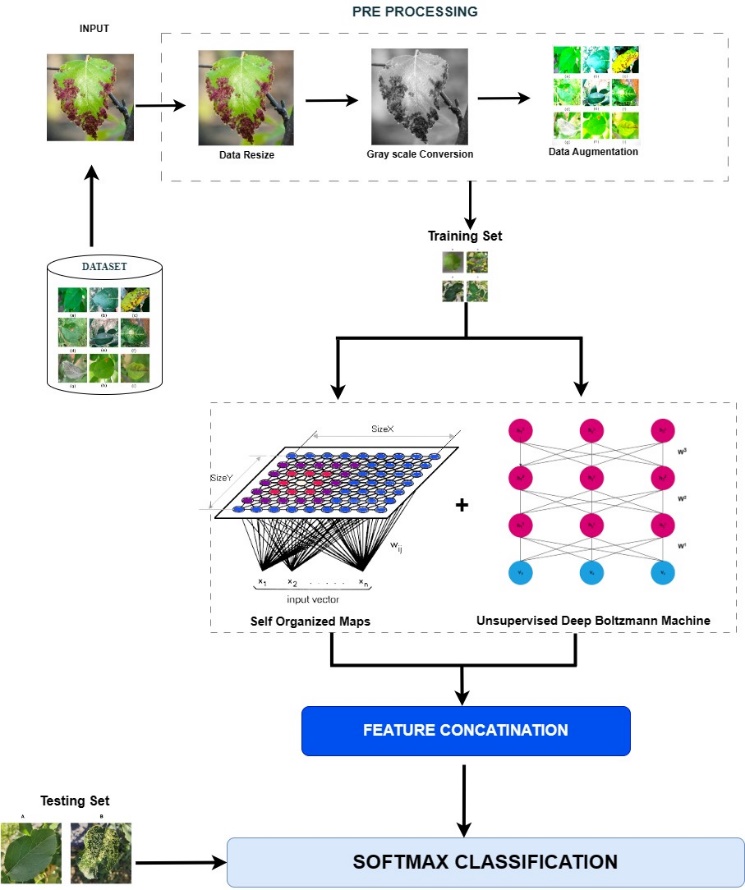
1. Select a random input sample Yi from Y
2. Find the winning neuron where W=argmin||yi -Kij||
3. Update the weights for all neurons based on the winning neuron and its neighbour.
4. For each neuron S= 1 to M2
5. Update learning rate and neighbourhood radius ΔKij= α(f)gjs(f)(Xi – Kij)
6. Increment iteration counter:
7. End the training loop.
8. Output the reduced dimensional data:

X’ = W\*X

1. End algorithm.

The algorithm starts by loading the input image with N sample having r features which defines a self-organized map with M\*N neurons. Initialize the synaptic weight matrix K randomly, where i=1 to M2  and j= 1 to r.

Define control parameters and set the initial and final values for the learning rate. Set the initial and final values for the neighborhood radius. Specify the maximum number of iterations. The learning rates vary from σinitial to σfinal. The aim is to make the neurons fit quickly to the input data. The calculated neighborhood function is a Gaussian function. Initialize the iteration counter f=0, while f<fmax .Randomly choose the input sample and compute the index of the winning neuron. Update the synaptic weights. Update the learning rate and the neighborhood radius. Increment f by 1. Generate the reduced dimensional data by multiplying the final weights with input data. In the training loop, the result is the input data compressed to lower dimensional space.



# IV.COMPARISION AND DISCUSSION

This methodology utilizes the apple and maize plant’s dataset to identify and categorize the plant leaf diseases. The input dataset is processed under various levels such as preprocessing , model building, feature extraction, and testing phases. The preprocessing step includes data resizing, gray scale conversion, and data augmentation. The combined utilization of the self-organised maps and unsupervised deep Boltzmann machine make the model more efficient in predicting the leaf diseases. The fusion of these machine learning models makes the result more robust and reliable .

This survey presents the comparison of the proposed model and all other existing models mainly to highlight the significance of the proposed model. The methods are deliberately selected from various research to make the comparison. Most abundantly used pretrained models are the CNN models. For instance in [1] which has utilized the existing CNN model leveraging augmentation techniques such as shifting, shearing, scaling, zooming and flipping has achieved an accuracy of 98%. This accuracy is achieved for identifying and predicting Scab, Cedar rust and black rot diseases that affects the apple leaves. This underscores the effectiveness of the CNN model. The research [11] concentrates on the DenseNet201 which is again a powerful Convolutional Neural Network and also Support vector machines. The CNNs are proven to be the most effective one in the image classification. The densenet201 mainly concentrates on the dense connective networks. By deploying the Bayesian optimization technique to the Support vector machines to fine tune the model adds more efficiency to the existing model. This model combing the Densenet201 and SVM has achieved an accuracy of 94.6% in detecting and predicting the diseases such as blight, common ruts m gray leaf spot of the maize leaves. The other paper [20] provides an accuracy of 9.5 which has utilized the VGGNET ,Inception V3, ResNet50 and InceptionResNetv2 are used and trained using the dataset . This model is then utilizes the supervised learning technique which may provide less accuracy and may lead to overfitting or underfitting. Another paper provides an overall accuracy of 96% where it mainly uses the Resnet-9 model for the detection of early and late blight disease in potato and tomato leaves. The model utilizes the Plant Village dataset which has both healthy and diseased plant leaves of potato and tomato. This model extracts the meaningful features from the input images and also employs the parameter tuning by adjusting the necessary and unnecessary parameters.

Similarly in [2] the algorithms that are used are AlexNet, LeNet, VGG, InceptionV3, ResNet50, Logistic Regression Analysis, Support Vector machines, MobileNetv2 for classification process. For the object detection it utilizes You only Look Oncev5.The object detection aims at predicting both the type of disease and the diseased regions with the plant images. This model provides an accuracy of 97% due to the ultilization of the object detection algorithm YOLOv5 and MobileNetv2, ResNet50 classification algorithms.

The accuracy of the proposed model has reached upto 99.8%. It is evident that the propsed model has obtained the highest accuracy of all the papers that are compared with.Ou proposed model leverages more advantages than any other model. By combining the diverse models such as self-organized maps and a restricted Boltzmann Machine that proposed model has achieved the highest accuracy with more precision, recall, F-1 score. These metrics indicate that the fusion of the models are more effective in correctly classifying images into respective diseases and proving high performance in prediction, minimizing false positives and false negatives. The incorporation on different and diverse techniques that utilizes types of artificial and generative neural networks to handle complex hierarchical representations have made the model more robust and reliable.This approaches removes the risk of overfitting or underfitting of testing and training data .Furthermore our approach in classifying handles the data inconsistencies that are caused by the outliers.However it is necessary to look through the limitations that this methodology provides. The SOMs and DBNs are complex models which consists of multiple layer of neurons. Which can require high computational resources and time. Understanding how the features learned by SOMs or DBNs relate to specific plant diseases or environmental factors may not be straightforward. Ensuring stable and robust training of DBNs can be challenging, particularly in the context of limited labeled data for plant disease detection. Despite all the limitations , the SOM and DBMs are capable of learning the complex non-linear relationships in the data. These are mainly suited for unsupervised learning tasks where it can discover intrinsic structures and features. SOMs can perform effective dimensionality reduction by mapping high-dimensional input data onto a lower-dimensional grid. This reduces the computational complexity of subsequent analyses and visualization tasks, making it easier to interpret and explore the data.

The Table 1 illustrates different diseases and also healthy leaf of Apple and Maize. Each disease and the healthy leaf are associated with different classes. These are done for classification purpose. Say Apple\_Black\_rot is associated with Class or label B. So when the model identifies Apple\_Black\_rot it classifies it to the B labelled class. In the same way if the model identifies Corn\_smut it classifies it to the W labelled class. The detailed classification of the diseases and their labels are mentioned in the Table 1.

| CLASS | TABLE 1: Apple and Maize plants classes and labels | | |
| --- | --- | --- | --- |
| LABEL | CLASS | LABEL |
| Apple\_Black\_rot | B | Corn\_smut | W |
| Apple\_healthy | D | Corn\_stallk\_rot | X |
| Apple\_LeafBlotch | V | Corn\_Diplodia\_Ear\_Rot | Y |
| Apple\_powdery\_mildew | T | Maize\_Southern\_Rust | Z |
| Apple\_Alternaria\_Leafspot | U | Corn\_Goss’sWilt | AB |

The efficiency and the performance of the proposed model has been calculated and evaluated using certain performance metrics. These metrics are calculated, compared, contrasted and has been recorded in the Table 1 with experiment 1 and 2.

Efficiency is calculated to depict how often the model predicts the output correctly and consistently.

Efficiency = (Actual Output/Standard output) \*100

Sensitivity= TP/TP + FN

Precision = TP/TP + FP

Specificity = TN/TN + FP

Where TP is the True Positive, FN is the False Negative, FP is the False Positive, TN is the True Negative

F1 Score= 2\*Precision\*Recall / Precision + Recall

FPR is the False Positive Rate which measures the false positives among the actual positives.

FNR is the False Negative rate which measures the false negatives among the actual negatives.

These measures are calculated to evaluated the proposed model which uses the Self organized maps and unsupervised Deep Boltzmann machine to predict the output.

Each label is represented as identifier which is used for identifying the diseases that are proposed to the model and the model recognizes these labels and classifies them to the respective classes .The below table is a structured way of categorizing and identifying the different types of diseases.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Classifier | TABLE 2: Experiment 1 | | | | | | |
|  | Acc (%) | Sen (%) | Pre (%) | Spe (%) | F1 (%) | FPR (%) | FNR (%) |
| AutoML | 98.30 | 98.35 | 98.36 | 99.88 | 98.35 | 0.11 | 1.65 |
| Reinforcement Learning(RL) | 98.80 | 98.83 | 98.83 | 99.92 | 98.83 | 0.08 | 1.17 |
| LightBGM | 98.70 | 98.73 | 98.74 | 99.91 | 98.73 | 0.09 | 1.26 |
| XGBoost | 97.10 | 97.27 | 97.36 | 99.80 | 97.27 | 0.19 | 2.73 |
| Graph Neural Networks(GNNs) | 90.80 | 90.84 | 91.04 | 99.35 | 90.90 | 0.65 | 9.16 |
| Self-Supervised Learning | 44.80 | 44.87 | 74.68 | 96.06 | 42.69 | 3.93 | 55.1 |
| Spatial Pyramid Pooling(SPP) | 89.30 | 89.28 | 90.13 | 99.23 | 89.21 | 0.76 | 10.7 |

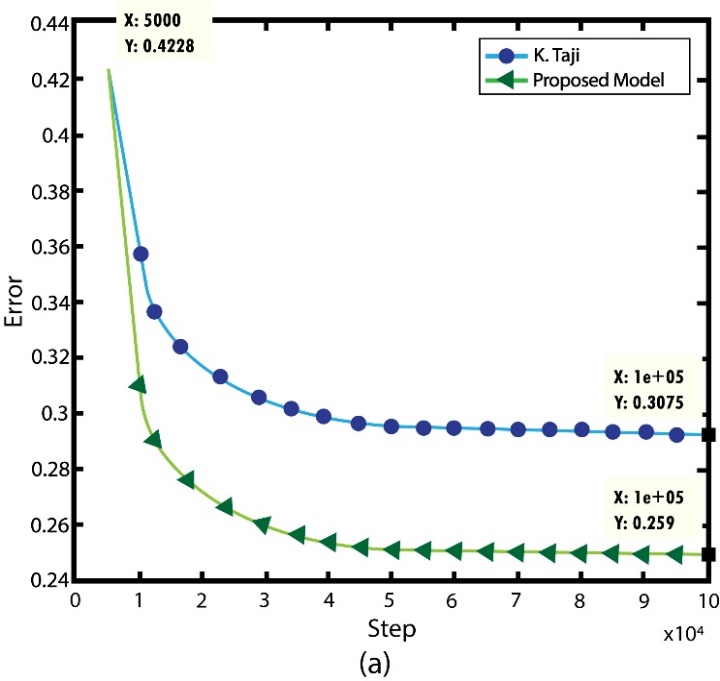
The below table represents the second experiment of the same model but a modified model based on the previous experiment. It also provides the performance measures to

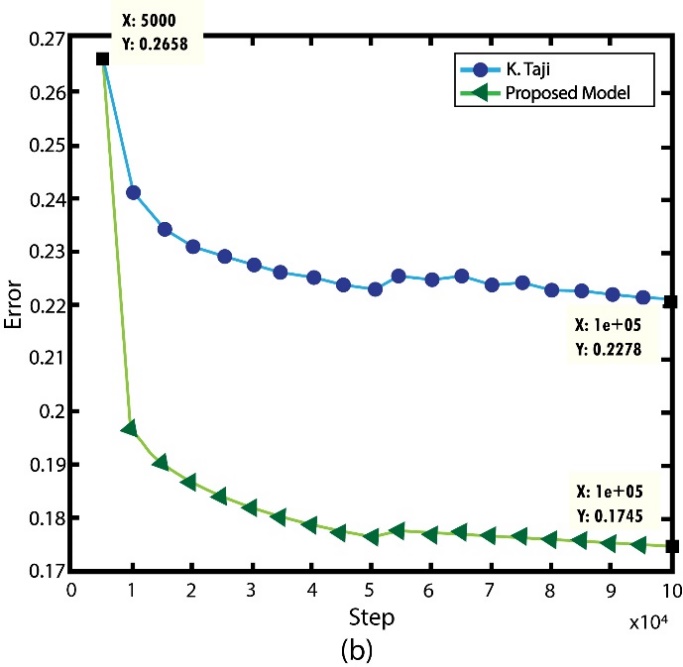
and the comparison can be made with the previous one.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CLASSIFIER | TABLE 2: Experiment 2 | | | | | | | |
|  | Acc  (%) | Sen (%) | Pre (%) | Spe (%) | FPR (%) | F1 (%) | FNR(%) |
| AutoML | 88.30 | 91.35 | 91.36 | 99.88 | 94.35 | 0.118 | 1.653 |
| Reinforcement Learning(RL) | 87.80 | 95.83 | 93.83 | 99.92 | 95.83 | 0.084 | 1.173 |
| LightBGM | 77.70 | 91.73 | 91.74 | 99.91 | 78.73 | 0.090 | 1.267 |
| XGBoost | 76.10 | 92.27 | 94.36 | 99.80 | 97.27 | 0.195 | 2.733 |
| Graph Neural Networks(GNNs) | 89.80 | 88.84 | 96.04 | 99.35 | 90.90 | 0.654 | 9.160 |
| Self-Supervised Learning | 42.80 | 41.87 | 74.68 | 99.06 | 42.69 | 3.938 | 55.13 |
| Spatial Pyramid Pooling(SPP) | 86.30 | 84.28 | 97.13 | 99.23 | 89.21 | 0.766 | 10.72 |
|  |  |  |  |  |  |  |  |

The visualization to analyze the efficiency of the Restricted Boltzmann Machines(RBMs) for plant leaf disease detection is shown in the above graphs. The Learning progression and the error reduction have been drawn. The X-axis is the Learning steps that have been calculated with the experiment data with the help of the experiment table data. The Y-axis is the Error Rate which could be the mean-square error, log-likelihood, or any appropriate metrics that would depend upon the goals of the RBM. The graph visualization of the graph represents the Restricted Boltzmann Machine (RDBM) which is the proposed model and the existing proposed model. The dotted line is the traced representation of the existing model and the arrowed line is the representation of the proposed Restricted Boltzmann Machine. It highlights the capability of the RBMs in the extract and retain information about the features from the plant leaf data, possibly reducing the dimensionality significantly while maintaining and improving the model’s performance. It also traces and compares how well the error rates evolve around the data throughout the learning process for the model. It coveys about the importance of how quickly the RBM model covers the error rates when compared to the existing models.

The graph mainly compares the errors of the proposed model and the existing baseline method. We can clearly see that until 5000 steps both the models use all the same features so that they reach the same error levels which is 0.4228 and 0.2658 on both the training and the evaluation sets. The proposed model reaches an approximate error level 0.1825 with the training data which is lower than the base method of the execution set. The errors for the proposed model and the existing model RBM are 35% and 14% lower than the error which is at step 5000 which is clearly visible in the graph above. The extracted feature selection removes almost every redundant and the irrelevant features od the data and subsequently prevents the overfitting of the data set which is useful when the model is trained on the testing dataset. This proposed model can find upto 775 repeated and irrelevant features of the given data and thus it removes the features contrast about 20% of the initial 3870 features.





##### V. CONCLUSION

Plant leaf disease detection and classification greatly impact various factors in the agricultural sector, reduce crop destruction, and control financial losses and food shortages thereby increasing food quality. The traditional methods to manually identify plant diseases can involve processes like observation, symptom recognition, and monitoring the periodical growth of a plant and these can be compared with other plants based on the similarities. Even though manual analysis of plants is done for the identification of the disease-causing specimen, still we need to improve the accuracy and efficiency of the plant disease diagnosis that can prevent the disease and help towards the growth of the plant. The employment of machine learning models enhances precision agriculture by targeting disease prediction and management and ensuring efficiency in the agricultural field. The study emphasizes the analysis of the plant diseases that affect the apple and maize plants, the utilization of this proposed technique yields an optimizing accuracy. The plant disease classification is accurate and involves a timely optimized output that improves crop health, increasing yields and the loss of the plants. The accuracy of the proposed model has reached up-to 99.8%. We believe that this study would provide additional innovation in agricultural technology as it employs effective agricultural techniques that produce optimized outcomes in real-time.

##### VI.References

…

[1] 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,” *IEEE Access*, vol. 11, pp. 6594–6609, 2023, doi: 10.1109/ACCESS.2022.3232917.

[2] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, “Machine Learning and Deep Learning for Plant Disease Classification and Detection,” *IEEE Access*, vol. 11. Institute of Electrical and Electronics Engineers Inc., pp. 114352–114377, 2023. doi: 10.1109/ACCESS.2023.3324722.

[3] C. Madhurya and E. A. Jubilson, “YR2S: Efficient Deep Learning Technique for Detecting and Classifying Plant Leaf Diseases,” *IEEE Access*, vol. 12, pp. 3790–3804, 2024, doi: 10.1109/ACCESS.2023.3343450.

[4] A. Bhargava, A. Shukla, O. P. Goswami, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, “Plant Leaf Disease Detection, Classification, and Diagnosis Using Computer Vision and Artificial Intelligence: A Review,” *IEEE Access*, vol. 12, pp. 37443–37469, 2024, doi: 10.1109/ACCESS.2024.3373001.

[5] K. Taji, A. Sohail, T. Shehzad, B. Shoaib Khan, M. A. Khan, and K. Ouahada, “Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000. An Ensemble Hybrid Framework: A Comparative Analysis of Metaheuristic Algorithms for Ensemble Hybrid CNN features for Plants Disease Classification”, doi: 10.1109/ACCESS.2017.DOI.

[6] F. Abdelghafour *et al.*, “Including measurement effects and temporal variations in VIS-NIRS models to improve early detection of plant disease: Application to Alternaria solani in potatoes,” *Comput Electron Agric*, vol. 211, Aug. 2023, doi: 10.1016/j.compag.2023.107947.

[7] S. Dahiya, T. Gulati, and D. Gupta, “Performance analysis of deep learning architectures for plant leaves disease detection,” *Measurement: Sensors*, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100581.

[8] C. Jackulin and S. Murugavalli, “A comprehensive review on detection of plant disease using machine learning and deep learning approaches,” *Measurement: Sensors*, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100441.

[9] A. O. Anim-Ayeko, C. Schillaci, and A. Lipani, “Automatic blight disease detection in potato (Solanum tuberosum L.) and tomato (Solanum lycopersicum, L. 1753) plants using deep learning,” *Smart Agricultural Technology*, vol. 4, Aug. 2023, doi: 10.1016/j.atech.2023.100178.

[10] M. Gomez Selvaraj *et al.*, “Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 110–124, Nov. 2020, doi: 10.1016/j.isprsjprs.2020.08.025.

[11] A. Dash, P. K. Sethy, and S. K. Behera, “Maize disease identification based on optimized support vector machine using deep feature of DenseNet201,” *J Agric Food Res*, vol. 14, Dec. 2023, doi: 10.1016/j.jafr.2023.100824.

[12] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, “Plant leaf disease detection using computer vision and machine learning algorithms,” *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, Jun. 2022, doi: 10.1016/j.gltp.2022.03.016.

[13] A. S. Paymode and V. B. Malode, “Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG,” *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, Jan. 2022, doi: 10.1016/j.aiia.2021.12.002.

[14] M. S. A. M. Al-Gaashani *et al.*, “Deep transfer learning with gravitational search algorithm for enhanced plant disease classification,” *Heliyon*, vol. 10, no. 7, Apr. 2024, doi: 10.1016/j.heliyon.2024.e28967.

[15] R. A. Rizvee *et al.*, “LeafNet: A proficient convolutional neural network for detecting seven prominent mango leaf diseases,” *J Agric Food Res*, vol. 14, Dec. 2023, doi: 10.1016/j.jafr.2023.100787.

[16] S. Strachan, S. A. Bhuiyan, N. Thompson, N. T. Nguyen, R. Ford, and M. J. A. Shiddiky, “Latent potential of current plant diagnostics for detection of sugarcane diseases,” *Current Research in Biotechnology*, vol. 4. Elsevier B.V., pp. 475–492, Jan. 01, 2022. doi: 10.1016/j.crbiot.2022.10.002.

[17] U. Sanath Rao *et al.*, “Deep Learning Precision Farming: Grapes and Mango Leaf Disease Detection by Transfer Learning,” *Global Transitions Proceedings*, vol. 2, no. 2, pp. 535–544, Nov. 2021, doi: 10.1016/j.gltp.2021.08.002.

[18] W. Macdonald, Y. A. Sari, and M. Pahlevani, “Grow-light smart monitoring system leveraging lightweight deep learning for plant disease classification,” *Artificial Intelligence in Agriculture*, vol. 12, pp. 44–56, Jun. 2024, doi: 10.1016/j.aiia.2024.03.003.

[19] F. O. Isinkaye, M. O. Olusanya, and P. K. Singh, “Deep learning and content-based filtering techniques for improving plant disease identification and treatment recommendations: A comprehensive review,” *Heliyon*, vol. 10, no. 9, p. e29583, May 2024, doi: 10.1016/j.heliyon.2024.e29583.

[20] I. Khan, S. S. Sohail, D. Ø. Madsen, and B. K. Khare, “Deep transfer learning for fine-grained maize leaf disease classification,” *J Agric Food Res*, vol. 16, Jun. 2024, doi: 10.1016/j.jafr.2024.101148.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**