A NEW APPROACH FOR PLANT DISEASE DETECTION USING CONVOLUTIONAL NEURAL NETWORK AND CONVOLUTIONAL AUTOENCODER

Dr. B. Rebecca
Associate Professor, Dept. of Computer Science and Engineering (Data Science), Marri Laxman Reddy Institute of technology and management,Dundigal, Hyderabad, Telangana , INDIA, 500043.
EMail id: baderebecca@yahoo.co.in

P. Siva Padmini
Assistant Professor, Dept. of Computer Science and Engineering (Data Science), Marri Laxman Reddy Institute of technology and management, Dundigal, Hyderabad, Telangana , INDIA, 500043.
Email: sivapadmini83@gmail.com

Dr V Sangeetha
Dept of Computer Science and Engineering
MRCET,Maisammaguda, Medchal, Hyderabad Telangana , INDIA-500100
Mail ID: sangeetha12@gmail.com

Wankhede Mahendra Pandhari
Dept. of Computer Engineering, Bharati Vidyapeeth (Deemed to be University) .
College of Engineering, Pune, Bharati Vidyapeeth Campus, Maharashtra, INDIA- 411043
EMail ID: wankhede.mahen@gmail.com

Abstract:

Detection of plant disease is crucial for efficient and precise disease prevention in a complex environment. Enhanced decision support, astute analysis, and strategic planning are made possible by the rapidly developing field of "smart farming," in which the process of identifying plant diseases is digitized and data-driven. A plant's vulnerability to several diseases increases as it ages. One of the most challenging tasks in agriculture and in life in general is the detection of plantdiseases at an early stage. Productivity drops may occur if an illness is not promptly detected and treated. If infections are not detected immediately, farmers' incomes might be impacted. Researchers have employed deep learning and other machine learning methods to address this issue. Similar systems have been used in testing. Most of these algorithms are either very difficult to implement or need several tuning knobs. They are equally incorrect. This approach uses a CNN andCAE (Convolutional autoencoder) network to develop a tailored model for diagnosing plant diseases. Using a CAE and CNN system to detect plant diseases is not yet possible with state-of-the-art technology. Extensive research confirms this to this day. Bacterial Spot was discovered using this study and the model described here. In any case, this method is utilized to detect plant illnesses. Using a free online photo sharing service, we compiled images of leaves from various plants. The proposed approach yields a remarkable 99.59% training accuracy and 98.77% testing accuracy across a total of 9,714 parameters. Compared to traditional models, the hybrid one under investigation needs fewer parameters for training. Everything we know about the world's workings contradicts this. This work's suggested deep learning technique is, thus, essential for sustainable agriculture, efficient food production, and the preservation of natural resources.

Keywords: ***Classification, Extraction, Convolutional Neural Networks, CAE, Image Segmentation***

1. Introduction:

Plant disease may cause stunted development, which has negative consequences on production [1-3]. Worldwide economic losses are predicted to reach $20 billion annually [4-6]. Due to regional variances that might make an accurate identification more difficult, researchers face their most challenging obstacle when studying diverse situations [7, 8]. Traditional approaches also depend heavily on experts, expertise, and manuals [9], but the bulk of them are costly, labor- intensive, and difficult to detect accurately [10]. Therefore, a quick and reliable method of identifying plant diseases is necessary for agriculture's economic and ecological well-being. The use of multimodality data from multiple sensors, suchas the Internet of Things and sensor networks, has quickly advanced in internet technologies [11]. Here, a brand-new deeplearning-based plant leaf recognition model is created to address the aforementioned problems. Leaf retrieval, picture

segmentation, and identification are all included in the function, and the integrated deep learning algorithm is used throughout the whole process.

Agriculture contributes significantly to India's GDP (GDP). 10% of India's GDP comes from food and agricultural exports[9]. 75% of India's population depends on farming. High-quality, disease-free agriculture drives economic growth. Out- of-phase plants are vulnerable to pests and diseases. This policy change hurts farm production and profit. Plant diseases require early diagnosis. Farmers or experts inspect plants for pests or disease.
Global research efforts have produced cutting-edge technologies (Ashraf and Khan, 2020[17]). Many of these systems require a powerful computer or a lot of user expertise. We test the hypothesis that omitting variables from the CAE network affects plant disease classification accuracy. Eliminating prerequisites halved training time.
The brain's neural network architecture can inform deep learning. ANNs and variants can be used to uncover hidden data. Artificial neural networks are linked to recurrent and convolutional networks (ANN). Deep Learning helps analyze large, complex datasets. Image data allows for many approaches. Image convolution can extract temporal and spatial information. CNNs categorize input images while CAEs reduce them. This research suggests combining methods to quickly identify plant diseases. It uses CAE and CNN, two cutting-edge training methods.

Our research is the first to show how CNN and CAE may work together. Obviously, we are aware of everything that is going on. Numerous articles have been written on using machines to detect plant diseases. By using CAE, the dimensionality of the model may be reduced and the associated training parameters can be lowered as well. These kinds of simulations are possible using CAE. Pin that can detect the presence of microbes The hybrid model proposed is analyzed by looking at health problems. Xanthomonas cerecospora is responsible for this illness.
The report is divided into four sections for clarity and convenience. In Section 2, you can find information on the most up-to-date, fully or partially automated systems for determining which plant diseases to treat. Section 3 describes the workingcomponents, materials, and methods of this work and delves further into the origins of the hybrid framework. Section 4 describes the proposed methodology. Section 5: Model data is used to detect and diagnose the microbial illness Pinpointin peach plants. We must now wrap up Section 6 with a conclusion and future scope.

2. Literature Review:

All ResNet prototypes that existed prior to the VGG-16 were taken into account.ResNet-152 has an unsurpassed 99.2% accuracy percentage when compared to other networks. By taking images of the leaves with a smartphone and uploading them to an app, it's feasible to find banana plant disease quickly. The InceptionV3 model allowed this software to forecast illness with a 98.79 percent accuracy rate[5]. They needed 60 million parameters to train ResNet-152, their top- performing model.

Plant Village was analyzed using cutting-edge CNN architectures and classifiers to look for feature extraction using VGG-16, ResNet-50, and Google NET. SVM and kNN classified it (SVM). ResNet-50-trained SVM outperformed competitors,study authors said (98 percent). ResNet-50 reportedly has 25 million training parameters. Tiwari's employees were in agreement (Tiwari et al., 2020) For the automated identification of illnesses in potato crops, many approaches were put forward. VG-19, InceptionV3, Logistic Regression, k-Nearest Neighbors, and

Support Vector Machines (SVM) are only a few of the CNN designs that were used for disease diagnosis (NNs). They discovered that VGG-19 had a success rate of 97.8% using Logistic Regression. We categorise images of tomato, potato, and maize leaves into six different groups (i.e., infected and healthy). During training, they employed a model with a flawless accuracy of 100 percent, and during testing, it had an accuracy of 86.78 percent. In contrast to the 9,914 parameters employed in their study, the suggested system has 3.3 million total training parameters.Agricultural illnesses that manifestthroughout certain seasons have been identified using CNN and the autoencoder.

New CAE-CNN hybrid approaches are being developed. Khamparia et al. believe their model offers more reliable test results (Khamparia et al., 2020). Liu, Bin, et al. used DCNs to diagnose apple leaf disease. Liu proposes a unique deep convolution network model in this paper [10]. The study's approach for detecting character swaps is excellent. Principal component analysis and oscillation yielded 13,689 images. We also advise extending AlexNet using the NAG Algorithm. Future research may employ FCNN, RCNN, and SSD to predict apple leaf disease. This research [8] discusses leaf data extraction and disease categorization. For predictive modeling, convolutional neural networks (CNNs) are utilized. Instructions include training photographs, pre-testing images, image augmentation, CNN deep and optimizer training. By studying these photos, we can determine their processing technique and distinguish between plant illnesses. This research [10] uses thermal, digital, and hyperspectral imaging to categorize foliar diseases. We segment to concentrate on this goal's important parts. This approach removes outside effects [15]. The threshold value separates color and black-and- white photos. Matrix usage include feature extraction, associated values, and histogram intensities. Artificial neural networks can predict a disease's winter recurrence.

The use of the vector engine consistently produces better results for all sorts of images, and vector networks and machineshave a wide range of applications. The research cited above all had the same flaw: they all made use of training environments. Another key consideration is how much time and resources your computer will need to train a truly complex model with many training parameters. In order to improve the accuracy of findings, several techniques have beencreated and researched in the field of exact identification. To create a sophisticated system of picture classification, the identification model trained images using class labels. For the purpose of identifying plant disease pictures in the context of leaf photos, Zhang et al. suggested a hybrid clustering-based technique. Using a support vector machine (SVM) classifier to extract texture data and a mean value to determine colour attributes, The aforementioned study set out to create classification algorithms and perform image analysis to extract and recognise characteristics[22]. A more precise and unambiguous diagnosis is now possible because to recent improvements in the science and technology of illness detection. The selection of images and succinct textual descriptions has made it easier for non-experts to identify plant diseases. This innovative technique may be remotely applied from a computer created a method for real-time, on-location photography of afflicted plants utilising mobile devices[24]. They also created a strategy for segmenting leaf pictures and locating disease patches using an enhanced form of k-means clustering. Using a decision tree-confusion matrix and a synergistic evaluation of texture and shape features, Yang et al. provided a technique for detecting microscope pictures in their research.

Based on images of sick maize plants collected in the field, Chad et al. devised a system for automatically detecting such plants. To create an autonomous corn detector, Ni et al. used 1632 pictures of corn kernels to build a deep convolution neural network. Recently, a method for identifying illnesses in rice was described by Lu et al. using deep convolutional neural network (CNN) technology. Zhang et al. used deep learning to train a network to recognise agricultural equipment in images. To increase the accuracy of identifying illnesses of maize leaves, Zhang et al.[23] enhanced deep convolution neural networks.

In Sharma, (2020)[20], machine learning methods and plant applications are reviewed. CNN and KNN were not compared for tomato leaf detection in earlier studies. There are few articles comparing CNN and KNN, and none comparetomatoleaf. The authors of Hatuwal, Shakya, and Joshi (2020) stress the need to use their methods on more plant speciesin future research.'k' in KNN impacts algorithm predictions. If k is too low, noise is likely and the model may be overfit. Alternatively, if k is too large, the model labels [19].

Most machine learning challenges employ black box models. Black-box models only consider input and output. Uncontrollable black box systems. As machine learning advances, explaining model behavior becomes harder (Loyola- González, 2019). XAI makes machine learning discoveries transparent and explainable. Since users can now explain the AI's choice, they may better understand their model and improve its performance or eliminate bias. Machine learning users trust XAI. 2020 [21]

Another model is used to identify plant diseases with 96.5 percent accuracy by Kiranne (2020) [8]. The study classifies plant diseases using AlexNet. AlexNet uses eight layers of learnable characteristics to classify images. The study dataset comprises 54,323 pictures of plant diseases in 38 categories from plant village. Hatuwal, Shakya, Ribon, and Joshi (2020) identified plant illnesses using SVM, KNN, RFC, and CNN.

Table 1: Studies comparing the effectiveness of various CNN approaches used to the diagnosis of plant leaf diseases

Reference	Task	Dataset	Method	Accuracy	Pros and Cons
Zhang et al.(2020) [29]	Leaf diseases on rice and maize	500 rice, 466 maize photos	VGGNet, Inception	92%	In the future, we will concentrate our efforts on deploying the module on mobile devices and applying it to a wider range of situations. Practical implications and uses in the actual world
Mithra et al. (2020) [26]	Identify two leaf diseases	Real-time Photos and Plant Villagedata	DCNN (Deep Convolutional Neural Network)	88.46%	There have only been two corn illnesses detected and categorized thus far, and the datasetis insufficient
Ezzat et al.(2019) [27]	Diagnose 3 maize illnesses	15,408 images from Kaggle	VGG16&19	98.2%	The variety of datasets is not sufficient on itsown.

uryawati,et al. (2019) [28]	Train deeper convolutional neural networks	5632 images of tea	Multicondition training (MCT), AlexNet, GoogLeNet	None-	Only two techniques for segmentation—a blurwith a kernel size of fiveand a rotation of forty—were used.

3. Materials and Methods: Training a neural network

It was suggested that one may train a neural network after first preparing a deep convolutional neural network and extracting an image clustering model from a dataset. The open-source programme Tensor Flow employs data-flow graphs to carry out mathematical operations. The edges in the picture in this case indicate the multidimensional information exhibits (tensors) sent between the hubs, which are the numerical jobs. The adaptable technology lets you to distribute compute across several CPUs or GPUs on a desktop, workstation, or mobile phone using a single API. To improve the company's efforts in artificial intelligence (AI) and deep neural network (DNN) studies, Google scientists and designers working in the Google Brain Group created Tensor Flow as part of the company's Machine Intelligence research organization.

However, the framework is sufficiently all-encompassing to be used in a wide variety of other settings.Artificial intelligence uses convolutional neural networks, a kind of feed-forward imitation neural network. The connections in the creature's visual brain served as inspiration for its network design. The "responsive field" that each cerebral cortex cell possesses allows it to track changes in its immediate surroundings. The response fields of individual neurons only partly overlap in order to tile the visual field. An accurate numerical representation of a neuron's response to stimulation within its open field is represented by convolutional activity.

Convolutional networks are variations on multilayer perceptrons that are driven by natural cycles and have as their primary objective the use of very little pre-handling. They are often used in a variety of industries, including as picture recognition, search engines, and language instruction. Active field-layered convolutional neural networks are made from by (CNNs). These are groups of neurons that repeatedly display certain informational fragments. The outcomes of these collections are tiled such that their information districts overlap in order to provide a more accurate representation of the original picture; this procedure is repeated for each such layer. CNNs may provide enduring interpretation of the information picture because of tiling. Pooling layers may also be used in convolutional architectures to integrate the local or global outputs of groups of neurons. Point aware nonlinearity is introduced at the end of or in between each of these layers, which are also constructed using a mix of convolutional and fully connected layers.

A convolution approach is used in limited information regions to improve speculation and reduce open borders. The usageof shared weight in convolutional layers, which indicates that a similar channel (load bank) is utilized for each pixel in thelayer, is a critical component of flexibility in convolutional networks. Performance is increased while the memory footprint is reduced using this method. There is a set of teachable chunks that span the whole depth of the data volume and are limited by the limits of the layer. Amended Linear Units (Re LU) are employed instead of water logging.
This first method employs adaptive learning of the rectifier constraints to boost accuracy with a little increase in computing cost. In the specific situation of artificial neural networks defined by f (x) = max, the rectifier is an enactment job (0, x). where the contribution to a neuron is represented by x. Similar to halfwave rectification in electrical design, thisis also known as an inclination. Hahn et al. created a dynamical framework for this enactment

difficulty in a 2000 study that published in Nature, drawing on both strong biological inspirations and numerical avocations.

Convolutional networks have effectively employed it in lieu of the more common strategic sigmoid (which is triggered bythe probability hypothesis; for more detail, see strategic relapse) or its more grounded cousin, the exaggerated digression. The rectifier has been the most well known actuation mechanism for very complicated neural networks as of the year2015.Impressive results are produced by combining a deep CNN with ReLU's quicker training rates. The output of eachconvolutional and completely related layer is altered in this way. Despite the yield, standardising the input is not essential since employing ReLU nonlinearity after the first and second convolutional layers lowers the top-1 and top-5 error rates. CNN divides neurons into smaller units known as "include maps" that are located in a hidden layer.An element mapcontains neurons that are all around the same size and leanness. The neurons in the component map are constantlysearching for a matching component.

Convolutional neural network (CNN)

The convolution process is the foundation of the Deep Learning method known as Convolutional Neural Networks, as opposed to the more typical matrix multiplication. For processing visual information, CNN is the most effective deep learning method.It derives a range of spatial and temporal features from input pictures in computer vision applications like image categorization.Input, output, a number of convolutional layers (each with an activation function), a pooling layer, and a fully connected layer make up a standard CNN, as shown in Fig. 1. (each having an activation feature). Convolution is handled by the convolutional layer of CNN. In contrast to the ultimate convolutional levels, which extract more intricate image elements, the earlier convolutional layers of a CNN retrieve basic visual information. We maintain multidimensional arrays of the input, kernel/filter, and feature map. The output feature map has a dimension of mk+1mk+1 if the input matrix is mm in size and the filter is kk in size (where k m). We may thus assume that every convolution process produces a smaller output feature map. With additional convolutions applied, the input picture shrinks in size until it eventually hits zero, to put it another way.

As a consequence, it limits CNNs by limiting the amount of Convolutional layers that may be used. Additionally, the centre components of the input matrix are used more often than its edge and corner components. The CNN's Convolutional layers use padding to overcome these two issues. The padding is done by padding the edges of the input matrix with increasing layers of zeros. By enlarging the area of the input matrix on which the convolution operation must be carried out, this ensures that the input matrix's size does not decrease after the convolution operation is carried out.

To ensure that the items in the margins and corners are used, it also employs a variety of padding layers. Both proper padding and identical padding are possibilities. A pooling strategy is used to minimize the dimensionality of the feature map it receives from the layer above it and lower the number of training parameters. It gathers information about its immediate environment and applies it to produce a single output value. These statistics include maximum pooling, average pooling, and other comparable measurements. While Average Pooling calculates the mean, Max Pooling chooses the biggest value within its region.

It's not always clear how seizures and other disease-related issues will manifest in plant life. The most common causes are bacteria, viruses, and fungi, and their effects on plants may help differentiate between them. Environmental conditions such as temperature and humidity, as well as insufficient or excess food and light, can also play a role. The proposed system employs the Convolutional Neural Network (CNN) method for disease detection in plant leaves because, with CNN, the best accuracy may be reached if the data is good.

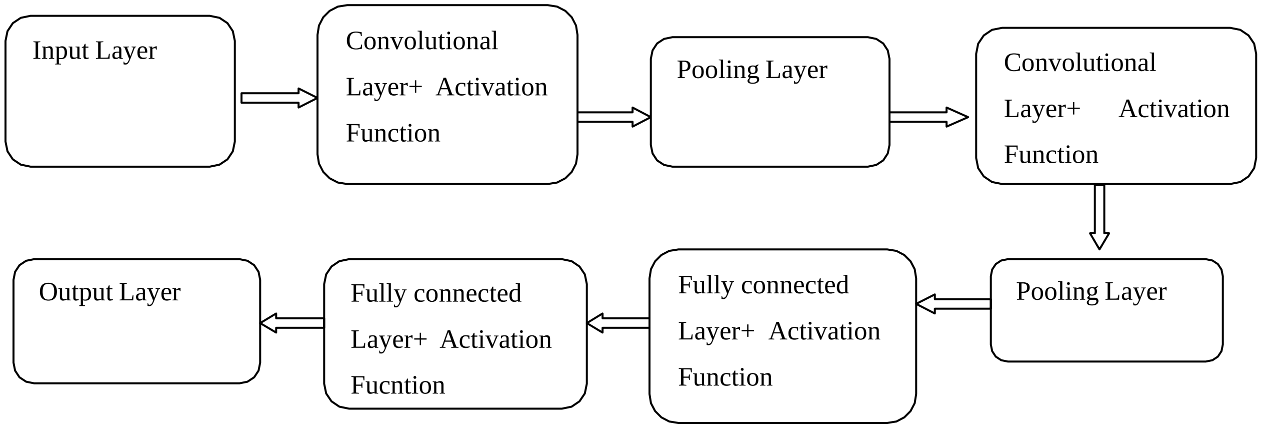
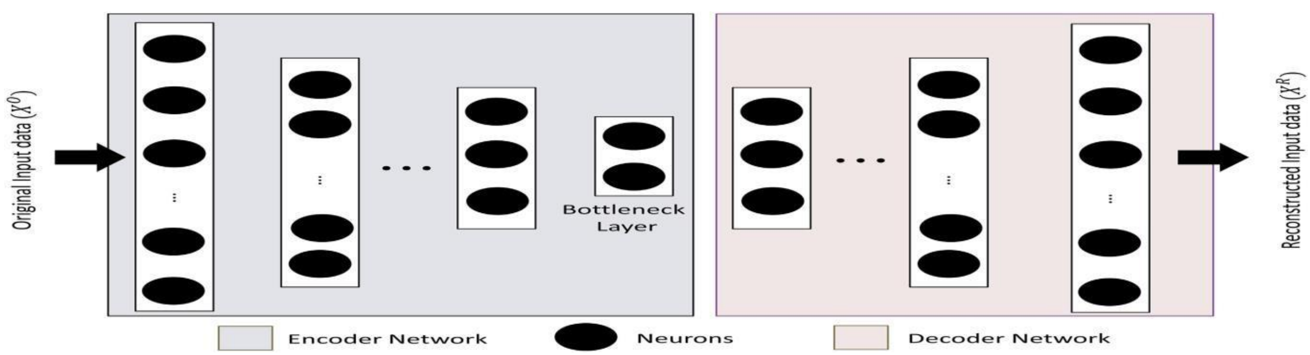


Fig. 1. The architecture of a typical CNN.

3.3. Convolutional autoencoder (CAE):

Auto encoders are neural network-based self-supervised learning algorithms. This system provides detailed about encoding and decoder process. In an N-layer encoder network, the Nth layer is the bottleneck. Fig. 2 demonstrates N-layer

auto encoder



design.

Fig. 2. The architecture of a typical CAE.

4. Proposed Methodology:

4.1 Image Acquisition:

Intelligent visualization and classification systems need a large amount of data for training. In general, the results of machine learning and deep learning systems benefit from training on massive datasets. It is recommended to split the data volume into research sets and evaluation sets in order to train and evaluate deep learning systems. There has never been anything quite like the data collected for this investigation. It offers a wider variety of picture sizes and more power for sensors. For the purposes of preprocessing, feature extraction, and feature selection, we collected plant leaf photos from a variety of species Choice and categorization. Presented in Tables 2–6, the following are some Apple, Strawberry, potato, Pepper, and Corn all suffer from a wide range of leaf diseases plants, in each case.

Table 2: Apple plant leaf with disease

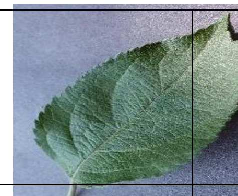
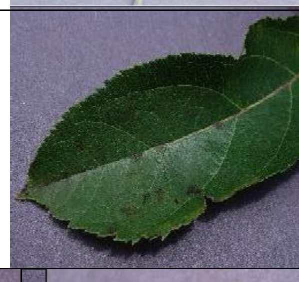
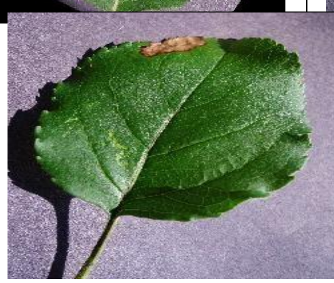
Healthy apple leaf				
Scab-infected apple leaf				
Black rot disease on anapple leaf				
Diseased apple leaf, often known as cedar apple				

Table 3: Corn plant leaf with disease

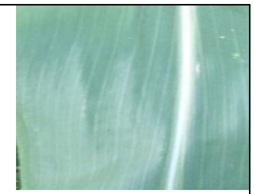
Healthy corn leaf				
Common rust on a corn leaf.				
Northern leaf blight on a corn leaf.				

Table 4: Potato plant leaf with disease

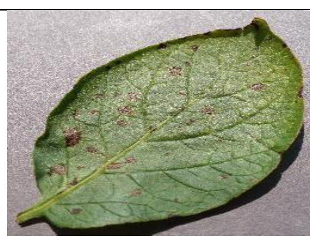
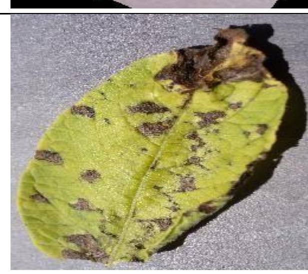
Healthy Potato leaf				
Early-Blight on a Potato Leaf				
Late-Blight on a Potato Leaf				

Table 5: Strawberry plant leaf with disease

Healthy Strawberry leaf				
Strawberry Leaf Scorch				

Table 6: Pepper plant leaf with disease

Healthy Pepper leaf				
Pepper BacterialSpot				

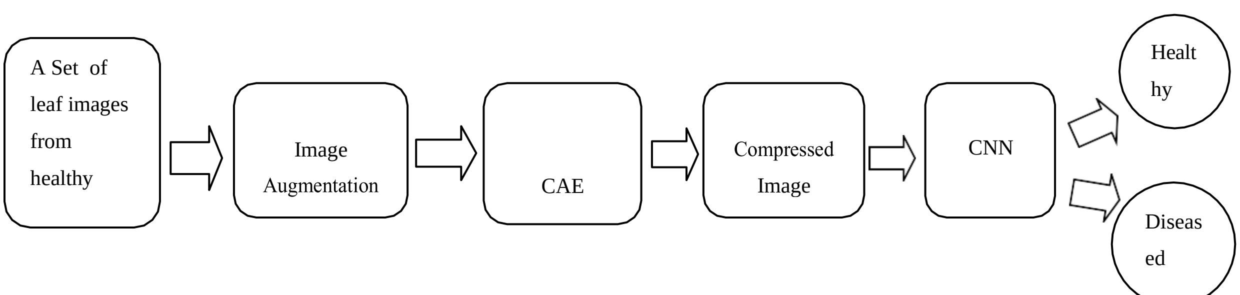


Fig 3 Schematic diagram on Proposed Model

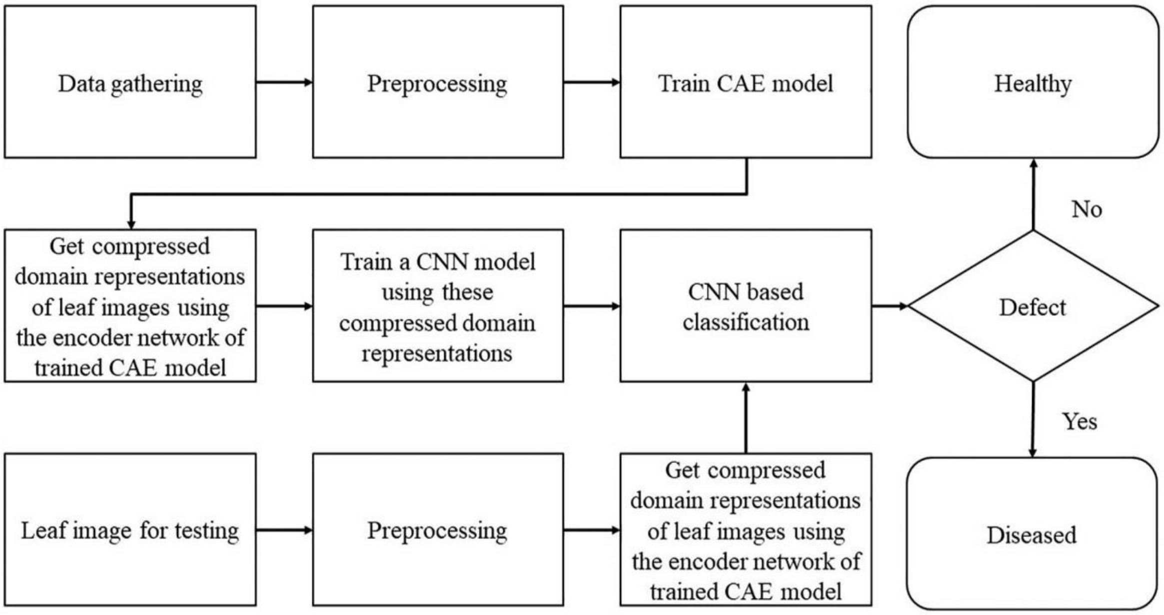


Fig 4. Flowchart for Proposed Model

4.2 Image augmentation:

Image augmentation is crucial for any serious image classifier. Datasets may include hundreds or even thousands of training examples, yet this may not be enough to build a trustworthy model. You may resize the image, rotate it, and flip itvertically or horizontally, just to name a few of the many editing options for your photos. Including these augmentations increases a dataset's useful information content. The pixel dimensions of each individual Plant image are 256 ×256. Data processing and image augmentation are carried out with the aid of the Keras deep-learning framework.

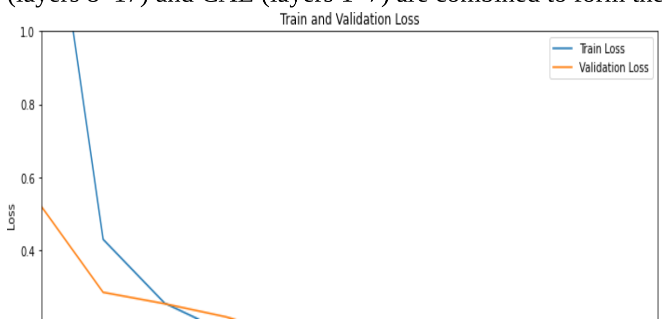
The following options for augmenting training are available:

- The feature "Rotation" allows you to randomly change the orientation of a sample image for training purposes.
- In order to better adapt to different lighting conditions, the model is trained using images of varying brightness.
- In order to shear, the angle of shearing must be altered.

It is the goal of the CAE training stage to reduce the dimensionality of the supplied leaf diseases. The original photographs of leaves have been diminished in size, but their basic characteristics have been maintained. To be sure of this, we employed the strictest Restoration Degradation of CAE constraint attainable. The CAE encoder network delivers compressed domain representations of the leaf images, which are subsequently fed into the CNN for fine-tuning. The input image of a leaf was used to determine whether the leaf was ill or healthy using a convolutional neural network (CNN). As may be seen in Figure 3, a block schematic of the suggested hybrid model has been provided. Figure 4 also includes a flowchart used to illustrate the proposed method. The development of the suggested hybrid model consists of two phases. In the initial stage, while a CAE network is being constructed, the input leaf images' dimension is decreased from 256 ×256 to 32× 32. In this study, we use the Normalized Root Mean Squared Error (NRMSE) loss function to determine the amount by which the reconstructed leaf pictures deviate from the original. NRMSE may be calculated using

Eq (1). $$NRMSE = \frac{RMSE}{\bar{O}}$$ ---------(1)

The input pictures of leaves were downsampled and then classified by the CNN as either unhealthy or healthy. The CNNreceives output from the CAE's Bottleneck Layer. The suggested hybrid model is the result of merging the encoder networks of a conventional neural network (CNN) with a convolutional neural network (CAE). CNN (layers 8–17) and CAE (layers 1–7) are combined to form the optimal hybrid model (layers 1 to layer 8).



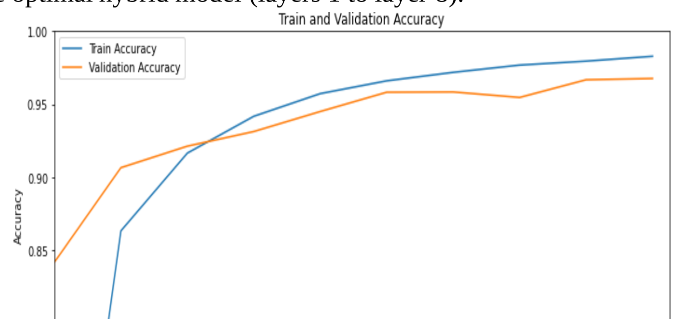


Fig 5. Train and Validation Loss **Fig 6 Train and Validation Accuracy**

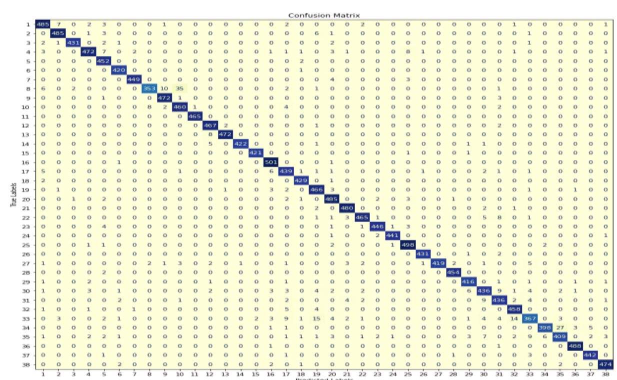


Fig 7.Confusion Matrix

```
Epoch 5/10
2197/2197 [==============================] - 193s 88ms/step - loss: 0.1299 - accuracy: 0.9573 -
val_loss: 0.1694 - val_accuracy: 0.9451
Epoch 6/10
2197/2197 [==============================] - 194s 88ms/step - loss: 0.1014 - accuracy: 0.9661 -
val_loss: 0.1315 - val_accuracy: 0.9583
Epoch 7/10
2197/2197 [==============================] - 194s 88ms/step - loss: 0.0823 - accuracy: 0.9719 -
val_loss: 0.1392 - val_accuracy: 0.9585
Epoch 8/10
2197/2197 [==============================] - 204s 93ms/step - loss: 0.0691 - accuracy: 0.9769 -
val_loss: 0.1477 - val_accuracy: 0.9547
Epoch 9/10
2197/2197 [==============================] - 194s 88ms/step - loss: 0.0623 - accuracy: 0.9796 -
val_loss: 0.1131 - val_accuracy: 0.9668
Epoch 10/10
2197/2197 [==============================] - 194s 88ms/step - loss: 0.0532 - accuracy: 0.9959 -
val_loss: 0.1166 - val_accuracy: 0.9959
```

Fig 8 Sample Execution Screenshots

5. Results and discussion

5.1 Dataset:

In all, there are 61,486 photos of leaves across 39 categories in the dataset used for this work. The number of pictures in each group is equal. The 256x256 pixel photos of plant diseases are complete and accurate in the collection. The dataset, known as "plant village dataset," is available to the general public via a publicly accessible website. The experiment results are described below. First, CAE's results are given. The hybrid model's results are next. NRMSE loss measures CAE network efficiency. Calculations employ original and rebuilt leaf pictures. Train and test NRMSE losses are 0.0597 and 0.0607 for the original and reconstructed leaf. The recommended hybrid model achieves 99.35% accuracy in training time, whereas Figure 5 shows testing. Determining the model's F1-measure, accuracy, and recall (Mohameth,2020). The hybrid model's accuracy is 97.83%. 96.65% recall and F1 brilliance. The suggested hybrid model had a testing accuracyof 98.57 percent, which was greater than the research by Khaparia et al. (2020) [17]. The study's accuracy is higher than the proposed model's.

Studies demonstrate 99.2% (Sanga et al., 2020 [21]), 95.48% (Ngugi, L. C., et al., 2020[15]), and 91.83% accuracy (Chen et al., 2020[5]).. So, the suggested hybrid model has low training and prediction times. The recommended paradigm is shown by two important applications. Low-power devices may be used for automated plant disease detection, and trainingand prediction take less time. The model may be utilized on smart phones, too. Instead of transmitting photos of plant leaves to a cloud or server, farmers may utilize Deep Learning on a mobile app to obtain sensitive data with minimum latency.

Table 7: Performance Metrics of proposed Model

Accuracy	Precision	Recall	F1-Score
Training Accuracy:99.59(%) Testing Accuracy: 98.77(%)	98.76(%)	98.77(%)	98.12(%)

Table 8: Comparison of state-of-the-art testing accuracies and training parameters with the suggested work

Author(s) name and year	Proposed approach	Testing accuracy	Number of training parameters (approximately)
Khaparia et al. (2020)	Convolutional EncoderNetwork	88.68%	8,716
Sang et al. (2020)	ResNet-152	91.45%	7,781
Twari et al. (2020)	VGG-19 + SVM	97.56%	6.789
Mhameth et al. (2020)	ResNet-50 + SVM	98.11%	7,897
Chohn et al. (2020)	VGG-19	96.77%	8,956
Proposed approach	CAE + CNN	98.77%	9,714

Following are the proposed model's performance metrics such as accuracy, precision, recall, and F1-score. Before understanding the evaluation metrics offered thus far, you must understand True positives (TP), True negatives (TN), False positives (FP), and False negatives (FN).

5.1.2 Accuracy:

This metric is generated by dividing the total number of samples by the total number of correct predictions, both positive and negative. Table 7 shows the accuracy of the present model; the proposed model scored 99.59 percent. Accuracy seems to be a useful measure for comparing the two models in this study, given the dataset picked contains labels with thesame number of samples as the dataset itself. The accuracy paradox makes it hard to show that accuracy may be desirable in imbalanced datasets.

5.1.3 Precision:

This statistic is calculated by dividing positive predictions by accurate forecasts (Truepositives) (Truepositives and false positives). This statistic shows how accurate a classifier is by answering the question: how many "early blight" samples did the classifier really predict? Table 7 shows the model's 98.76% accuracy. This shows that, compared to the current model, the proposed model would provide more relevant findings.

5.1.4 Recall:

This statistic shows how sensitive the model is; how many of the "Early blight" samples in the dataset were also predicted to have "Early blight"? Table 7 shows that the model has 98.77% recall. Compared to the current model, the suggested model may generate most of the significant plant village dataset results.

5.1.5 F1-score:

Since the suggested model has superior accuracy and recall than KNN, comparing F1-Scores is unnecessary. The F1- Score is a harmonic mean of accuracy and recall that helps decide if a model with high precision and poor recall is better than one with low precision and high recall. Since CNN has superior accuracy and recall than KNN, compare Table 7 shows that the recommended model beats the present classifier models in terms of F1-score accuracy, with 98.12%.

6. Conclusions and Future scope:

A plant disease's early stages are notoriously difficult to identify. Many scientists have been hard at work on selfsufficient plant disease diagnosis systems, using a broad variety of machine learning and deep learning technologies. Mostof these methods, nevertheless, either misclassify objects or need a large number of parameters for training. Using the (CAE) and the CNN, this research developed a unique hybrid model for autonomously diagnosing plant illnesses (CNN). The proposed hybrid method employs the CAE encoder network to generate compressed domain representations of leaf pictures, which are then classified by the CNN. When compared to current state-of-the-art systems, the quantity of features and, by extension, training parameters may be drastically reduced by using CAE to reduce dimensionality. The Bacterial Spot disease in peach plants was used as a case study for the model. The model achieves 99.59 percent accuracy during training and 98.77 percent accuracy during testing using just 9,714 training parameters. Because the suggested hybrid model uses fewer training parameters, both the time needed to train the model for automated plant disease detectionand the time needed to diagnose the illness in plants using the trained model are drastically reduced. The plant village dataset served as the basis for the one that was utilized in this investigation. In the future, using the new dataset, which contains a large number of images of plant diseases captured in real surroundings, new algorithms will be developed to improve plant disease diagnosis in real-world photos. These new algorithms will produce results quickly, making them suitable for use in applications that require real-time processing.

References:

1. [1]S. Momen, S. Irfanul Alam, T.R. Shahidi, and K. Ahmed (2019). Machine learning for rice leaf disease diagnosis. Global Summit on Green Technologies for the Fourth Industrial Revolution (STI). Dhaka, Bangladesh: IEEE, pages 1-5.
2. T. Ashraf & Y. N. Khan 2020. Using AI to categorise the weediness of a rice field. 175, 105590 (Comput. Electron. Agric.). https://doi.org/10.1016/j. \scompag.2020.105590.
3. 2019 E. Bisong. Autoencoders. A Step-by-Step Tutorial for Novices Using Google Cloud to Construct Machine Learning and Deep Learning Models. Pp. 475–482 in Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4470-8 37.
4. 2020: Chen, Junde and Jinxiu Chen, D. Zhang, Y. Sun, and Y.A. Nanehkaran. Isolating plant diseases from photographs using deep transfer learning.173, 105393, Comp. Electron. Agric. https://doi.org/10.1016/j.compag.2020.105393.
5. Chohan, M., A. Khan, S. Katper, and M. Mahar. Using deep learning, we can identify plant diseases. 9(1), 909–914. International Journal of Emerging Technologies and Engineering. https://doi.org/10.35940/ijrte. \sA2139.059120.
6. ToLeD: Tomato Leaf Disease Detection Using Convolution Neural Network" Agarwal, M. et al., Procedia Computer Science, 167, 293–301, 2020. doi:10.1016/j.procs.2020.03.225
7. R. Ault, "Optimization Study of an Image Classification Deep NeuralNetwork," Final Report, 2020, p.
8. Dieber,J and Kirrane, S.(2020) ‘Why model why? Assessing the strengths andlimitations of LIME.’, CoRR.
9. Garbin, C., Zhu, X. and Marques, O. (2020) ‘Dropout vs. batch normalization: an empirical study of their impact to deep learning’, Multimedia Toolsand Applications, 79(19), pp. 12777–12815. doi: 10.1007/s11042-019-08453-9.
10. J, A. P. and Gopal, G. (2019) ‘Data for: Identification of Plant Leaf Diseases Using a 9-layer Deep Convolutional Neural Network’, 1. doi: 10.17632/tywbtsjrjv.1
11. Khodabandehloo, E., Riboni, D. and Alimohammadi, A. (2021) ‘HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline’, Future Generation Computer Systems, 116, pp. 168–189. doi: 10.1016/j.future.2020.10.030.
12. Kinkead, M. et al. (2021) ‘Towards Explainable CNNs for Android Malware Detection’, Procedia Computer Science, 184, pp. 959–965. doi: 10.1016/j.procs.2021.03.118.
13. Linardatos, P., Papastefanopoulos, V. and Kotsiantis, S. (2021) ‘Explainable AI: A Review of Machine Learning Interpretability Methods’, Entropy.doi: 10.3390/e23010018.
14. David P. Hughes & Mary Salathe 2020. An image database on plant health that is freely available to researchers in order to facilitate the creation ofportable disease diagnostics. Preprints on arXiv are available at arXiv (1-13).
15. Ngugi, L. C., Abelwahab, M. and Abo-Zahhad, M. (2021) ‘Recent advances in image processing techniques for automated leaf pest and disease recognition – A review’, Information Processing in Agriculture, 8(1), pp. 27–51. doi: 10.1016/j.inpa.2020.04.004.
16. Andres, F., C. Anutariya, and W. Jearanaiwongkul (2020). The development of a plant disease identification system using ontology. A Record of Presentations from the 10th Annual International Conference on Advances in Information Technology (IAIT 2018). New York, NY: ACM Press, pages 1-8.
17. Karlekar Khaparia, Ashraf A., & Seal, A. (2020). https://doi.org/10.1145/3291280.329178. Diseases of the soybean leaf are catalogued in SoyNet. Computing and Electronics in Agriculture 172 (2010):105342. https://doi.org/10.1016/j.compag.2020.105342.
18. Authors: Khamparia, A.; Saini, G.; Gupta, D.; Khanna, A.; Tiwari, S.; de Albuquerque, V.H.C.; 2020.
19. W.-S. Kim, D.-H. Lee, and Y.-J. Kim. 2020. Automatic identification of onion downy mildew symptoms using machine vision. Computers and Electronics in Agriculture, 168, 105099. https://doi. \sorg/10.1016/j.compag.2019.105099.

20. Paryudi, I. (2019) 'What Affects K Value Selection In K-Nearest Neighbor', International Journal of Scientific & Technology Research. Available at: /paper/What-Affects-K-Value-Selection-In-K-NearestParyudi/34c2817df22843d5dc1e4047086617d2d28dd7b6 (Accessed: 6 June 2021). [20]Sharma, V., Verma, A. and Goel, N. (2020) 'Classification Techniques for Plant Disease Detection', International Journal of Recent Technology and Engineering, 8(6), pp. 5423–5430. doi: 10.35940/ijrte.F9902.038620.
21. Tjoa, E. and Guan, C. (2020) 'A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI', IEEE Transactions on Neural Networks and Learning Systems, pp. 1–21. doi: 10.1109/TNNLS.2020.3027314
22. In 2020, Mohameth, F., Bingcai, C., and Sada, K.A. Plant disease detection with deep learning and feature extraction using Plant Village. To cite this article: J. Comp. Comm. https://doi.org/ \s10.4236/jcc.2020.86002 .
23. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad Zhang, M.; 2020. Here we take a look at the recent developments in image processing methods that have enabled fully automated detection of leaf pests and diseases. 1-25 in Inform. Proc. Agric. https://doi.org/10.1016/j.inpa.2020.04.004.
24. ResearchGate. Zhang, H., Zhang, L. and Jiang, Y. (2019) 'Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems', in 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP). 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6. doi: 10.1109/WCSP.2019.8927876.
25. Zafar, M. R. and Khan, N. (2019) DLIME: A Deterministic Local Interpretable Model-Agnostic Explanations Approach for Computer-Aided Diagnosis Systems.
26. Mishra, S.; Sachan, R.; Rajpal, D. Deep Convolutional Neural Network based Detection System for Real-time Corn Plant Disease Recognition. Procedia Comput. Ence 2020, 167, 2003–2010. [CrossRef]
27. Darwish, A.; Ezzat, D.; Hassanien, A.E. An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. Swarm Evol. Comput. 2019, 52, 100616. [CrossRef]
28. Yuwana, R.S.; uryawati, E.; Zilvan, V.; Ramdan, A.; Fauziah, F. Multi-Condition Training on Deep Convolutional Neural Networks for Robust Plant Diseases Detection. In Proceedings of the 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Tangerang, Indonesia, 23–24 October 2019.
29. Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y.A. Using deep transfer learning for image-based plant disease identification. Comput. Electron. Agric. 2020, 173, 105393. [CrossRef]