A Survey on Plant Leaf Disease Detection

Tannu Kumari
Department of Computer Scienceand Engineering
Jain Deemed To be University, Bangalore, India

M.K. Jayanthi Kannan, PhD
Professor, Department of Computer Scienceand Engineering
Jain Deemed To be University, Bangalore, India

Vinutha N., PhD
Department of Computer Scienceand Engineering
Jain Deemed To be University, Bangalore, India

ABSTRACT
Climate change and pests' attacks have a high impact on the healthy harvest in our country. A farmer can ensure the healthy production of his labor if he accurately and timely detects plant diseases. Earlier, disease detection was a manual job and it was impractical also. Now with the advancement of technology various Machine Leaning and Deep Leaning technologies are on a verge of replacing manual labor for leaf disease detection. Plant leaf disease detection using leaf images has now become a hot topic for many researchers. Here we have analyzed the model's efficiency in various research papers which has incorporated Machine learning (ML) or Deep Learning (DL) technique. Some papers used only Image segmentation for leaf disease detection but these models were quite inefficient and inaccurate but managed to differentiate infected parts from healthy leaves. However, Image segmentation in combination with either machine learning or deep learning models improves leaf disease detection. Image segmentation helps in segmenting the infected part of the leaf from large data set and then feeds this result to aMachine Learning or Deep Learning model. The large data set usually has high noise. Hence, prepossessing also plays a vital role in the prediction of disease.

Keywords
Agriculture; Deep Learning; Leaf Disease; Machine Learning, Pre-processing

1. INTRODUCTION
Agriculture is the prime economic source of the Indian economy. There are numerous technologies utilized by farmers to increase the production of food crops. But the production is hampered due to various reasons such as climate change or the plants may be infected with various diseases. Though the frequent occurrence of natural calamities is beyond our control. However, the plants can be protected from pests. Therefore, the identification of the infected part of the plant at the early stage is essential to incorporate the necessary measures to increase crop production. Disease identification has a significant impact on crop production. Researchers and scientists are developing and incorporating various algorithms of deep learning, image segmentation, and machine learning to create a framework that has high efficiency in the identification of multiple crop diseases affecting different types of plants. With the advancement of technologies, the focus has been shifted to the development of devices that can accurately detect plant disease and help farmers in a healthy harvest. Both machine learning and deep learning have a vast difference in their way of learning from their environment. The models utilize machine learning algorithms, dissect data and then learn from this data and make an informed decision on what it has learned. Whereas in a Deep Learning based model, algorithms are structured in different layers to create an "Artificial Neural Network" (ANN). In this model, data is parsed in different layers and the model learns from it and makes intelligent decisions on its own.

Convolutional Neural networks and Generative Adversarial networks are very significant for plant disease detection. Some of the CNN architecture used in leaf disease detection are VGG16, Alex Net, and Dense Net. VGG16 has a good architecture for benchmarking on a particular task. But, VGG16 has two major drawbacks, it performs slowly during training, and network weights assigned to the architecture are quite large. Alex Net allows for multi-GPU training but it is not suitable for color images. And since this model is not very deep enough, hence feature extraction is not very efficient. A Dense Net appeared in the year 2017 to resolve the vanishing gradient problem and it also reinforced feature extraction propagation. The only snag for this model is that for each L layer there are L(L-1)2 connections. This connection has an impact on network computation and parameter efficiency.

2. LITERATURE REVIEW
In most of the papers, both image segmentation and CNN models perform disease detection in plant leaves. Some papers have used only the image processing technique to improve disease detection. However, the accuracy improves with the fusion of image segmentation and CNN models.

M. Bhange et al [1],The authorused only the image segmentation technique for Pomegranate disease detection. The problem chosen in this paper is Bacterial Blight disease on fruit. This disease is one of the most common occurrences in plants. Based onColor, Morphology, and Color Coherence Vector (CCV), the input image is first processed to extract its features, and then it is further processed by clustering using the k-means algorithm. Out of the three features extracted from the plant, morphology gave the best result. Finally, the selected feature is subjected to a Support Vector Machine (SVM) for binary image classification as infected and non-infected. The model is then subjected to captured images as test images. But the model failed to detect the disease. Therefore to solve the problem on test images, the author provided an alternate option of intent search technique. In this technique, the model provides the user with the most similar images concerning the test images from its database and the user selects one among the option, and based on his/her choice further remedial action was proposed by the system. The accuracy obtained by this model is 88%.

Monika Jhuria et al [2], The authorproposed efficient farming and fruit grading by using the image processing technique. Plants used for disease detection are namely Grapes, Apples, and Mangoes. In this paper, the following steps are carried out: The image was first pre-processed and then the features were extracted from the pre-processed images. Finally, the accuracy was evaluated using the extracted features. From the extracted features, morphological features achieved 90% accuracy.

Shiv Ram Dubey et al [3], The authorworked on Apple Fruit disease detection using the Image Local Binary Patterns method. The images of Apple Rot and Apple Blotch diseases were focused on in this paper. For the considered images, the Image segmentation technique was used along with K-means clustering to create clusters. The clustering is followed by the feature extraction: Histogram, Color Coherence Vector, and Local Binary Pattern. Summarization of the above techniques is as follows: Color Coherence Vector (CCV) is a bit complicated method. In CCV, each pixel is categorized as Coherent or Incoherent. A small connected component forms part of an incoherent pixel while a big connected component forms a part of a coherent pixel. Several pixels with a particular brightness or toner value are plotted on a horizontal and vertical axis by the Histogram technique. Then the Digital Editor algorithm is applied to adjust the brightness of every pixel. After applying the Digital Editor algorithm, a texture operator called Local Binary Pattern is utilized for labeling the pixel based on the concept of thresholding the neighborhood of each pixel in the image. As the last step, K-means Clustering is used to partition the observations into K clusters.

***Mohanty et al*. [4]** The authorimplemented the AlexNet CNN model along with Image Segmentation. 'Alex Krizhevsky' along with 'Geoffrey Hinton and 'Ilya Sutskever' designed AlexNet. AlexNet is made up of eight layers the first five layers are convolutional and max-pooling layers and the last three are fully connected layers. The author in this paper used a non-saturating ReLu activation function which proved to have better training performance over tanh and sigmoid. Rectified Linear Activation Function (ReLU) is a piecewise linear function, which outputs the inputif it has a positive value. Otherwise, it will output zero. ReLU prevails over the vanishing gradient problem and allows the model to learn faster and perform better.

***Aravind Krishnaswamy Rangarajan et al* [5],** the Author used both ALexNet and VGG16 CNN architecture. The accuracy of both the models was compared and it was revealed that the accuracy of both models depends on the number of given sample images. When the number of sample images increased, it was found that accuracy also increased. And it was also found that both the models yielded an accuracy of 96.19% for VGG16 and 95.81 for AlexNet when 373 images were used. Simonyan and Zisserman introduced the VGG network. A 3*3 convolutional layer stacked on top of each other with increasing depth is used in this network. Max pooling reduces the volume size. A SoftMax Classifier is followed up with two fully connected layers each with 4096 nodes. The "16" and "19" symbolize several weight layers in the network. Training is slow in VGG Network and also it weighs much in terms of disk space. Vgg16 weighs around 533MB and Vgg19 around 574MB.

***R. Sujatha et al.*[6],**performed the comparison of deep learning and machine learning technique for the detection of Grape plant leaf disease. The disease dataset comprises Black Spot, Canker, Greening, Melanose, and Healthy plant leaves. This data set is then subjected to both ML and DL models. The performance of ML models like SVM, RF, and SGD are compared with DL models like VGG#16, Inception-V3, and VGG#19. The comparison result shows, that the DL models outperform the ML models. The accuracy achieved for DL Models is: - VGG-16 gives an accuracy of 89.5%, VGG-19 gives an accuracy of 87.4%, and Inception –V3 gives an accuracy of 89%. The accuracy achieved for Machine Learning Models is:- Stochastic Gradient Descent gives an accuracy of 86.5%, Rainforest gives an accuracy of 76.8%, and Support Vector Machine gives an accuracy of 87%. As a Future Enhancement, the author aimed to improve computational accuracy even for the small-sized dataset by using Fuzzy Logic and Bio-Inspired methods.

***Miaomiao Ji, Lei Zhang, Qiufeng Wu* [7],** Theauthors worked on various Grape Leaf diseases. The collected dataset comprises various images of leaf diseases like black rot, ESCA, leaf spot, and Isariopsis. A combined model having both InceptionV3 and Resnet50 was proposed as the CNN model by Author. The proposed model distinguishes a healthy leaf from a diseased leaf by achieving a training accuracy of 99.17% and a test accuracy of 98.57%.

***Ashraf Darwish, Dalia Ezzat, Aboul Ella Hassanien* [8],** presented a model for plant disease diagnosis based on Convolutional Neural Network and Orthogonal Learning Particle Swarm Optimization. In the proposed model, the authors used two CNNs, namely VGG16 and VGG 19. These models were pre-trained for plant leaf disease diagnosis by classifying the images of leaves as healthy and unhealthy. CNN in combination with Orthogonal Learning Particle Swarm Optimization is used due to its ability tooptimize various hyper parameters and these parameters are used for the classification of plant disease. CNN has different hyper parameters, thus, it becomes a challenge to identify and manually optimize these hyper parameters. The optimization problem is resolved using 'Orthogonal Learning Particle Swarm Optimization, where it searches for the optimal value that has an impact on the classification.

***Uday Pratap Singh et al* [9],** The author designed a model for the classification of Mango leaves with Anthracnose disease using a multilayer CNN. For this author used an Alex Net architecture for mango leaves classification and obtained an accuracy of 97.13%.

***Parul Sharmaa et al* [10],**developed a unique solution for classification, where a Convolutional Neural Network (CNN) model is trained on the segmented images. In this paper, a comparison of two CNN models was performed. The first technique is, F-CNN on the image without segmentation and S-CNN for the segmented image. When the comparison is performed, S-CNN doubles the performance as compared to F-CNN and gives the accuracy of 98.6% even when tested with previously unseen data of 10 disease classes.

***J. Sun et al* [11],** Author worked on the Maize plant with Northern Maize Leaf blight disease. The paper comprises three steps of data processing. The first step includes data processing by improved retinex which handles the problem of poor detection caused by a high-intensity light. An improved Region Proposal Network (RPN) network is used to adjust the anchor box of diseased leaves as the second step and this network reduces the search space of the classifier by deleting negative anchors and thereby creating a repository of the better initial information.

***Pranjali B. Padol; Anjali A. Yadav* [12],** In this paper, the authors applied K Means clustering to locate the infected region of plant leaves, and then it is further subjected to Support Vector Machine (SVM) to detect disease type. Using this framework, the authors obtained an accuracy of 88%. Generally, large data sets are a major drawback of the SVM algorithm. A large data set usually comes with more noise, which affects the performance of SVM that is because of the overlapping of target classes. Or in some cases, there may be a mismatch in several features that are in the training and original dataset. There is also some limitation with K-Means clustering, like predicting many clusters are slightly difficult over the considered initial Seed, and the order of the data need to be perfect as it has a high impact on the final result.

***Ümit ATİLA*, [13].** Theyworked on multiple plants and their diseases. The model used for plant leaf disease detection is the Efficient Net Deep learning model and its performance was compared with Deep Learning Models. The authors considered the images of apples, Grapes, Tomato, Cherry, Peach, and Potato. A disease that was targeted is Cedar Apple Rust of Apple Fruit, Powdery Mildew of Cherry, Black Measles of Grape Plant, Late Blight of Potato Plant, Bacterial Spot of Peach Fruit, Late Blight of Tomato. The augmented and original datasets of 55,448 and 61486 respectively are used to train the models. The Efficient Net Architecture was trained by the Transfer Learning approach in which all the layers of the model are trained with some weights. Both original and augmented dataset was used to test the Efficient Net model and Deep Learning model. The Efficient Net model showed the highest accuracy and precision. As a future enhancement, the author's focus was to increase the number of classes and also to expand the current plant leaf dataset. This helps the models to get trained on a huge data set and also to enhance the model performance.

***G. Sambasivam, G.D. Opiyo* [14],** The author worked on leaf disease detection in Cassava Plant. The author proposed the CNN model which comprises three convolutional layers and four fully connected layers. For extracting richer features, the author stacked the convolution and batch normalization layer of two sets before max pooling. Max Pooling layer reduces the spatial dimensions of the input images. The network consists of 4 fully connected layers with 512 neurons in the first layer, 1024 neurons in the second and third layers, and lastly 256 neurons in the fourth layer and a neuron per category in the output layer corresponding to five different classes. After hyperparameter tuning and optimization with grid search. Drop out was used in the fully connected layers which reduces the generalization error and over-fitting problems by encouraging the neural network to learn sparse features out of raw observations that always yields good performance by empowering the model's ability to generalize to new data. In general, the convolutional layers extract key features from the images and the fully connected layers focus uses the extracted features to classify images of cassava leaves into five different categories. Input image attributes take an order of 3 tensors. Example: an image with H rows, w columns, and 3 channels (R, G, B color channels) on the input layer and a neuron for every category in the output layer corresponds to five different classes: Bacteria Blight of Cassava, Green Mite infection, and healthy. The ReLU (Rectifier Linear Unit) activation function was used in the convolutional and hidden layers and SoftMax (for a multi-class case) is used in the output layer. The proposed model achieved an accuracy of 93%.

***A. Khattak et al.* [15],** worked on citrus leaves and fruit diseases like Cankers, black spots, and melanoses. The dataset comprises 2293 images of all the classes. Pixel scaling and data normalization are used as preprocessing techniques. Two layers of convolution are used, where the task of the first layer is to collect images of low-level features, and the task of the second layer is to collect high-level features. After each convolution layer, max-pooling is used to reduce the size of the feature map. The fourth step is flattening where the feature vector of three-dimension is converted to one dimension. The last step is classification in which the output of the flattening layer is used to predict the class with the help of the SoftMax activation function. An accuracy of 95.65% is obtained for the proposed model. In the future, the author planned to add more numbers of plant imagesto other deep learning models: RNN, LSTM, Bi-LSTM, and hybrid models.

***S. C. K. et al* [16]** This paper focuses author's works on the Cardamom (Colletotrichum Blight, Phyllosticta Leaf Spot diseases) and Grapes plant (Black Rot, ESCA, isariopsis Leaf Spot diseases).The disease of cardamom is identified with the help of 1724 real-time images.These images are affected by noise, light illumination, and different angles. 4026 images of grapes leaf are taken as an input. Cardamom images are not directly subjected to the models because they contain background, light illumination, and a lot of noises. These images reduce the accuracy of the model. Hence, for removal of background noise U2–Net is used. U2-Net has three parts, the six-layer encoder comes under the first part which creates the Residual U-Block. The second part contains a five-stage decoder and the third part creates Saliency Probability Maps. Saliency maps find the difference between the background and interested part of the leaf. CNN, Efficient-Net, Efficient-Net V2-S, Efficient-Net V2-M, and Efficient-Net V2-L are the different classification models considered in this work. The Efficient-Net V2-L model gives the highest accuracy of 98.26% for cardamom and Efficient-Net gives the highest accuracy of 97.81 for grapes.

***Z. Zinonos et al* [17]**, In this paper, authors focus on grapes plant leaf diseases like Black rot, Esca, and Leaf blight disease were identified by using LoRa(Low-Power Wide Area Networks) and deep learning technology which uses low-resolution images. Low-Power Wide-Area Networks are used to transmit Grapes leaf images from grape fields. Low-Power Wide-Area Networks convert images into packets. Bandwidth, Spreading Factor, Coding Rate, and Transmission Power are parameters of Low-Power Wide Area Networks. Bandwith is responsible for long-distance transmission. The spreading factor is responsible for deciding the length of packets. A coding rate is a code that corrects the packet errors before transmission. The last parameter is transmission power, which is the power consumed by transmitting a package of data. CNN models MobileNetV2 and ResNet50 are the deep learning models considered by the author. This results in the highest accuracy of 99.77% for the ResNet model with a 100% Packet Reception Ratio. And 99.65% accuracy is obtained for ResNet without augmentation with a 100% Packet Reception Ratio.

***K.K. Chakraborty et al.*[18],** The author worked on the Potato plant by optically recognizing Potato Leaf blight disease. The author collected the Potato leaf dataset from Plant Village Dataset. This dataset is augmented using label-preserving transformations, reducing overfitting and artificially increasing the number of data/samples. The author incorporated image pre-processing techniques such as resizing and scaling techniques. Resizing alters the size of the image to 224px height and 224px width which is also the input size for the first layer of the CNN model and rescaling standardizes each pixel value in the range of 0-1 and this helps to prevent the model to slow down. The processed image is then fed as an input to the CNN model. In this, the output of each layer acts as input forevery corresponding layer. Then the author tested the dataset on different models such as VGG16, VGG19, MobileNet, and ResNet50. However, VGG16 turned out to be the model with the highest efficiency 97.89%. VGG19 achieved 80.39%. Mobile net achieved 78.84%. ResNet 50 achieved 73.75% accuracy. The final output layer of VGG16 is fine-tuned with a sparse Adam optimizer and Binary Cross-Entropy loss function.

***M. Ahmad et al* [19],**The author Pepper a dataset from two different sources is considered and compared the efficiency of various models on the considered dataset. The author has taken the first dataset from the National Institute of Horticultural and Herbal Science, Republic of Korea, and another one from the Plant Village dataset. The dataset comprises 99507 images belonging to 24 different classes of diseases. Among these 24 classes, 6 are related to pulp and stem respectively, 9 are related to leaf, 2 are related to larva, and 1 class of healthy images for pulp, leaf, and larva. The imbalance dataset is handled by employing it to the different approaches such as the Adaptive Synthetic Sampling Approach and Synthetic Minority Oversampling Technique. These two techniques decrease the count of overrepresented classes and augment the class of underrepresented classes. After data processing, it is fed to different CNN models. Random Initialization, Fine Tuning, Feature Extraction using Random Initialization, and Feature Extraction using Fine-tuning are the different parameters considered for improving the performance of the model. InceptionV3 with Random Initiation and Fine-tuning achieved the highest efficiency 99.57% and 99.56% respectively.

***S. M. Hassan, A. K. Maji* [20],**The author used three different datasets of different plants. The author has used the Rice Plant dataset, Cassava Plant dataset, and Plant Village dataset. From the Plant Village dataset, the author has used a Corn, Potato, and Tomato plant disease image set. The rice plant dataset consists of 5932 images and the Cassava plant dataset consists of 5656 images. These images are resized to 256 x 256 pixels and rescaled before subjecting them to the CNN model. The Plant Village dataset comprises images that are uniform with most of the pictures taken from a uniform background. Class balancing is performed on both datasets. Class balancing is done on both the dataset was resized, and rescaled before being fed to the CNN model. To check the stability of the proposed CNN model, k-fold cross-validation was used. After running the model for 50 epochs, the author obtained an accuracy of 99.81% for the Plant Village dataset. An accuracy of 99.94% and 98.17% is obtained for Rice and Cassava plant images. As a future enhancement, the author planned to use the model for weed detection and pest identification.

***Z. Zinonos et al* [21],** The collection of leaf datasets for early-stage disease and rare diseases is a complicated task. A very smaller number of images are available for these types of plant diseases. This smaller number of images affects the efficiency of the DL model because of overfitting. Therefore the author proposed grained-GAN for data pre-processing. The grained-GAN works in two stages: first is to identifythe leaf spot location and after that segmentation is applied on the spot part. The second is the data augmentation stage. Four detection boxes of different sizes (32∗32, 64∗64, 96∗96, and 128∗128) were selected to find the ratio of area intersected to the area of candidate bounding boxes. The 64*64 have the largest ratio sample, so it is selected for segmentation. Generative Adversarial Networks (GAN) help in the synthesis of images that have the same feature as original images. Thus, Data Augmentation isperformed by Generative models (DCGAN, Info-GAN, WANG, LRGAN, original, Leaf GAN, WGAN-GP, CAVE-GAN, fine Gradient GAN). Then the original and generated dataset is subjected to the five CNN (AlexNet, ResNet-50, DenseNet-121, X-ception, VGG-16) models for the classification of images is used. The best accuracy of 96.7%is obtained by using fine gradient GAN with ResNet-50.

TABLE 1. Comparison Between Different Classifier

Article	Crop	Disease	Pre-processing technique	Classifier	Performance	Year
[23] J. Chen, et al	‘Rice plant, maize’	‘Rice disease like Stackburn, Scald, Smut, White Tip, Streak, Phaeosphaeria Spot, Maize Eyespot, Goss's Bacterial wilt’Gray, Leaf Spot	‘grey transformation, image filtering, image sharpening and resizing,’	INC-VGGN	92%	2020
[24] S. Hernández and J.L. López	‘Apple,grape,tomato, cherry,Peach,strawberry, blueberry, black paper,Raspberry Soybean’	‘26 diseases, 38 classes’	Stochastic Gradient Descent	‘Bayesian fine-tuning VGG16’	96	2020

[25] A. Waheed, et al	'Maize'	'Common rust, Healthy crop, Cercospora, spot Gray, spot, Northern blight '	data augmentation	'Optimized DenseNet, VGG19, NasNet, XceptionNet , EfficientNet '	'EfficientNet-99.84 XceptionNet-93.52 NasNet- 91.9 Optimized DenseNet-98.06 VGG19Net 96.36'	2020
[11] J. Sun et al	'maize'	'Northern leaf blight'	'high-pass filter reduces the reflection of the image'	'generalized intersection over union (GIoU), , RetinaNet 500, DSSD, RelationNet 600, SNIP'	GloU-91.83 RelationNet 600 – 86.5, SNIP – 89.79, RelationNet 500 – 77.43, DSSD 531 – 78.15	2020
[7] M. Ji, L. Zhang, Q. Wu	'grape'	black rot, esca, isariopsis, spot'	'Data Augmentation'	CNN architecture on InceptionV3 ResNet50;	validation accuracy is 99.17% and test accuracy is 98.57%'	2020
Article	**Crop**	**Disease**	**Pre-processing technique**	**Classifier**	**Performance**	**Year**
[29] O.E. Apolo-Apolo, et al.	'Citrus leaf dataset;	'Blackspot, Melanose, Canker, greening'	Squeeze net	ML SVM, Random Forest, Stochastic Gradient Descent, & DL inception-v3, VGG-16, VGG-19	'RF gives 76.8% SGD gives 86.5% SVM gives 87% VGG-19 gives 87.4% Inception-v3 gives 89% VGG-16 gives 89.5%'	2021
[26] Mohit Agarwal et al	Tomato	12 tomato disease	Data augmentation technique	'VGG16, InceptionV3, MobileNet, and A proposed model of CNN-based architecture with three convolution max-pooling layers and varying numbers of filters in each layer is applied'	'Mobilenet 63.75%, VGG 16 77.2%, InceptionV3 63.4%, Proposed Model 91.2%'	2020
[27] U. Atila et al	14 types of plant	26 disease	six types of augmentation method	'AlexNet, ResNet50, VGG16, Inception V3'	'AlexNet 99.67, ResNet50 99.88, VGG16 99.94, Inception V3 99.93'	2020
[30] G. G. and A.P. J.	Leaves of 13 different plant	29 type disease	image flipping, gamma correction, noise injection, principal	'support vector machine, K-NN, decision tree	'SVM – 50.6 decision tree– 72.2 logistic	2019

			component analysis, color, augmentation, rotation, and scaling	logistic regression, and deep convolutional neural network’	regression–80.9 K-NN-87.86 and deep CNN-96.46’	
[14] G. Sambasivam, et al	Cassava	Cassava green mite, CBSD, CMD, Cassava bacterial blight	‘class weight, focal loss, SMOTE, and different image dimensions’	three convolutional layers and ahead of four fully connected layers.’	93%	2020
[28] Y. Zhao, Z. Chen, X. Gao, W. Song, et al	Tomato(disease) Apple Corn (healthy) Grape (healthy) Potato (healthy)	‘bacterial spot, early blight, late blight, mold, partial, spot, Two-spotted, spider mites, target spot, mosaic virus, yellow leaf curl virus	Double GAN first images change into low-resolution by using(WGAN) and after that (SRGAN) used to reconstruct images into high-resolution.	VGG16, DenseNet121, and ResNet50	VGG16 - 98.58 ResNet50 – 99.14 DenseNet121-99.70	2021

3. MACHINE LEARNING

The machine learns itself without any programming. A child learns from its experience, in the same way, a machine learns from data to make predictions. ML is important because the human brain does not perform computation for big data. There are many models in ML which are used according to the type of the data. Supervised, Unsupervised, Semi-supervised, and Reinforcement learning is the different ML models. The term is called Supervised learning when the target class is labeled. The dataset which is a part of original big data is trained on algorithms, which help algorithms to learn and understand the relationshipbetween independent and target classes. When we do not know which class our target belongs to, then it is an example of unsupervised learning. We can label thousands of data manually, but if data is in million, it is impossible to label it. To resolve this issue, we can use Unsupervised learning. As a result of Unsupervised learning, the relationship between data points is obtained because of pattern matching. The structure of data is very complex in some cases, in those cases finding the best result between data points is a difficult task. We, humans, learn from our past mistakes because we get a punishment or reward, the same analogy is followed by Reinforcement Learning. When an algorithm finds a solution, then it is called a reward. If it fails to determine the solution then it becomes become a punishment. Machine Learning all models all models are shown in “figure-1”.

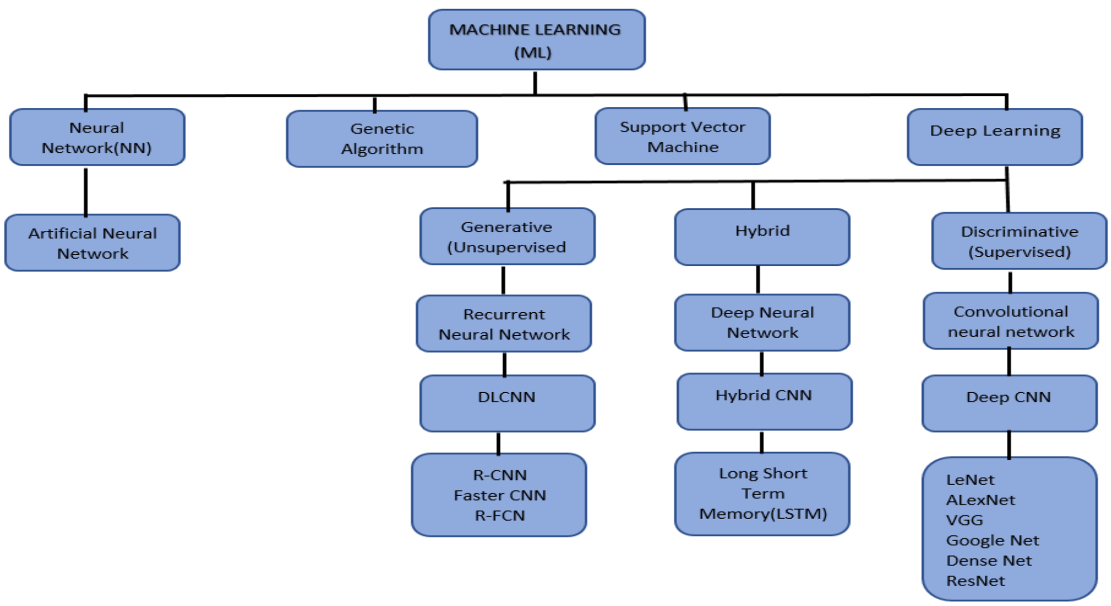

Fig-1. Machine Learning Models

4. DEEP LEARNING

Deep learning handles complex data with the help of hidden layers. It was first programmed in 1943. In 1960 backpropagation was added to it. In the year 1970, the concept of backpropagation was deployed. In the year 1979 Convolution Neural Network (CNN) came into the picture. In 2000 some neural networks face the problem of vanishing gradient descent. To overcome the vanishing gradient problem, Long-Short Term Memory came into the existence. In later years, the dataset started to increase at a high rate. Therefore, in 2009 ImageNet was discovered. In the year 2012, the CNN model- AlexNet played a major role by comprising different CNN layers. In 2014 Generative Adversarial Network (GAN) was developed by Google. GAN is very important when we have an imbalanced dataset because it is popular for the augmentation of the dataset. In the same year, GoogleNet was developed to decrease the error rate while training the images. All deep learning modeldss that are developed by time is as shown "Figure 2".

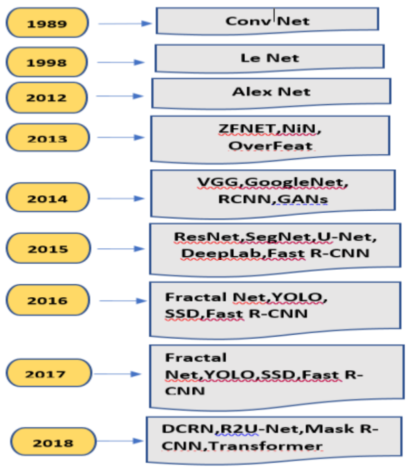

Fig-2 Deep Learning Models

5. CONVOLUTION NEURAL NETWORK

Our nervous system works on axons which pass signals to the brain, a similar analogy is followed concerning neural networks but the only difference here is instead of the axon, a perceptron does this task. The data cleaning and transformation is used as the data pre-processing technique. And Data augmentation and image segmentation are also a partof the pre-processing technique. Preprocessing is followed by the Feature extraction. It is very important when the data comprises images because a lot of features are present in image data. Reducing the dimensionality from the

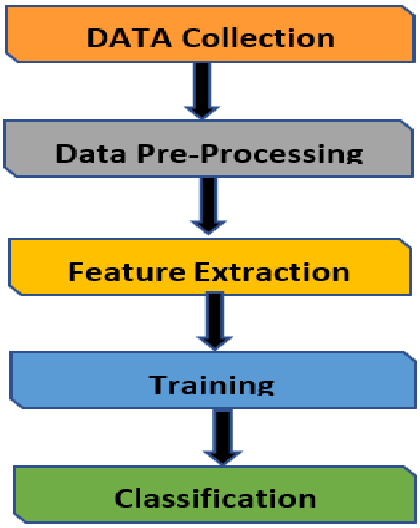

Fig-3. CNN MODEL

extracted features was a difficult task. So, the CNN layers are used in dimensionality reduction and also to extract the best features having an impact on the model's performance. There are many models of CNN which have different working architectures: VGG-16, VGG-19, DensNet, and LeNet. The steps required for modeling in CNN is shown in "Figure-3".

6. CONCLUSION

The incidents of Crop failure are very high. Farmers face huge losses due to crop failure. Farmers also find it difficult to appoint a person to carefully examine the plants and also provide the remedial solution for handling the leaf diseases. The rise of Neural Network techniques has a significant impact on disease detection. In this paper, We have summarized the various techniques utilized for plant disease detection. Discussion of machine learning techniques are performed but the problem with machine learning technique is that the spatial information of the image is not processed effectively. And the paper also briefs about Image processing and Segmentation and Deep Learning models for leaf disease detection.

7. ACKNOWLEDGMENT

I thank my project guide Dr. Vinutha N for her support and guidance in successfully completing this paper.

8. REFERENCES

[1] Manisha Bhangra,∗ , H.A.Hingoliwalab, "Smart Farming: Pomegranate Disease Detection Using Image Processing", Elsevier, 2015, Volume 58, Pages 280-288.

[2] Monika Jhuria, Ashwani Kumar, Rushikesh Borse, "Image Processing For Smart Farming: Detection of Disease and Fruit Grading", IEEE, 2013, Proceedings of the 2013 IEEE Second International Conference on Image Information Processing.

[3] Shiv Ram Dubey, Anand Singh Jalal, "Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns", IEEE, 2012, Third International Conference on Computer and Communication Technology.

[4] Mohanty, S.P., Hughes, D.P., Salathe, M.," Using deep learning for image-based plant disease detection", Frontiers in plant science, 2016, Volume 7, Article-1419.

[5] Rangarajan, A.K., Purushothaman, R., Ramesh, A, "Tomato crop disease classification using a pre-trained deep learning algorithm Procedia computer science", Elsevier,2018, Volume 133, 2018, Pages 1040-1047.

[6] R. Sujatha, Jyotir Moy Chatterjee, NZ Jhanjhi, Sarfraz Nawaz Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection", Elsevier, 2021, Volume 80, February 2021, 103615.

[7] Miaomiao Ji, Lei Zhang, Qiufeng Wu," Automatic Grape Leaf Diseases Identification via UnitedModel Based on Multiple Convolutional Neural Networks", Elsevier, Volume 7, Issue 3, September 2020, Pages 418-426.

[8] Ashraf Darwish, Dalia Ezzat, Aboul Ella Hassanien, "An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis", Elsevier, Volume 52, February 2020, 100616.

[9] UDAY PRATAP SINGH 1, SIDDHARTH SINGH CHOUHAN 2, (Student Member, IEEE), SUKIRTY JAIN 3, AND SANJEEV JAIN 2, "Multilayer

Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease", Volume: 7 Page(s): 43721 - 43729

[10] - Parul Sharma, Yash Paul Singh Berwal, Wiqas Ghai, "Performance Analysis of Deep Learning CNN Models for Disease Detection in Plants using Image Segmentation", Volume 7, Issue 4, December 2020, Pages 566-574.

[11] JUN SUN, YU YANG, XIAOFEI HE, AND XIAO HONG WU, "Northern Maize Leaf Blight Detection Under Complex Field Environment Based on Deep Learning", IEEE, 2020, Volume: 8 Page(s): 33679 – 33688.

[12] Pranjali B. Padol; Anjali A. Yadav," SVM classifier based grape leaf disease detection", CASP, June 2016, Conference on Advances in Signal Processing (CASP),DOI: 10.1109/CASP.2016.7746160

[13] Ümit ATİLAa*, Murat UÇARb, Kemal AKYOLc, Emine UÇAR, "Plant leaf disease classification using EfficientNet deep learning model", Elsevier, Volume 61, March 2021, 101182.

[14] G. Sambasivam, Geoffrey Duncan Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks", Elsevier, Volume 22, Issue 1, March 2021, Pages 27-34.

[15] ASAD KHATTAK, MUHAMMAD USAMA ASGHAR, ULFAT BATOOL, MUHAMMAD ZUBAIR ASGHAR, HAYAT ULLAH, MABROOK AL-RAKHAMI, AND ABDU GUMAEI 3, "Automatic Detection of Citrus Fruit and Leaves Diseases Using Deep Neural Network Model", IEEE,2021, Volume: 9 Page(s): 112942 – 112954

[16] SUNIL C. K., JAIDHAR C. D., AND NAGAMMA PATIL, "Cardamom Plant Disease Detection Approach Using EfficientNetV2", IEEE,2022, Volume: 10) Page(s): 789 – 804

[17] ZINON ZINONOS, SOCRATES GKELIOS, ALA F. KHALIFAH, DIOFANTOS G. HADJIMITSIS, YIANNIS S. BOUTALIS, AND SAVVAS A. CHATZICHRISTOFIS, "Grape Leaf Diseases Identification System Using Convolutional Neural Networks and LoRa Technology", IEEE,2021, Volume: 10 Page(s): 122 - 133

[18] Kundu Kashyap Chakraborty, Rashmi Mukherjee, Chandan Chakroborty, Kangkana Bora, "Automated recognition of optical image-based potato leaf blight diseases using deep learning", Elsevier,2022 Volume 117, January 2022, 101781.

[19] MOBEEN AHMAD, MUHAMMAD ABDULLAH, HYEONJOON MOON, AND DONGIL HAN, "Plant Disease Detection in Imbalanced Datasets Using Efficient Convolutional Neural Networks With Stepwise Transfer Learning", IEEE,2021, Volume: 9 Page(s): 140565 – 140580

[20] SK MAHMUDUL HASSAN AND ARNAB KUMAR MAJI, (Member, IEEE), "Plant Disease Identification Using a Novel Convolutional Neural Network", IEEE,2022, Volume: 10 Page(s): 5390 – 5401

[21] CHANGJIANG ZHOU, ZHIYAO ZHANG, SIHAN ZHOU, JINGE XING, QIUFENG WU, AND JIA SONG1, "Grape Leaf Spot Identification Under Limited Samples by Fine Grained-GAN", IEEE,2021, Volume: 9 Page(s): 100480 – 100489

[22] STEFANIA BARBURICEANU, SERBAN MEZA, (Member, IEEE), BOGDAN ORZA, RAUL MALUTAN, AND ROMULUS TEREBES, (Member, IEEE), "Convolutional Neural Networks for Texture Feature Extraction. Applications to Leaf Disease Classification in Precision Agriculture", IEEE, 2021, Volume: 9, Page(s): 160085 - 160103

[23] Junde Chena, Jinxiu Chena, Defu Zhanga,□, Yuandong Sunb, Y.A. Nanehkarana," Using deep transfer learning for image-based plant disease identification", Elsevier, June 2020, Volume 173, June 2020, 105393.

[24] S.Hernández[ab]Juan L.López[b] "Uncertainty quantification for plant disease detection using Bayesian deep learning", Elsevier, November 2020, Volume 96, November 2020, 106597.

[25] Abdul Waheed, Muskan Goyal, Deepak Gupta, Ashish Khanna, Aboul Ella Hassanien, Hari Mohan Pandey, "An optimized dense convolutional neural network model for disease recognition and classification in corn leaf,' Elsevier, 2020, Volume 175, August 2020, 105456.

[26] MohitAgarwal[a]AbhishekSingh[b]SiddharthaArjaria[c]AmitSinha[d]SuneetGupta[a,]" ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network", Elsevier, 2020, Volume 167, 2020, Pages 293-301.

[27] Ümit ATİLAa*, Murat UÇARb, Kemal AKYOLc, Emine UÇAR," Plant leaf disease classification using EfficientNet deep learning model", Elsevier, 2021, Volume 61, March 2021, 101182.

[28] Yafeng Zhao, Zhen Chen, Xuan Gao, Wenlong Song, Qiang Xiong, Junfeng Hu, Zhichao Zhang, "Plant Disease Detection using Generated Leaves Based on Double GAN", IEEE, 2021, IEEE/ACM Transactions on Computational Biology and Bioinformatics (Early Access) Page(s): 1 - 1

[29] O.E. Apolo-Apoloa, J. Martínez-Guantera, G. Egeaa, P. Rajab, M. Pérez-Ruiza," Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV", Elsevier, 2020, Volume 115, April 2020, 126030.

[30] Geetharamani G. a, Arun Pandian J. b, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network", Elsevier, 2020, Volume 76, June 2019, Pages 323-338.