DEEP LEARNING BASED PLANT SPECIES CLASSIFICATION USING DIGITAL MORPHOMETRIC MANAGEMENT

Devi Mahalakshmi S[1], Dr.Vanitha Sivagami S[2]

[1]Mepco Schlenk Engineering College (Autonomous), Sivakasi-626005
[2]Department of Computer Science and Engineering, Mepco Schlenk Engineering College,Sivakasi,Sivakasi-626005,Tamil Nadu, India

sdevi@mepcoeng.ac.in[1]

Abstract:
The number of plant species is extremely huge all over the world. Hence, it is impossible to identify and classify all the species. Plant species may be similar to each other, taking a long time to differentiate them. Hence, there is a need to develop an automated system. Deep learning has been widely employed for classification and recognition tasks in the biological fields. An automated plant species identification system could help botanists. Deep learning is used for feature extraction as it provides deeper information of images. Plant species identification includes preprocessing, feature extraction, and classification. Here we used many algorithms to analyze the image. The paper aims at understanding preprocessing, extracting the leaf features using contours, and classifying them using deep-learning techniques, namely CNN (AlexNet). It presents the leaves of various plant species from which the vein characteristics are extracted and presented to detect and classify various kinds of plant species and other artificial intelligence techniques used to perform pattern recognition. Here the classifier achieved the best performance of 82% for accuracy.

Keywords: Deep Learning, Convolutional Network, Leaf Vein Morphometric, Feature Extraction, Classification, Artificial Neural Network

I.INTRODUCTION

Plant species classification is an interesting field in which Deep Learning is applied. Plant species classification is done by extracting the features of plants. The most common feature used to develop automated plant classification systems is leaf shape. Other than shape, leaf features like color, texture, and veins can also provide additional features. Some plant species are very similar to one another, and it can take a long time to classify them. So the need for an automated system to identify and classify plants is required. CNN is applied to extract the features from leaf images. Three CNN models were used, such as the pre-trained AlexNet CNN model, the fine-tuned AlexNet CNN model, the DLeaf CNN model, and the proposed modified DLeaf CNN model. The extracted features were fed into a classification. Before the feature extraction, we performed vein morphological measurements and image preprocessing.

II.RELATED WORKS

Malarvizhi K *et al.* (2021) [1] proposed a shape-based approach for leaf vein morphometrics with several steps including sampling, image preprocessing, edge detection, feature extraction, classification, comparison, etc. A random forest classifier achieved the best performance of 90% for accuracy. However, SVM and KNN classifiers obtained an accuracy of not less than 85%. The data set used for this work is the Flavia dataset, which contains about 1907 images of 32 different species.

UmitAtila et al. (2021)[2] The dataset used for this work is the PlantVillage dataset, which contains 38 classes and 54,305 images of 14 different plant species in total, including 12 healthy and 26 diseased species The B5 model gave 99.91% accuracy and 98.42% precision in the original dataset, while the B4 model gave 99.97% accuracy and 99.39% precision in the augmented dataset.

Azath M et al. (2021)[3] proposed a cotton leaf disease and pests using the deep learning technique. used methodologies are Cotton Image Sample Digitization, Image Data Preprocessing, Feature Extraction, Dataset

Partitioning, and Model Selection Methodology. captured 2400 images distributed into four equal classes, such as bacterial blight, healthy, leaf miner, and spider mite. This used to dataset. The accuracy of this model is 96.4% for identifying classes of leaf diseases and pests in cotton plants

Paul Shekonya Kanda et al. (2021)[4] proposed a deep learning approach to classifying plants via leaf images. A Conditional Generative Adversarial Network was used to generate synthetic data, a Convolutional Neural Network was used for feature extraction, and the rich extracted features were fed into a classifier for efficient classification of the plant species. It gave 96.1% accuracy. The dataset used for this work is the MalayaKew dataset.

Jiang Huixian et al. (2020)[5] extract plant leaf features and identify plant species based on image analysis. Plant leaf images are first segmented using various methods, and then leaf shape and texture features are extracted from leaf sample images using a feature extraction algorithm. 50 plant leaf databases are tested and compared with KNN-based neighborhood classification, Kohonen network based on self-organizing feature mapping algorithm and SVM-based support vector machine.Dataset collected from the local garden.It contains 50 plant leaves and photographed,. Each leaf took an average of 200 samples, of which 180 were taken as training samples and 20 as test samples. Here, the BP Neural Network algorithm gives a high recognition rate (92.47%) compared to others.

Sujith*et al.* (2020)[6] used PHOG, LBP, and GLCM feature extraction techniques. The proposed result gives an average accuracy of 98.23%. Flavia and Swedish Leaves were used to evaluate the proposed work.

XiaojunJin*et al.* (2021)[7] proposed an approach to identify weeds in vegetable plantations using deep learning and image processing. A CenterNet model was trained to detect vegetables. The trained CenterNet achieved a precision of 95.6% and a recall of 95.0%.

Bor-HorngSheu et al. (2020)[8] performed the binary leaf contour is extracted by normalized sampling of leaf length and width through image preprocessing, and then the special geometric features of leaves, such as morphological convex hull, centroid-contour distance, and serrated shape segmentation, are extracted. A CCD converts the geometry of a leaf to a centroid distance and captures the distance (di) from the centroid point to its leaf contour.

S. Santhosh et al. (2020)[9] performed leaf image features to classify the species by using Support Vector Machines (SVM). This classification helps to recognize and identify the population of endangered and extint species for preservation. One or more planes separate the two or more different classes of data. The model is framed across 99 species of data with an SVM classifier. The RBF kernel makes the model looser to avoid poor fitting and *indicates the influence* of the data.

Lawrence C. Ngugi*et al.* (2021)[10]Proposed a work done in plant disease recognition using IPTs. Methodologies used are AlexNet, GoogLeNet, Inceptionv3, SqueezeNet, ResNet-101, VGG16, ShuffleNet, InceptionResnetv2, MobileNetv2, DenseNet201, and InceptionResnetv3. The dataset used in this comparative study contains 54,305 leaf images representing 14 crop species and 26 diseases. DenseNet201 gives more accuracy (0.9973).

Muammer Turkoglu et al. (2019)[11] proposed a fine-tuned based on plant disease and pest images.GoogleNet, ResNet50, ResNet101, Inceptionv3, InceptionResNetv2, and SqueezeNet are among the methodologies employed.The ResNet50 model and SVM classifier produced the highest accuracy among these other deep models, with an accuracy of 97.86%. The dataset was obtained by plant protection academics from the Agricultural Faculty of Bingol University and nonu University in Turkey.It contains 1965 images.

Munish Kumar et al. (2019)[12] proposed a plant leaf recognition system using morphological features. Adaptive boosting methodology has been presented. For classification, KNN, decision trees, and multilayer perceptrons are used. Used methodology was AdaBoost methodology. The used methodology was the AdaBoost methodology. A maximum precision rate of 95.42% has been achieved for 32 kinds of plant leaves.

Dhananjay Bisen *et al.* (2019)[13] proposed a plant species recognition through features of leaf. A Swedish leaf dataset was used in this project. It had an average accuracy of 97%. Some categories are classified correctly with an accuracy of 100%. The used methodology was CNN.

Dhaya*et al.* (2020)[14] determine FO disease in the tomato plant leaves.They are achieved better accurate results due to the two-factor identification method.used classification method is Naive Bayes.Datasetconsists of 87k images with 60% affected leaves images, 40% healthy plant leaves. Used algorithm is found the disease with 96% accuracy.

Saleem, G., et al.(2019)[16] proposed a automated analysis of visual shape features for plant species classification. Proposed algorithms are k-nearest neighbour (KNN), decision tree, naive Bayes, and multi-support vector machines (SVM). The proposed algorithm is evaluated on the standard dataset "Flavia" of 1600 leaf images and on a self-collected dataset of 625 leaf images. The proposed algorithm gives precision and recall values of 97.6% and 98.8%, respectively, when tested on the "Flavia" dataset. The proposed technique was also tested on their self-collected dataset, giving respectively 96.1% and 97.3% precision and recall measure results.

III.PROPOSED WORK

This plant species classification system consists of three steps. The three steps are: image preprocessing, feature extraction, and classification. In image preprocessing, the leaf images in a dataset are preprocessed for further analysis. The image processing step contains three steps: RGB to Grayscale conversion, Adaptive Thresholding, and Skeletonization. Using these steps, the leaf images are preprocessed. After the image processing, the next step is the feature extraction. From the preprocessed leaf images, the features are extracted using the convolutional neural networks or models: pretrained AlexNet, finetuned AlexNet, DLeaf, and Modified DLeaf. After feature extraction, using the extracted features, the leaves are classified into which species they belong to. So the third and last step in the plant species classification system is classification. The classification also uses the same convolutional networks or models as pretrained AlexNet, finetuned AlexNet, DLeaf and modified DLeaf. The proposed steps in the plant species classification system are shown in fig.1.

1.Image preprocessing:

For Analysis purpose we do the preprocessing methods that are

1.Converting the color image to gray image
2.Adaptive thresholding is done on the gray image
3.Skeletonization is done on the image.

First, all leaf images were converted from RGB images into grayscale images. The converted grey image is shown in Fig.2. After converting the colour images to grayscale images, edge detection is done on the leaf images. Adaptive thresholding is an edge detection method, which is a method where the threshold value is calculated for smaller regions, and therefore, there will be different threshold values for different regions. Adaptive thresholding is used to segment out the images.The edge detected image is shown in Fig. 3. After segmentation, the images were then post processed and skeletonized to verify that a clean vein architecture could be obtained. The skeletonized image is shown in Fig.4.

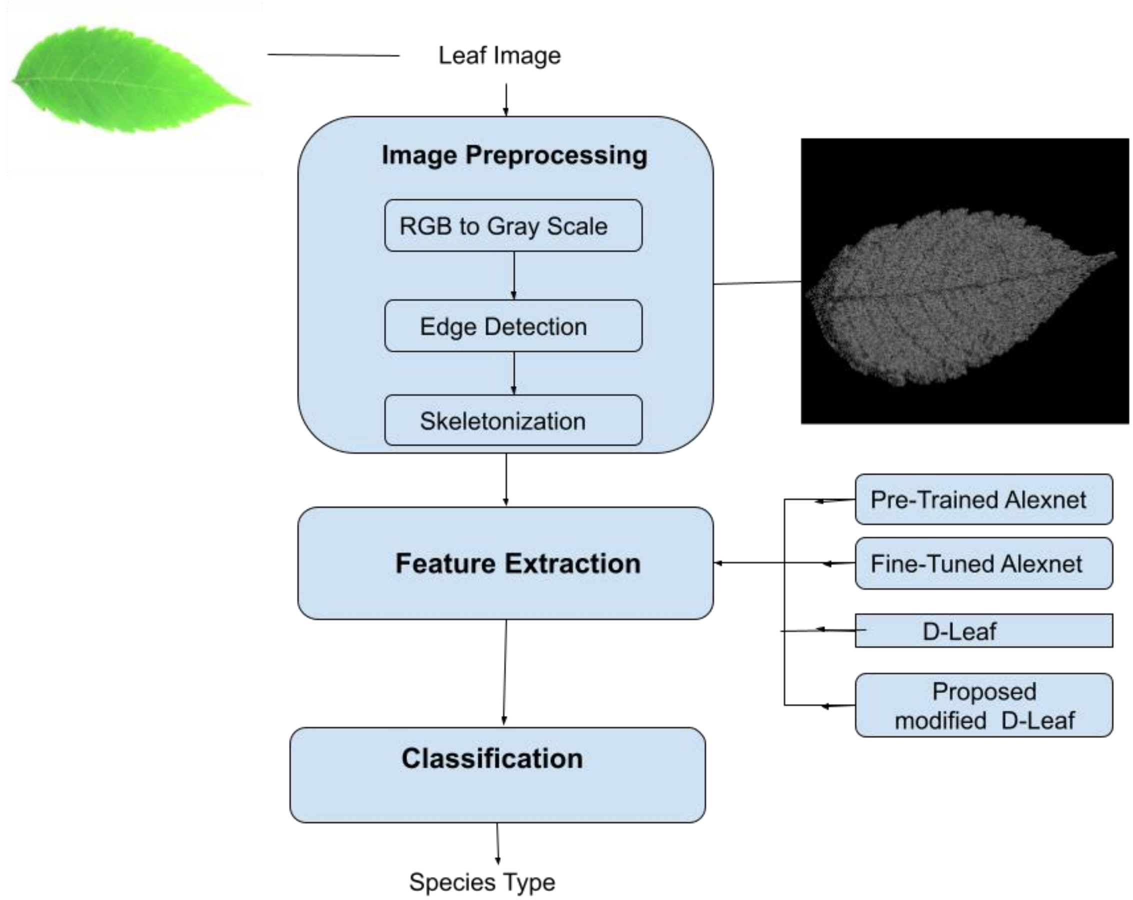

Fig. 1. Proposed System for Plant Species Classification

Fig. 2. Grayscale Image.

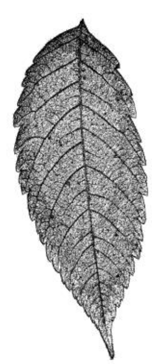

Fig. 3. Adaptive threshold.

Fig. 4. Skeletonized image.

2.Feature Extraction

In plant species classification, after the preprocessing, the next step is feature extraction. Convolutional Neural Network Models or Networks for Feature Extraction Pretrained AlexNet,FineTuned AlexNet ,DLeaf and modified DLeaf have been fine-tuned.

Convolutional Neural Networks (CNNs):

A typical CNN layer is composed of three layers: the first layer is the convolution stage, the second layer is the Activation Function, and the third layer is the Maxpooling layer. The first layer is the convolution layer, which performs convolution operations on an input image. The second layer is a non-linear activation function such as the Rectified Linear Unit (ReLU). It is also called a piecewise linear function. The third layer is the pooling layer stage, which employs a function to simplify and summarise the information of the particular layer. In this plant species classification model, we are using a kind of CNN model to classify the plant species.

Pre-trained AlexNet Model:

For Feature Extraction: First, using the pre-trained Alex net. The preprocessed image is given to AlexNet for feature extraction. Because AlexNet accepts input sizes of 227*227, the input image is resized to 227*227 *3.AlexNet was trained with over 1000 different classes of images from ImageNet. A pre-trained CNN AlexNet model was made up of nine layers, which included five convolution layers, three fully connected layers, and a softmax classification layer. Three fully-connected layers with 4096, 4096, and 1000 neurons, respectively, The AlexNet architecture is shown in Fig. 5. The parameters of the pretrained Alexnet layer are given in the table 1.

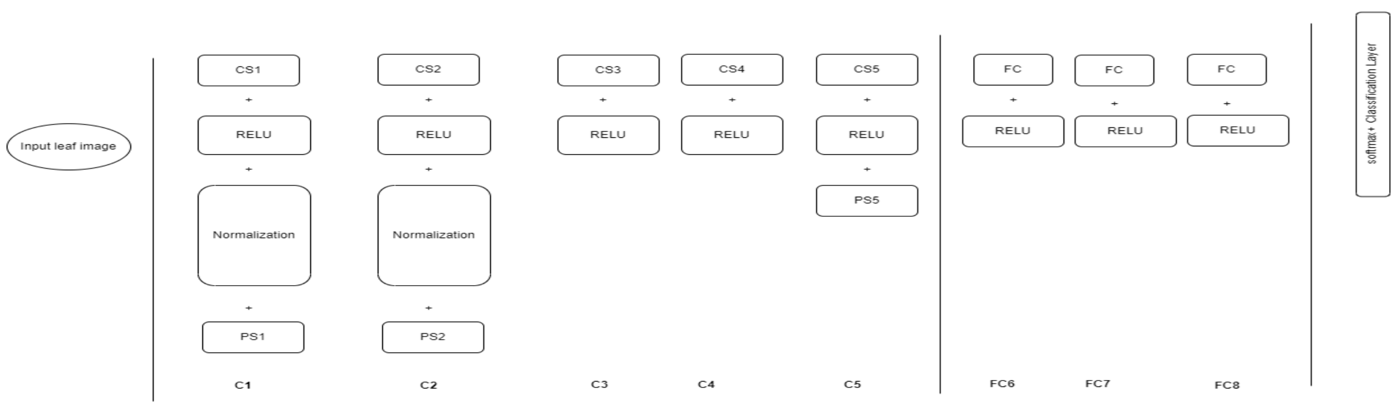

Fig. 5. AlexNet Architecture. CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, and FC - Fully-Connected Layer:

CNN Layer	CS1	PS1	CS2	PS2	CS3	CS4	CS5	PS5	FC6	FC7	FC8
Filter Size	11*11	3*3	5*5	3*3	3*3	3*3	3*3	3*3	_	_	_
No Of Kernel	96	_	256	_	384	384	256	_	4096	4096	1000
Size Of Stride	[4 4]	[2 2]	[1 1]	[2 2]	[1 1]	[1 1]	[1 1]	[2 2]	_	_	_

TABLE 1.Parameters for PreTrained AlexNetLayers

Fine-tuned AlexNet Model:
The architecture of the finetuned AlexNet model is the same as the original AlexNet. The input image size is 227*227*3.The first convolutional stage was fine-tuned with the 7 x 7 filter size and [2 2] stride. Also, three fully connected layers were fine-tuned. Those values are 1290, 1290, and 43. The other parameters of this model

CNN Layer	CS1	PS1	CS2	PS2	CS3	PS3	FC4	FC5	FC6
Filter Size	11*11	2*2	5*5	2*2	4*4	2*2	_	_	_
No Of Kernel	64	_	96	_	256	_	1290	1290	43
Size Of Stride	[4 4]	[2 2]	[2 2]	[2 2]	[1 1]	[2 2]	_	_	_

remained unchanged as in the AlexNet models. The parameters of the FineTuned Alexnet layer are given in table 2.

TABLE 2.Parameters for Fine-Tuned Alexnet Layers

DLeaf Model:

CNN Layer	CS1	PS1	CS2	PS2	CS3	CS4	CS5	PS5	FC6		FC7	FC8
Filter Size	7*7	3*3	5*5	3*3	3*3	3*3	3*3	3*3	_		_	_
No Of Kernel	96	_	256	_	384	384	256	_	1290		1290	43
Size Of Stride	[2 2]	[2 2]	[1 1]	[2 2]	[1 1]	[1 1]	[1 1]	[2 2]	_		_	_

For the plant species classification system, here we are using another model named Dleaf. The Dleaf model consists of six layers. The six layers are three convolutional layers and three fully connected layers. The convolutional layer is followed by the convolution stage, relu, and pooling stage. After a fully connected layer, a softmax layer is used. Fully connected, it contains 1290, 1290, and 43 neurons, respectively. The parameters of the Dleaf layer are given in Table 3. Dleaf architecture is shown in Fig. 6.

TABLE 3.Parameters of DLeaf Layers

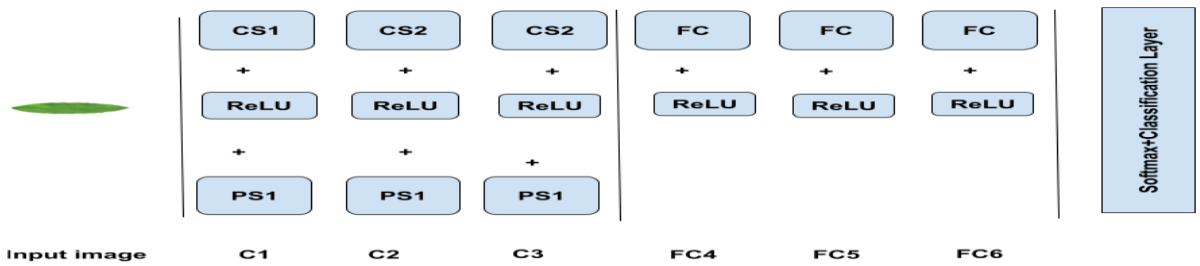

Fig. 6.DLeaf Architecture. CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, and FC - Fully-Connected Layer:

Proposed Modified DLeaf model:

This CNN model was developed to extract the features from the images instead of fine-tuning the AlexNet model. It consists of three convolution layers and a softmax classification layer. Here in the proposed model D-leaf architecture, we are doing some modifications in order to make the accuracy high compared to the given model. The proposed model D-leaf architecture is shown in Fig.7. Here we add some layers, like using a flat layer and a dense layer. A flattening layer is used to convert the output which we are getting from the multiple layer into a 1D array. The dense layer receives information from the flat layer. also referred to as a fully connected layer, is a layer that is used as the final stage of the neural network. This layer helps in changing the dimensionality of the output from the preceding layer so that the model can easily define the relationship between the values of the data set on which the model is working. Table 3 displays the parameters of the DLeaf Layers.

CNN Layer	CS1	PS1	CS2	PS2	CS3	PS3
Filter Size	5*5	2*2	5*5	2*2	3*3	2*2
No Of Kernal	128	_	64	_	32	_

TABLE 4.Parameters of Proposed model modified D-Leaf Layers

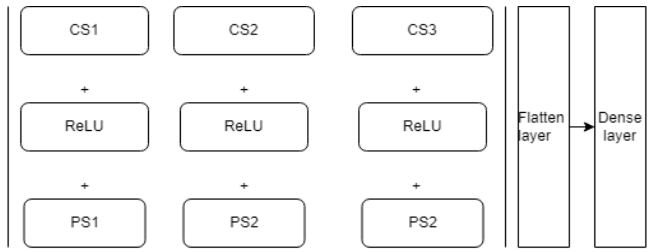

Fig. 7. Proposed model modified DLeaf Architecture. CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, and FC - Fully-Connected Layer:

3.Classification:

Classification, as the last step for an automated plant recognition. Here we are classify the plant species using the same pre trained alexnet ,fine-tuned alexnet and modified Proposed Model D-Leaf .Softmax classification layer was used to classification.

IV. RESULTS AND DISCUSSION

In this work, the proposed automated plant identification system has been tested on the Flavia dataset containing leaf images of 32 plant species, totaling 1907 images. Feature extraction was done by three CNN models. For performance evaluation, the data was partitioned into training and testing sets in the ratio of 80:20 and fitted to the same CNN model classifiers. The performance of each CNN model classifier is discussed in Table 5. A training dataset is utilized to train the model, and it helps the model gain knowledge related to leaf images. It is observed that the proposed Model DLeaf classifier achieved the best performance of 82% for accuracy. However, the other classifiers obtained an accuracy of not less than 70% (PreTrained AlexNet, FineTuned AlexNet, and DLeaf). Hence, in this research, we can claim that the proposed Model Dleaf is highly suitable with contours for feature extraction models. In contrast, pre-trained AlexNet and fine-tuned AlexNet were less suitable for the contours. On the other hand, the conventional feature extraction method extracts each type of feature separately and manually, consuming a lot of time. For example, if vein features are considered in this work, different sets of processes will be required for segmentation followed by vein feature extraction. Thus, the machine learning approach is more practical and appropriate than conventional methods for developing an automated plant species classification system.

CNN MODELS	ACCURACY
D-Leaf	**80%**
Proposed Model modified D-leaf	**82%**

TABLE 5. Classification accuracy of different CNN models

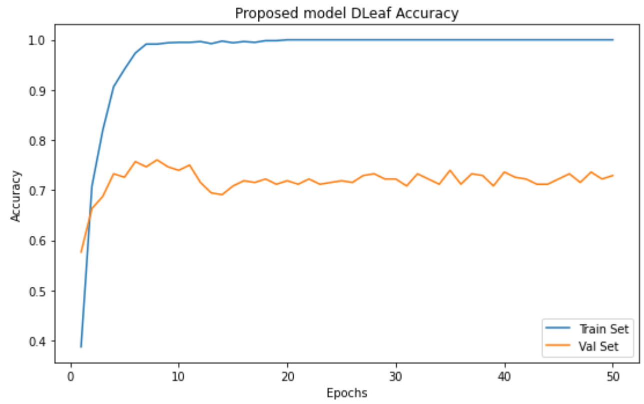

Fig.8.Training accuracy and validation accuracy of the proposed model modified DLeaf.

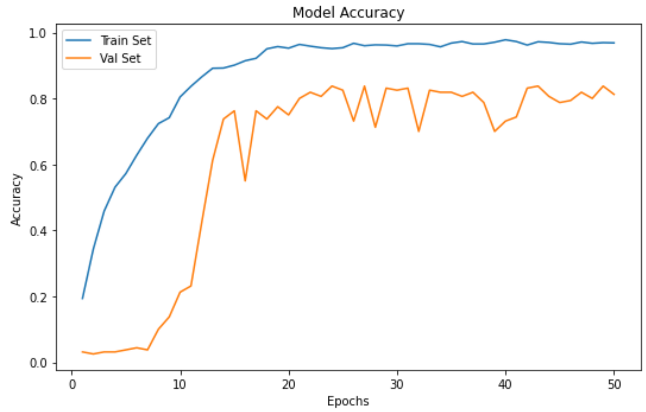

Fig.10.Training accuracy and validation accuracy of the DLeaf.

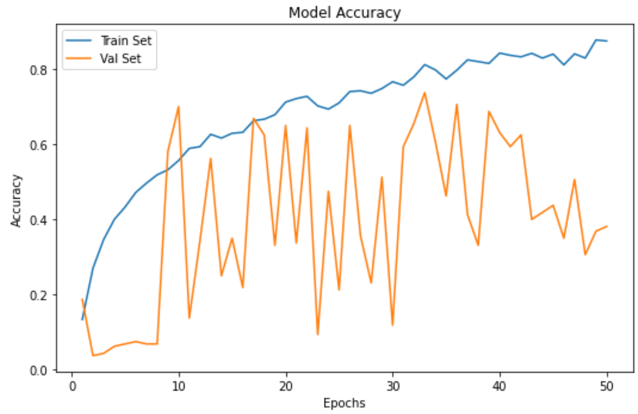

Fig.11.Training accuracy and validation accuracy of the FineTuned AlexNet model.

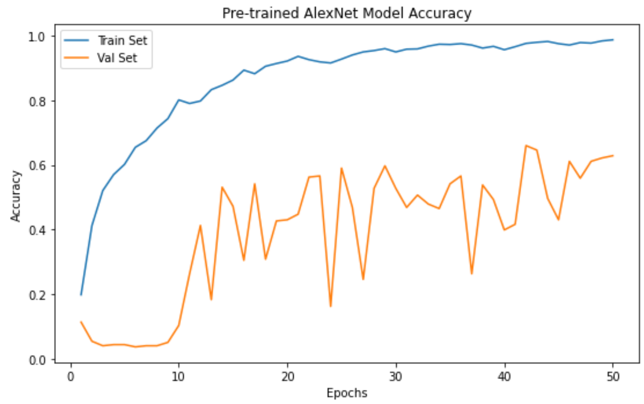

Fig.12.Training accuracy and validation accuracy of the Pretrained AlexNet.

V.CONCLUSION

We proposed a plant species classification model using leaf vein features using deep learning techniques. Here we are doing several steps like image preprocessing, feature extraction and classification. Image preprocessing steps RGB to Gray, Adaptive thresholding and Skeletonization. For Feature extraction using deep learning models pretrained alexnet , fine tuned alexnet and modified D-leaf .Extracted features are classified using the same deep learning models. In this compare the accuracy of these models. The Dleaf model give more accuracy compared to other models pretrained and fine-tuned alexnet.

V.REFERENCES

[1]MalarvizhiK,SowmithraM,Gokula Priya D,Kabila B,"Machine Learning for Plant Species Classification using Leaf
Vein Morphometric"International Journal of Engineering Research & Technology (IJERT)http: ISSN: 2278-0181 IJERTV10IS040273 (This work is licensed under a Creative Commons Attribution 4.0 International License.)Published by :www.ijert.org,Vol. 10 Issue 04, April-2021.

[2]UmitAtila, Murat Ucar , Kemal Akyol , EmineUcar,"
Plant leaf disease classification using EfficientNet deep learning model" Ecological Informatics 61 (2021) 101182.

[3]Azath M,MeleseZekiwos,andAbey Bruck"Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis" Journal of Electrical and Computer Engineering,Volume 2021, Article ID 9981437, 10 pages
https://doi.org/10.1155/2021/9981437.

[4] P S Kanda, Kewen Xia, Olanrewaju Hazzan Sanusi,"A Deep Learning-Based Recognition Technique for Plant Leaf Classification*Digita"l Object Identifier 10.1109/ACCESS.2021.3131726.*

[5] Jiang Huixian" The Analysis of Plants Image Recognition Based on Deep Learning and Artificial Neural Network". *Digital Object Identifier 10.1109/ACCESS.2020.2986946.*

*[6]*A. Sujith and R. Neethu," Classification of Plant Leaf Using Shape and Texture Features", 28 October 2020.

[7] XiaojunJin, Jun Che, YongChen,"Weed Identification Using Deep Learning and Image Processing in Vegetable Plantation" *Digital Object Identifier 10.1109/ACCESS.2021.3050296.*

[8]B.H.Sheu,F.SChou,"A Study On Species Identification Based On Leaf Contours Of Taiwan Lauraceae and Fagaceae Plants",2020.

[9] S Santhosh,A Fahima Zulfath,"Classification of Leaf Images For Species Identification",June 2020.

[10] Lawrence C. Ngugi, MoatazAbelwahab, Mohammed Abo-Zahhad," Recent advances in image processing techniques for automated leaf pest and disease recognition"

INFORMATION PROCESSING IN AGRICULTURE 8 (2021) 27– 51.
[11] Muammer Turkoglu, DavutHanbay,” Plant disease and pest detection using deep learning-based features”, Turk J Elec Eng& Comp Sci (2019) 27:1636 – 1651 © TÜBİTAK doi:10.3906/elk-1809-181.
[12] Munish Kumar, Surbhi Gupta , Xiao-Zhi Gao, And Amitoj Singh,” Plant Species Recognition Using Morphological Features and Adaptive Boosting Methodology”, *Digital Object Identifier 10.1109/ACCESS.2019.2952176.*
[13] Dhananjay Bisen,” Deep convolutional neural network based plant species recognition through features of leaf”,# Springer Science+Business Media, LLC, part of Springer Nature 2020.
[14]Dhaya,”Flawless Identification of Fusarium Oxysporum in Tomato Plant Leaves by Machine Learning Algorithm”, journal of Innovative Image Processing (JIIP) 2020) Vol.02/ No. 04 Pages: 194-201.
[15] Azlah, Muhammad Azfar Firdaus, et al. "Review on techniques for plant leaf classification and recognition." Computers 8.4 (2019): 77.
[16] Saleem, G., et al. "Automated analysis of visual leaf shape features for plant classification." Computers and Electronics in Agriculture 157 (2019): 270-280.
[17] Nagamani H S,”Leaf Disease Identification By Extracting Gradient Local Ternary Pattern & Zernike Moment” Indian Journal of Computer Science and Engineering (IJCSE),Jan-Feb 2022.
[18]Chen, Junde, et al. "Using deep transfer learning for image-based plant disease identification." Computers and Electronics in Agriculture 173 (2020): 105393.
[19] Omeer, Abdulla & Deshmukh, Ratnadeep. (2021). Improving the classification of invasive plant species by using continuous wavelet analysis and feature reduction techniques. Ecological Informatics. 61. 101181. 10.1016/j.ecoinf.2020.101181.
[20] Shin, Jihoon& Yoon, Seong Hyeon & Kim, YoungWoo& Kim, Taeho& Go, ByeongGeon& Cha, Yoonkyung. (2021). Effects of class imbalance on resampling and ensemble learning for improved prediction of cyanobacteria blooms. Ecological Informatics. 61. 101202. 10.1016/j.ecoinf.2020.101202.
[21]M. A. Jasim and J. M. AL-Tuwaijari, "Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques," 2020 International Conference on Computer Science and Software Engineering (CSASE), Duhok, Iraq, 2020, pp. 259-265, doi: 10.1109/CSASE48920.2020.9142097.