



Deep Learning for Plant Species Classification

Dasari Akash¹, Kurivella Venkata Vijay Rahul², Kamisetty Markandeya Swamy³, Madala Revanth Venkata Sai⁴, Lankoji Venkata Samba Siva Rao⁵

^{1,2,3,4,5}Department of CSE Assistant professor KL University, Green Fields, Vaddeswaram, Andhra Pradesh 52230

Abstract:

Neither the ecology of the world nor the existence of humans can exist without plants. It is necessary to have an automated technique that makes use of deep learning in order to safeguard endangered species. In order to pre-process leaf images and extract important features, a new CNN-based technique known as D-Leaf was introduced. Specifically, this approach takes use of three distinct CNN models: D-Leaf, pre-trained AlexNet, and fine-tuned AlexNet. The support vector machine (SVM), artificial neural network (ANN), k-nearest neighbour (k-NN), naive bayes, and convolutional neural network (CNN) were the five various machine learning approaches that were used in order to establish the classification of these qualities. The D-Leaf model surpassed both the raw AlexNet model (93.26% accuracy) and the fine-tuned AlexNet model (95.54% accuracy) in the testing process. Additionally, the ANN classifier was an excellent choice for the CNN that was emphasised. According to the findings of the empirical research, D-Leaf has the potential to serve as an effective automated method for the classification of plant species.

Keywords: CNN (Convolutional Neural Network), ANN (Artificial Neural Network), Feature extraction, Classification, D-Leaf method, testing accuracy

1. Introduction

Worldwide, there are around 391,000 vascular plant species [1]. This is an enormously large number of plant species. As a result, no experienced botanist can hope to be able to name and categorize every species. It could take a long time to tell certain plant species apart because of how similar they are. Furthermore, the issue of extinction affects a great number of plant species. To lessen the likelihood of extinction, it is necessary to protect and conserve plant species, whether they are endangered or not. Therefore, a computerized or automated method for plant identification and classification is required. When creating these kinds of automated plant categorization systems, the shape of the leaf is by far the most popular trait to employ. The leaf may reveal more than just its shape; it can also reveal details like veins, colors, and textures.

The application of machine learning for identification and classification tasks has grown throughout industries, particularly in the life sciences, thanks to advancements in technology. There is a vast use of machine learning techniques in pattern recognition, including k-Nearest



Neighbour, Artificial Neural Networks, Support Vector Machines, and a great deal of other techniques.

Image-based methods for identifying animal species are promising. Customers can use their phone's camera to capture a plant image, which can then be identified using software. This process can be done even without technical knowledge, using a PC-assisted plant recognized proof framework. The goal is to obtain a complete image of a plant or its parts for classification studies.

Deep learning, a branch of AI, is now a popular and frequently used method with several applications in fields such as biology, medicine, computer vision, voice recognition, and more. Including both ends, the numerical range is 2–5. A state-of-the-art AI technique, deep learning provides a strong foundation for supervised learning [6]. Even when dealing with a big dataset, the system can quickly and efficiently connect an input vector to an output vector [6]. Deep Belief Networks (DBNs), Convolutional Neural Networks (CNNs), and similar architectures are some ways to categorise deep learning. Because of its superior ability to retrieve complicated and precise information, deep learning surpasses traditional machine learning approaches.

The goal of this study is to identify certain tree species by analysing images of their leaves using Convolutional Neural Networks (CNN). There were three different convolutional neural network (CNN) models used: the pre-trained AlexNet model, the fine-tuned pre-trained AlexNet model, and the suggested D-Leaf CNN model. In order to learn and make training easier, the collected characteristics were fed into several categorization algorithms. The segmentation of veins in leaves was assessed using a standard approach. Among the first attempts to classify tropical tree species using Convolutional Neural Networks (CNN), this article stands out.

2. Related Review

The basic steps of a system for automatically classifying plants are shown in Figure 1. At first, images of leaves are taken with the help of digital cameras, scanners, or other comparable equipment. These photographs undergo pre-processing to remove noise and improve quality. Noise, stemming from inaccuracies in pixel values during image capture, can distort image intensities. Image enhancement techniques are employed to highlight image characteristics [7]. Eliminating image noise is an essential process to accentuate or improve the significant characteristics of a picture. Afterwards, the area of interest (ROI) was separated from the photos, and then the features were extracted. Ultimately, the collected characteristics are inputted into a categorization or recognition system.

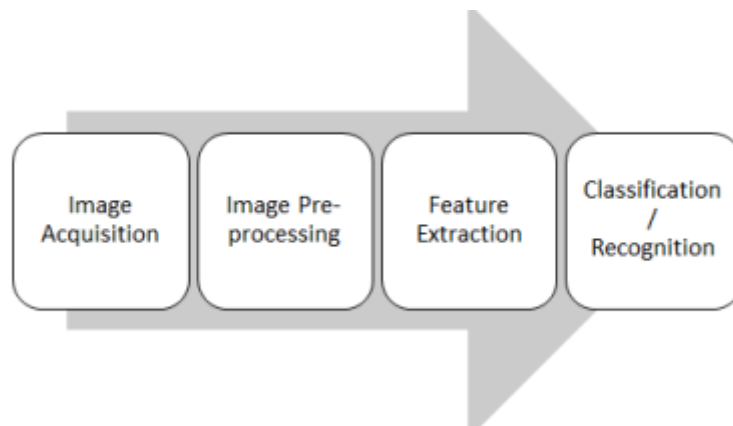


Figure 1: General Approach for Automated Plant Classification.

Leaves are often used to identify plant species since they are readily available throughout the year, particularly in tropical regions. Form, texture, venation pattern, color, and more may all be gleaned from just one leaf. Various approaches, including conventional morphometric measurement and machine learning, may be used to get each of these features. Nevertheless, there are ways to get two traits at once. Shape and texture, for instance, may be provided by Zernike Moment.

The main characteristic used in the development of systems for identifying plants is their morphology. Evaluating a form often involves taking its slenderness, roundness, compactness, and rectangularity into account, among other relevant characteristics. Because it allows for the categorization of leaves according to their surface structure, texture is an essential component of the system for identifying plants. Typically, the uneven spatial distribution pattern of varying picture intensities impacts certain pixels within an image [8, 9]. Different plant species may be identified based on the pattern of their veins. There are three distinct kinds of veins found in leaves: main, secondary, and tertiary. Since different kinds of plants have distinctive leaf colors, comparing them might help identify them. Even within the same species, you may find leaves with a wide range of colors. Take *Cinnamomum inners* as an example; its crimson juvenile leaves eventually become green.

2.1 Conventional Approach

Cope et al. [10] introduced a genetic algorithm (GA) and Ant Colony algorithm (AC) based evolving vein classifier for vein structure extraction. Almost all primary and secondary vein patterns can be extracted by the created vein classifier with little noise. Its performance in reconstructing highly discontinuous venation structures was superior to that of the Ant Colony algorithms. On the other hand, continual venation extraction is better handled by Ant Colony algorithms; but, they have the potential to generate larger and more interconnected noise regions, which might make real vein identification challenging. Put simply, the ant algorithms



couldn't compare to the evolved classifiers. Therefore, it's possible that a hybrid approach that incorporates both approaches will provide more accurate results.

Prior to this, Anami et al. [11] had suggested an identification method. We used the Sobel operator to extract the color histogram as a texture feature and the edge direction histogram as a color feature from a thousand photos of various plants, trees, and herbs. After feature extraction, a radial basis exact fit neural network (RBENN) and support vector machine (SVM) were used for training.

An automatic system for recognizing legume leaves was created by Larese et al. [12] using just the vein architecture. After applying some basic measures to the vein anatomy, a Random Forests method was used to identify them. After extracting 39 vein characteristics, the Random Forests were used for classification. In addition, the scientists found that employing a subset of 7 characteristics is just as effective as using all 39 traits.

Using the Flavia and Foliage datasets, Kadir et al. [13] suggested an alternative approach. Various models were developed utilizing various combinations of shape features, color moments, texture characteristics derived using GLCM, and vein features. The Polar Fourier Transform and three geometric features represent shape characteristics.

2.2 CNN-based Plant Identification

In the absence of structured data, the Polar Fourier Transform and three geometric aspects serve as substitutes. The unique feature filtering and picture visualization were done using a de-convolutional network (DL). The characteristics that were retrieved were further categorized using an SVM and a Multilayer Perceptron (MLP). D1 was the complete picture collection, while D2 was the set of leaf patches. A precision of almost 97% was attained by both sets of data. More than 91% accuracy was also reached by researchers in [17] when they integrated global and local features.

As an additional point, Sladojevic et al. [18] used CNNs for the purpose of herbicide detection. Thirteen distinct plant species were classified by the study using the open-source CaffeNet model. Five convolutional layers are among the eight total learning layers.

Grinblat et al. [19] utilized deep learning techniques to differentiate between various plant types by analysing the physical attributes of leaf veins. They trained CNN models with an increasing number of layers, ranging from two (including one convolutional and one softmax) to six levels (including five convolutional and one softmax). Among these, the CNN model with five layers, incorporating veins at three different scale factors, achieved the highest performance with an average accuracy of 96.9%.



3. Methods And Metrics

Figure 2 displays the four main processes that were used in the study: sampling, picture pre-processing, feature extraction, and classification. Taking pictures and collecting leaf samples was the first step. Following the pre-processing of the images, they proceeded to the feature extraction step, where important information about the leaves was extracted using a mixture of Convolutional Neural Networks (CNN). The last step was to train and categorize the acquired attributes using a variety of machine learning approaches.

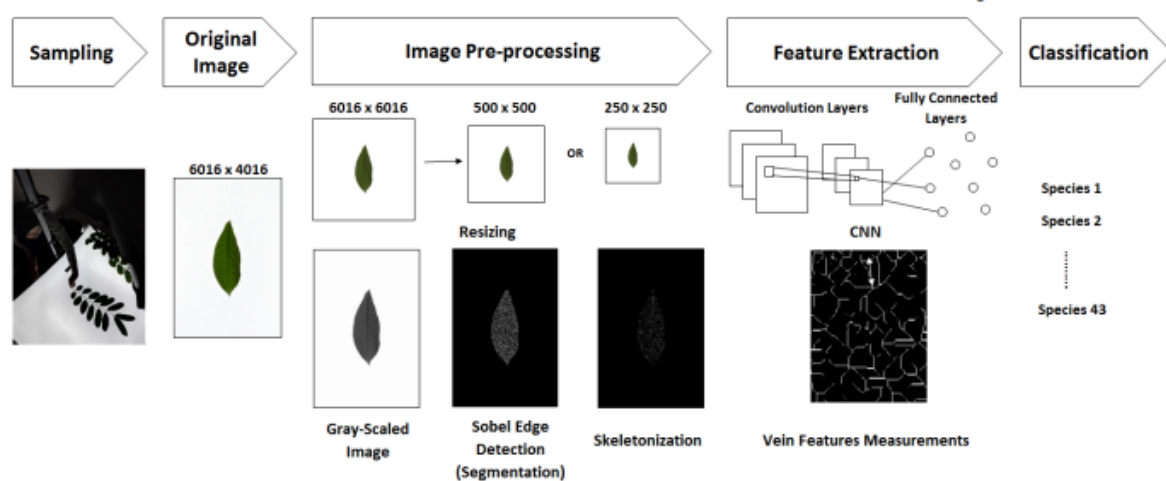


Figure 2: Schematic of the Suggested Model

3.1 Sampling

The images of leaves in this collection were taken from tropical trees that are readily available in Malaysia and may be visited with relative ease at the University of Malaya. Unlike other parts of the tree, such as fruits or flowers, leaves are always accessible, making them an ideal sample for scientific investigation. Thirty samples were collected from each of the forty-three species of tropical plants. Hence, a total of 1,290 images of leaves were collected for this analysis. A Nikon D750 camera was used to take the images of the leaves. The specimens were placed on top of a white-back dropped container, and fluorescent lights were placed underneath the container to provide uniform background illumination for high-quality images. Reduced glazing and shadow on the leaf are the results of lighting coming from below in this arrangement.

3.2 Image Pre-processing

To make raw photographs usable for research, one needs convert them to a processed format like jpeg, jpg, or tiff. Utilizing a Nikon camera, the acquired images were recorded in the NEF format, boasting a pixel resolution of 6016 x 4016. Tagged Image File Format was the end result of the processing performed on the raw photographs in Adobe Photoshop. Also, using



Adobe Photoshop, we were able to reduce the amount of background noise in the images. The pictures were also resized and converted using MATLAB R2016a. Vein morphometric measurements and image reconstruction for CNN were the two separate pre-processing approaches used.

To prepare the images for convolutional neural networks (CNNs), we first scaled the leaf photographs to square meters by square meters, according to the requirements of the CNN models we selected. By include 1000 pixels of padding at the top and bottom, the initial photographs with dimensions 6016 x 4016 were enlarged to a resolution of 6016×6016 . This was carried out prior to reducing the size of the images so that the leaf shape could be maintained. In order to reduce the computer time, the photographs were then resized to a resolution of 250 by 250. After then, the RGB format was used to store the images.

3.3 Feature Extraction

Important features from the leaf images are extracted using feature extraction, which is the backbone of this research. Zernike Moments, Hu's Moment, Histogram of Oriented Gradient (HOG), and more methods are available for feature extraction.

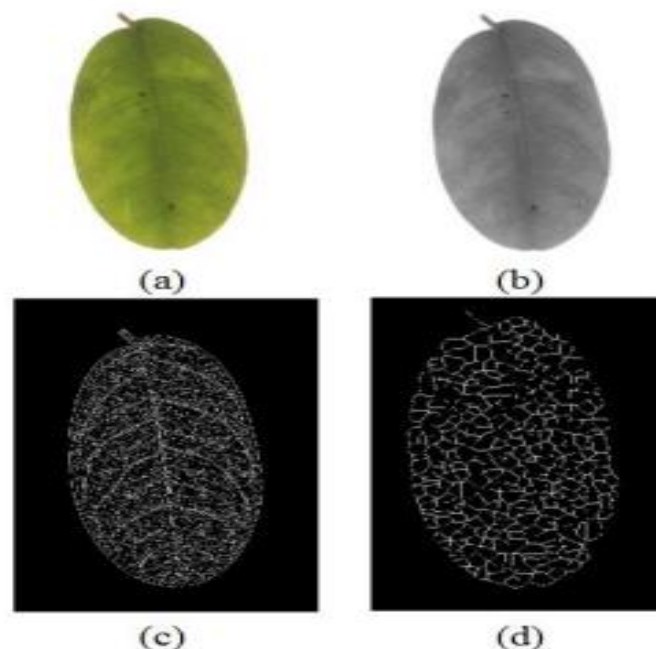


Figure 3: Images in the following formats: (a) RGB (b) Greyscaled (c) Segmented (d) Skeletonized

In this particular research, the characteristics were obtained by the use of a deep learning strategy known as a Convolutional Neural Network. Combining Sobel edge detection with morphological data enabled researchers to test the method's ability to segment leaf venation.



This paper suggests using convolutional neural networks, a deep learning approach, for feature extraction. This research also includes the usage of deep learning algorithms. Convolutional neural networks are basically a multilayer perceptron version due to the fact that they gather feature information from pictures by using many layers of perceptron's. A convolutional neural network (CNN) layer typically consists of three layers, as shown in Figure 4 [18, 19]. In the convolution stage of the first layer, filters and kernels are used to execute convolution operations on feature maps, processing the input image. For the second layer's feature maps, the non-linear activation function known as the Rectified Linear Unit is used. Using a piecewise linear technique, this function would retain its almost linear structure. The result is that this approach will preserve many characteristics that improve the usability and efficacy of gradient-based techniques to linear models [6]. Finally, the data from the third layer is summarized and compressed during a pooling phase in the third layer.

A CNN model named AlexNet, which had already been trained, was used in the research. ImageNet supplied the dataset that AlexNet was trained on, which included one thousand images classified into different types [2, 25]. Figure 4 shows the nine layers that make up AlexNet. Among these layers, you may find a softmax classification feature, three completely connected layers, and five convolution layers. The softmax classification layer is an algorithmic component that makes use of the softmax function to predict probabilities across all categories. Between it and the last fully linked layer, there is an absolute connection. The input picture was changed to a $227 \times 227 \times 3$ resolution before AlexNet was trained. Table I displays the setup of the layers used by AlexNet. There were 4096, 4096, and 1000 neurons in the model's three fully-connected layers, in that order. Next, get the softmax classification layer. AlexNet's revised design is identical to the original. As can be seen in TABLE II, we tweaked a number of the AlexNet model's parameters. Like the AlexNet models, these fine-tuned models had their input dimensions set to $227 \times 227 \times 3$. To fine-tune the first convolutional layer, we changed the stride to [2 2] and the filter size to 7×7 . Additionally, the number of neurons in the three completely linked layers was adjusted to 1290, 1290, and 43 by fine-tuning.

This model's other parameters were left unchanged, in line with the AlexNet model. D-Leaf Model Proposed As seen in Figure 5, a Convolutional Neural Network (CNN) model was created as part of this research to extract visual attributes. This method was used in lieu of fine-tuning the AlexNet model. Figure 5 shows the result of using MATLAB to build a convolutional neural network (CNN) model with six layers. A pooling stage and ReLU activation functions follow each of the model's three convolution layers. In addition, a softmax classification layer and three completely linked layers are present.

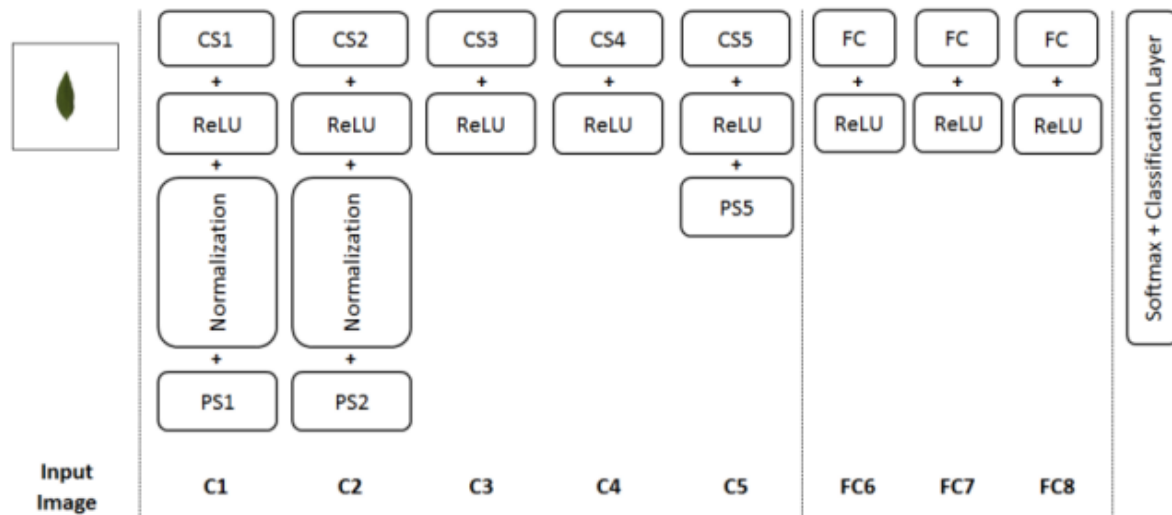


Figure 4: Architecture with different layers

Table 1: Factors of AlexNet Layers

CNN layer	CS1	PS1	CS2	PS2	CS3	CS4	FC6	FC7	FC8
Filter Size	11*11	3*3	5*5	3*3	3*3	3*3	-	-	-
No. of Kernels	96	-	256	-	384	384	4096	4096	1000
Stride size	[22]	[22]	[11]	[22]	[11]	[11]	-	-	-

Table 2: Factors of D-Leaf Layers

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 Kernels	64	-	96	-	256	-	1290	1290	43
Stride size	[44]	[22]	[22]	[22]	[11]	[22]	-	-	-

In Table 2 you can see the D-Leaf model's layer parameters. The model's three tiers were interdependent on one another. Two completely connected layers, each having 620 neurons,



received the output of the third layer. One layer had 1290 neurons. There were 43 neurons in the FC6 layer, which was the third completely connected layer. This number was indicative of the amount of classes that were present in the dataset. A softmax classification layer was the last component that was included into the model. The 100-batch stochastic gradient descent in conjunction with momentum served as the foundation for the training choices that were made for the D-Leaf model.

3.4 Classifiers

The last step in an automated system for identifying plants is the classification process. During this stage, an algorithm that uses artificial intelligence is used in order to identify certain properties that are unique to each species. Support Vector Machines, Artificial Neural Networks, and k-Nearest Neighbour are only few of the methods that are used extensively in the area of machine learning. There were five different classification methods that were used in the study. Support vector machines, artificial neural networks, k-nearest neighbours, neural networks, and convolutional neural networks were used. Support Vector Machines are widely recognised as effective classification methods. They are supervised machine learning techniques that are resilient in high-dimensional areas and handle input points that are not linearly spaced. Applying linear Support Vector Machines to feature-mapped data has the potential to improve classification performance, streamline execution, and reduce storage requirements [14]. This study used a Linear Support Vector Machine using the 'One versus all' technique to handle the dataset's many classifications.

In several research, an Artificial Neural Network is the most popular way [3, 13, 15, and 17]. The behavioural and anatomical features of neurons in the human brain inspired the development of ANN. One way to tweak an ANN model's classification performance is to play around with the model's neuron and hidden layer counts. The study used a feed-forward neural network that has a single 80-neuron hidden layer. Achieving a low gradient served as the stopping condition in the training technique, which used the default Scaled Conjugate Gradient algorithm.

Classification methods such as k-NN take the majority vote from nearby samples into account when deciding which class to place a given sample in [13]. Using the city block distance metric, the number of neighbours is fixed to 1 in this research. To determine the category of an unlabelled sample, a Bayesian classifier use probability [21]. NB means that there is no correlation between any of the sample variables [24]. The conditional independence of Bayes' theorem lowers classification accuracy than expected.



Results and discussion

4.1 AlexNet and Fine-tuned AlexNet

The research used an 80:20 split between the data sets used for training and testing. Ten rounds of feature extraction and classification were applied to each model to achieve optimal performance. The pre-trained AlexNet model was used to retrieve characteristics from the FC7 layer. These characteristics were then used by four distinct classifiers. Five distinct classifiers were trained and assessed after features were extracted using an improved version of the AlexNet model. With a score of 93.26 percent, the ANN model was the most accurate classifier. Nevertheless, the SVM classifier that made use of AlexNet features only managed to achieve an accuracy of 79.40%.

Among all the models that were trained using these characteristics, the one that performed the best was an artificial neural network classifier that was trained with fine-tuned AlexNet features, with an accuracy rate of 95.54%. Although a few classifiers managed to achieve 89% accuracy, others were utterly unsuccessful. Prior to using the refined AlexNet model for feature extraction, it is necessary to retrain it with further training time. However, retraining is unnecessary to use the AlexNet model for feature extraction. It was the fine-tuned AlexNet models that outperformed the conventional ones.

Table 3: Discoveries Made by AlexNet and Fine-Tuned AlexNet System

Classifier	AlexNet Accuracy	Fine-tuned AlexNet Accuracy
ANN	93.31	95.57
CNN	-	88.32
KNN	85.63	87.35
SVM	79.40	87.83
NB	83.34	87.34

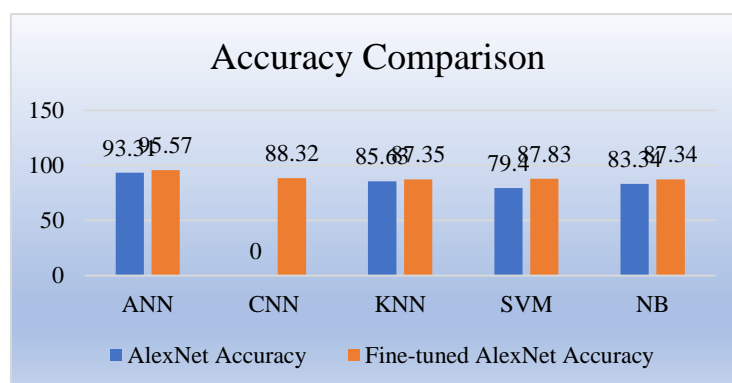


Figure 5: accuracy comparison for AlexNet with different classifiers



4.2 D-Leaf

A 250 x 250 picture input size was used in the research carried out on the proposed D-Leaf model. After that, we used a CNN to compare the outcomes. Next, we used SVM, ANN, k-NN, NB, and CNN classifiers to organise D-Leaf's FC5 layer characteristics. See Table 4 for classifier names. The study used a CNN AlexNet that had already been trained, as well as an AlexNet model that had been fine-tuned, to extract features for classification tasks. The D-Leaf architecture, a new kind of convolutional neural network (CNN), is more complicated than the AlexNet model. Among the most impressive results was the 95.54% testing accuracy attained by the fine-tuned AlexNet model, which was closely followed by the 94.88% D-Leaf model.

Table 4 Comparison of Classifier Performance Using D-Leaf Features

Classifier	D-Leaf Accuracy
ANN	94.89
CNN	79.56
KNN	82.47
SVM	82.78
NB	81.96

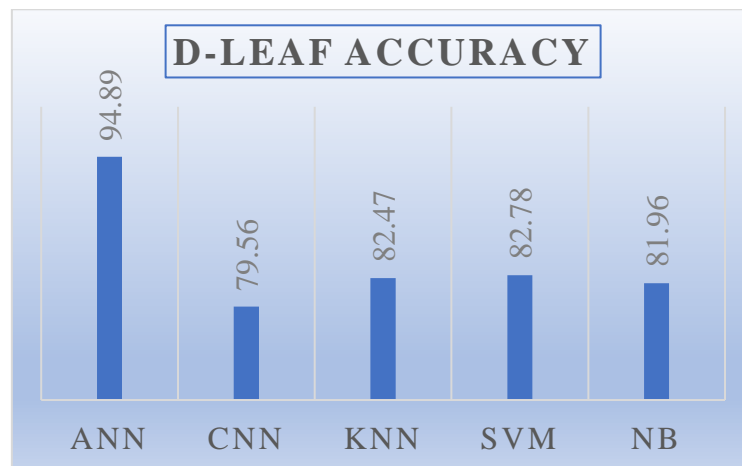


Figure 6: D_Leaf accuracy with classifiers

Unfortunately, AlexNet couldn't handle all of the plant species because of its final fully-connected layer layout, which had 1000 neurons (43). The use of AlexNet as a classifier necessitated the execution of many refining methods. As a result, D-Leaf came to light as a powerful automated method for plant species categorization, with results comparable to those of the refined AlexNet model. The research failed to undertake statistical analysis to determine the relevance of classifier performance variances.



5. Conclusion

Based on this study, we may infer that CNN outperforms traditional morphological techniques in extracting features from plant species. The traditional approaches need more pre-processing tasks in comparison to the CNN. This study establishes that CNN proves more advantageous for feature extraction than classification. Among all the classifiers examined, for optimal performance, use an ANN classifier in conjunction with a CNN feature extractor. The greatest testing accuracy that this research was able to achieve was 94.88%. A combination of the D-Leaf feature extraction technique and the ANN classification algorithm was used in order to achieve this goal.

6. References

- [1] Willis, K.J. (ed.) 2017. "State of the World's Plants 2017. Report. Royal Botanic Gardens, Kew".
- [2] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "ImageNet classification with deep convolutional neural networks". In Advances in neural information processing systems (pp. 1097-1105).
- [3] Wang, Z. (2015). "The applications of deep learning on traffic identification". BlackHat USA.
- [4] Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. H. (2016). "Deep learning for identifying metastatic breast cancer". arXiv preprint arXiv:1606.05718.
- [5] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N. & Kingsbury, B. (2012). "Deep neural networks for acoustic modelling in speech recognition: The shared views of four research groups". IEEE Signal Processing Magazine, 29(6), 82-97.
- [6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep learning. MIT Press".
- [7] Malladi, R., & Sethian, J. A. (1996). "A unified approach to noise removal, image enhancement, and shape recovery". IEEE Transactions on Image Processing, 5(11), 1554-1568.
- [8] Metre, V., & Ghorpade, J. (2013). "An overview of the research on texture based plant leaf classification". arXiv preprint arXiv:1306.4345.
- [9] Arun, C., Emmanuel, W. S., & Durairaj, D. C. (2013). "Texture feature extraction for identification of medicinal plants and comparison of different classifiers". International Journal of Computer Applications, 62(12).
- [10] Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010, December). "The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms". In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 135-144). Springer Berlin Heidelberg.



- [11] Anami, B. S., Suvarna, S. N., & Govardhan, A. (2010). "A combined color, texture and edge features based approach for identification and classification of Indian medicinal plants". *International Journal of Computer Applications*, 6(12), 45-51.
- [12] Larese, M., Craviotto, R., Arango, M., Gallo, C., & Granitto, P. (2012). "Legume identification by leaf vein images classification". *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 447-454.
- [13] Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I. (2013). "Neural network application on foliage plant identification". *arXiv preprint arXiv:1311.5829*.
- [14] Sulc, M., & Matas, J. (2014). "Texture-Based Leaf Identification". Paper presented at the *European Conference on Computer Vision*.
- [15] Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015, September). "Deep-plant: Plant identification with convolutional neural networks". In *Image Processing (ICIP), 2015 IEEE International Conference on* 2015 Sep 27 (pp. 452-456). IEEE.
- [16] Lee, S. H., Chang, Y. L., Chan, C. S., & Remagnino, P. (2016). "Plant identification system based on a convolutional neural network for the lifeclef 2016 plant classification task". In *Working notes of CLEF 2016 conference*.
- [17] Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). "How deep learning extracts and learns leaf features for plant classification". *Pattern Recognition*, 71, 1-13.
- [18] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification" *Computational Intelligence and Neuroscience*, 2016.
- [19] Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). "Deep learning for plant identification using vein morphological patterns". *Computers and Electronics in Agriculture*, 127, 418-424.
- [20] Larese, M. G., Namías, R., Craviotto, R. M., Arango, M. R., Gallo, C., & Granitto, P. M. (2014). "Automatic classification of legumes using leaf vein image features". *Pattern Recognition*, 47(1), 158-168.
- [21] Reyes, A. K., Caicedo, J. C., & Camargo, J. E. (2015). "Fine-tuning Deep Convolutional Networks for Plant Recognition". In *CLEF (Working Notes)*.
- [22] Gupta, S., & Mazumdar, S. G. (2013). "Sobel edge detection algorithm". *International journal of computer science and management Research*, 2(2), 1578-1583.
- [23] Vincent, O. R., & Folorunso, O. (2009, June). "A descriptive algorithm for sobel image edge detection". In *Proceedings of Informing Science & IT Education Conference (InSITE) (Vol. 40, pp. 97-107)*.
- [24] Yadav, A. R., Dewal, M., Anand, R., & Gupta, S. Classification of hardwood species.
- [25] Prasvita, D. S., & Herdiyeni, Y. (2013). "MedLeaf: Mobile Application for Medicinal Plant Identification Based on Leaf Image". *International Journal on Advanced Science, Engineering and Information Technology*, 3(2), 103-106.