IDENTIFICATION OF MEDICINAL PLANT'S AND THEIR USAGE BY USING DEEP LEARNING

C.Amuthalingeswaran
Department of CSE
Thiagarajar College Of Engineering
Madurai, India
amuthalingeswaranbose@gmail.com

Mr. M.Sivakumar
Department of CSE
Thiagarajar College Of Engineering
Madurai, India
mskcse@tce.edu

Dr. P.Renuga
Department of EEE
Thiagarajar College Of Engineering
Madurai, India
preee@tce.edu

S.Alexpandi
Department of CSE
Thiagarajar College Of Engineering
Madurai, India
alexpandi165@gmail.com

J.Elamathi
Department of CSE
Thiagarajar College Of Engineering
Madurai, India
elamathijayaraj97@gmail.com

S.Santhana Hari
Department of CSE
Thiagarajar College Of Engineering
Madurai, India
santhanahari95@gmail.com

Abstract - **Nowadays, most of the peoples are not aware of their rural and urban medicinal plants and their uses. If we want to find the plant details for their medicinal values, we have to be exposed to it previously. However, some people will find it to be a tedious process since they are not aware of these plants before. Moreover, at some times, we have to depend on others for the identification of medicinal plants. (i.e.) Botanist, Peoples who aware of it. So in order to avoid all these kinds of situations and with the availability of modern computing devices and technology we had built a model (Deep Neural Networks) for the identification of medicinal plants. To train the model we used around 8,000 images belonging to four different classes. Finally, we arrived with good accuracy of 85% when testing with images taken from the open field land areas.**

Keywords – Medicinal Plants, Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Network, Fully Connected, Neurons.

I. INTRODUCTION

Today, AI (Artificial Intelligence) plays a major role in Computer Vision, Robotics, Digital Transformation of marketing, Medical industries and banking sectors. Artificial Intelligence was mainly devolved for making machines to think and act like a human.

The term ML (Machine Learning) is a sub-set of Artificial Intelligence(AI), and also scientific study of the algorithm and a statistical model that perform a specific test without explicitly programmed, relying on the inference and patterns. The types of machine learning techniques are Supervised Learning, Unsupervised Learning, Reinforcement Learning. The Supervised Learning Algorithms consists of data, which are labeled. Unsupervised Learning Algorithm takes a set of data; it contains only unlabeled inputs and finds the structure in the data. Learning can be a Supervised Learning, Semi-Supervised Learning, and Unsupervised Learning. The Amount of data being increased then the Machine Learning algorithms are insufficient in terms of its performance but Deep Learning provides better performance even on a larger dataset. Deep Learning is also called Deep Structured Learning. Deep Learning used in Computer Vision and Image Identification and etc., Deep Learning use Artificial Neurons that similar to our human Neurons. Deep Neural Networks, Deep Belief Networks, and Recurrent Neural Network, these are used in Speech Recognition, Natural Language Processing, Machine Translations, Audio Recognitions, bioinformatics, and Drug design, Medical Image Identifications.

ANN (Artificial Neural Network) is inspired by biological neural network based Matrix Mathematical model, ANN was created by the collection of the artificial neural network in which the computer is fed with the number of inputs, which machine learns and perform certain based on the previously feed information and finally produces an output from the learned pattern.

II. EXISTING WORK

The Transfer Learning for Deep Learning based Plants Classification model proposed by paper [1]. The results of this paper show the effect of four different Transfer Learning Models for DNN based plant classification deals with four different public datasets. Finally, their experimental study shows the Transfer Learning provide a self-estimating and Analyzing plant classification model. They use some general schema like End-to-End CNN, Fine Tuning, Cross Dataset Fine Tuning, Deep Featured Fine-tuning, RNN-CNN Classification.

The Paper [2] investigates into the main factor affecting the design and the effectiveness of deep neural network used in plant pathology. They discussed in- depth manner and also advantage and disadvantage. They also used a transfer learning methodology from the previously trained CNN Model. List of Factors affected CNN plant disease recognition. (i.e.) Limited Annotated Datasets, Covariate Shift, Image Background and Image Captured Conditions. Symptoms Segmentation, Symptom Variation, Simultaneous

Disorders, Disorder with Similar Symptoms, these are intensively connected to the problem.

In paper [3] the author used VGG Architecture to train the model. Here, the Main objective is plant disease detection and diagnosis. The use around 87,848 images containing 25 different plants in a set and 58 individual class includes diseased and healthy to train the model, the model to predict the test samples and provide better accuracy. The accuracy of the model 99.57% is the best accuracy of the above-mentioned Deep Learning Models for Plant Disease Detection and Diagnosis.

In Paper [4] collection of plant leaf images is given as input and used CNN to find the pattern of every individual plant details. Here, CNN was mainly used for better feature representation and for efficient findings of Leaves species used in DN (Deconvolutional Network). It provides better Identification of plant leaves and its species.

In paper [5] the author investigates with the application of deep learning and the application of transfer learning problems. CNN is a tool used in plant pathology identification problem. In [6] the team proposed CNN model for a Plant vein identification model and it uses three different legumes datasets: White Beans, Red beans, Soya Beans. And, they proved that deep learning approach is providing better accuracy and performance.

They [7] have applied three transfer learning model to identify the plant species identification. They used LIFECLEF 2015 to evaluate the network. Here the team used AlexNet, GoogLeNet, and VGGNet for their proposal. In [8] the author talks more about fine-tuning and evaluated state of the art the Deep Convolutional architecture for plant classification and identification. The author proposed the model by evaluating more transfer learning architectures like VGG 16, Inception V4, ResNet Layers, and DenseNet. Dense layer requires some reason to perform state of the art. This model achieves the accuracy of 99.57%.

In the work done by [9], semantic segmentation by using deep learning paradigm is done. The surveyed method divided into ten classes. Their survey shows accuracy improvement, performance, and speed. The main objective of this paper is to analyze performance and reach a goal of accuracy.

The purpose of the paper [10] is to identify the small objected plant leaves by using CNN (Complex Background). The proposed system applied the inception V2 with batch-normalization, which improves the accuracy of RCNN (Region Convolutional Neural Network). The High-quality images are divided into a hundred sub-samples for testing and the remaining images are given back to the final output. This proposed method is much faster than conventional RCNN.

The author of the paper [11] focused on the semantic segmentation by using deep learning techniques. They proved brief details about deep learning topics. And provides the necessary and required knowledge about deep learning for the upcoming task. Paper [12] deals with a real-time scenario like organic farming using deep learning. The test scenario is done field areas and not under laboratory. They use image classification and segmentation method.

Here, the hyperspectral data with plant disease detection is discussed with the use of NN techniques [13]. (pre-symptomatic, symptomatic, asymptomatic disease from a single plant).

From the work of paper [14] deeper learning and machine-learning tool for identifying plant stress phenotyping is done. They used the End-to-End DL approach to provide image-based classification and segmentation.

In paper [15], to train the model they used different colors of the leaf, as an input. This paper proposed a new methodology called TCCNN. TCCNN (Three Channel Convolutional Neural Network). TCCNN is to find a color combination of the leaf and to find the leaf disease from the inferred pattern. These mainly used for vegetable leaf disease detection problem and finally, it provides better performance and result.

III. DNN (DEEP NEURAL NETWORK)

DNN or DBN (Deep Belief network) is used in Computer Vision; Speech Reorganization, Natural Language Translation, and many other applications.

DNN has an Input, Hidden and Output Layers. The DNN [1] deals with multiple neural networks. Fig.1 shows how the layers and nodes in the Neural Network are connected and share informations.

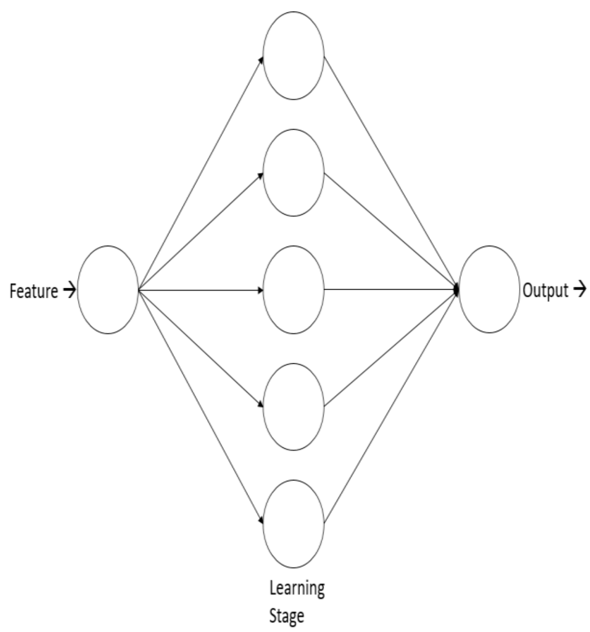

Fig.1. Connectivity of NN

A. Neural Network (NN)

NN is a network or connection of neurons. The NN [13] have an Input Layer, Hidden Layer, and Output Layer. Every input Layer will have some values, that values are multiplied with the weights. The results are sent to other hidden neurons. Applications of NN - Regression Analysis, Time Series prediction, Classification, Data Processing, Decision Making Application, Blind Decision Separations.

B. CNN (Convolutional Neural Network)

CNN [2] is a class of Deep Neural Network. It mainly used in CV (Computer Vision) to find the pattern given in the subjected object. The main objective of CNN [3] is to find the pattern of given input samples and to predict it. It provides better accuracy and improved performance.

C. Layers In CNN

- Convolutional Layer is an important and necessary part of Convolutional Neural Network. ConvNet has a set of self-learnable filters.
- ReLU stands for Rectified Linear Unit. ReLU is an Activation Function
- Pooling Layer is used to reduce the dimension of the given input matrixs. There are three kinds, Min-Pooling, Max-Pooling, Avg-pooling
- Fully Connected Layer is connected by a series of nodes (neurons)

IV. Proposed Methodology

This proposed work is mainly to focus on the detection of plants based upon their medicinal values. Our model MNN (Medicinal Neural Networks) is trained with the dataset, collected by us, manually. This work is not done by using the concept of transfer learning; rather we had trained our model from scratch. During the training [3,16] phase, the particular supplied model will learn some kind of different properties that are presented in the input image and at some cases while testing, can result in misclassification. Therefore, to avoid misclassification we had trained our model from scratch.

Layer, used for our model

- 0^{th} Layer - CNN of 16 filters (2*2) with Stride value = 1
- 1^{st} Layer - CNN of 16 filters (2*2) with Stride value = 1
- 2^{nd} Layer - MaxPooling of (2*2)
- 3^{rd} Layer - CNN of 32 filters (2*2) with Stride value = 1
- 4^{th} Layer - MaxPooling of (2*2)
- 5^{th} Layer - CNN of 32 filters (2*2) with Stride value = 1
- 6^{th} Layer - MaxPooling of (2*2)
- 7^{th} Layer – CNN of 64 filters (2*2) with Stride value = 1
- 8^{th} Layer - MaxPooling of (2*2)
- Dense Node (64)
- SoftMax (4)

A. Dropout

In many works of CNN, a feature Dropout is used. This can result in our model from the problem of overfitting. By dynamically removing certain connections existing between the nodes randomly.

B. Image Augmentation

Augmentation is done with the help of Python code. This can make the images to transform according to the fixed category of size valued parameters. So that the model will learn the images present in different angles and perception.

Fig. 2 represents the architecture of our work. SoftMax function is used at the very last layer; this will split the inputs into a vector array of probability values of (n). In addition, the position that is having a high probability shows that the image belongs to that corresponding class.

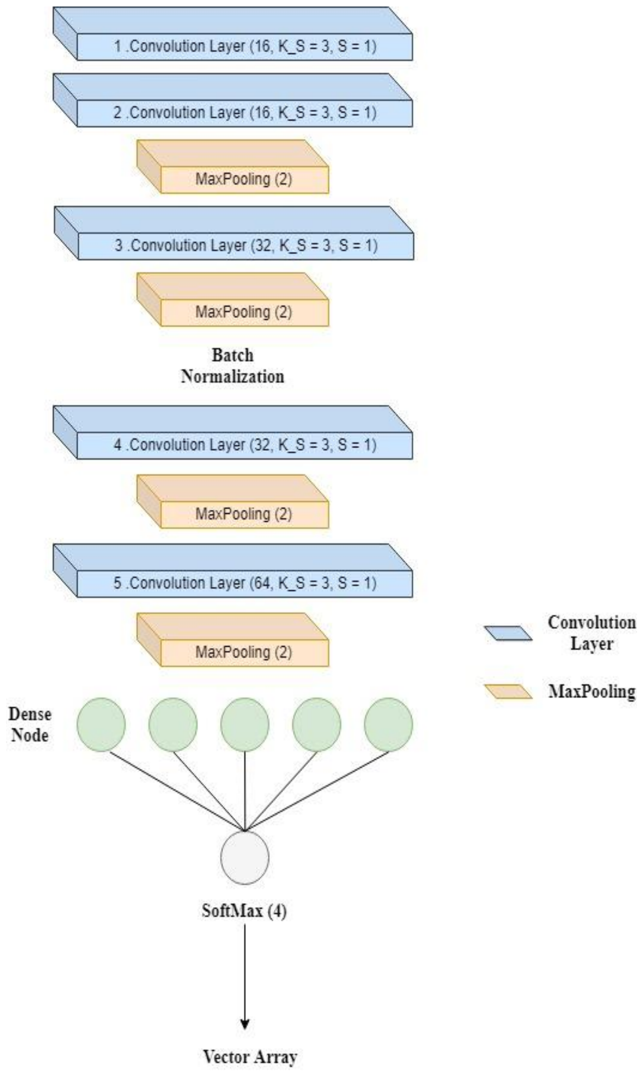

Fig. 2. The architecture of our model (MNN)

V. Experiments and Evaluations

A. Dataset

Dataset consists of four (4) different varieties of medicinal plants. Moreover, all the images that are collected are segmented by hiding the background information with a white screen. Table I cited below shows the classes present in the dataset and the number of samples along with their traditional Tamil names and botanical names. The image samples in the dataset are illustrated in Fig. 3. ((i) Catharanthus Roseus (ii) Tephrosia Purpurea) and Fig. 4. ((i) Phyllanthus amarus (ii) Abutilon Indicum)

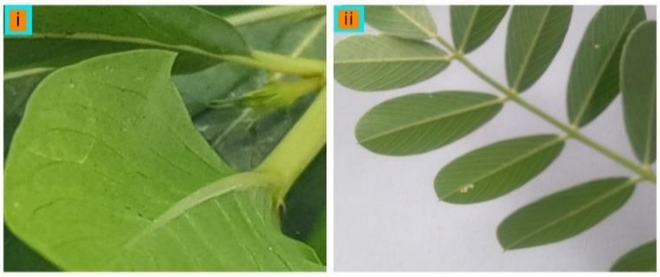

Fig. 3. Image samples in the dataset

Fig. 4. Image samples in the dataset

TABLE I
Classes Present in the Dataset

Class No	Tamil Name	Equivalent Botanical Name	Number of samples per class
0	Nithiya_Kalyani	Catharanthus Roseus	2089
1	Keela Nelli	Phyllanthus amarus	2525
2	Kolunchi	Tephrosia Purpurea	1725
3	Thuthi	Abutilon Indicum	1920

In our model, we had used the Confusion Matrix for evaluating the metrics. Table II shows the evaluation metrics results.

TABLE II
Overall Confusion Matrix for all Four (4) Class

No. of Samples (8259)	**Predicted No**	**Predicted Yes**	**Values**
Actual No	354	570	0.3831
Actual Yes	656	6679	0.9105
Values	0.3504	0.9213	0.8515

VI. Results and Discussion

We had made our training by considering the images taken under different lighting conditions. This property is adopted in order to produce a very good classification accuracy when tested in real time scenario.

Fig. 5. Represents the analysis of accuracy values produced by the two different models. MNN of 85% and Mobile Net of 72%.

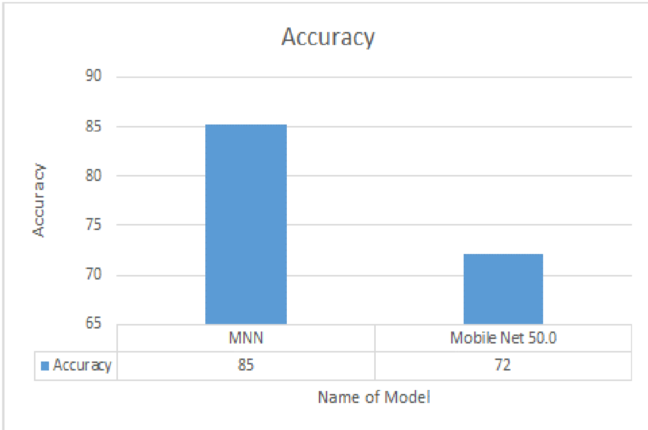

Fig. 5. Comparative analysis of Medicinal Neural Network and Mobilenet 50.0

VII. Conclusion And Future Work

In this work, we had developed a model for the efficient classification of medicinal plants, and this work is implemented by having four (4) number of plant disease classes. Finally, this whole work is implemented from scratch and produces an accuracy percentage of 85.15%. The future work is to increase the size of the dataset by increasing the samples as well as by adding new kinds of medicinal plants.

References

[1] Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H. and Tekinerdogan, B., 2019. Analysis of transfer learning for deep neural network based plant classification models. Computers and Electronics in Agriculture, 158, pp.20-29. https://doi.org/10.1016/j.compag.2019.01.041

[2] Barbedo, J.G., 2018. Factors influencing the use of deep learning for plant disease recognition. Biosystems engineering, 172, pp.84-91. https://doi.org/10.1016/j.biosystemseng.2018.05.013

[3] Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, pp.311-318. https://doi.org/10.1016/j.compag.2018.01.009

[4] Lee, S.H., Chan, C.S., Mayo, S.J. and Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 71, pp.1-13. http://dx.doi.org/10.1016/j.patcog.2017.05.015

[5] Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture, 153, pp.46-53. https://doi.org/10.1016/j.compag.2018.08.013

[6] Grinblat, G.L., Uzal, L.C., Larese, M.G. and Granitto, P.M., 2016. Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture, 127, pp.418-424. http://dx.doi.org/10.1016/j.compag.2016.07.003

[7] Ghazi, M.M., Yanikoglu, B. and Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing, 235, pp.228-235. http://dx.doi.org/10.1016/j.neucom.2017.01.018

[8] Too, E.C., Yujian, L., Njuki, S. and Yingchun, L., 2018. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2018.03.032

[9] Lateef, F. and Ruichek, Y., 2019. Survey on Semantic Segmentation using Deep Learning Techniques. Neurocomputing. https://doi.org/10.1016/j.neucom.2019.02.003

[10] Zhu, X., Zhu, M. and Ren, H., 2018. Method of plant leaf recognition based on improved deep convolutional neural network. Cognitive Systems Research, 52, pp.223-233. https://doi.org/10.1016/j.cogsys.2018.06.008

[11] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P. and Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing, 70, pp.41-65. https://doi.org/10.1016/j.asoc.2018.05.018

[12] Knoll, F.J., Czymmek, V., Poczihoski, S., Holtorf, T. and Hussmann, S., 2018. Improving efficiency of organic farming by using a deep learning classification approach. Computers and electronics in agriculture, 153, pp.347-356. https://doi.org/10.1016/j.compag.2018.08.032

[13] Golhani, K., Balasundram, S.K., Vadamalai, G. and Pradhan, B., 2018. A review of neural networks in plant disease detection using hyperspectral data. Information Processing in Agriculture. https://doi.org/10.1016/j.inpa.2018.05.002

[14] Singh, A.K., Ganapathysubramanian, B., Sarkar, S. and Singh, A., 2018. Deep learning for plant stress phenotyping: trends and future perspectives. Trends in plant science. https://doi.org/10.1016/j.tplants.2018.07.004

[15] Zhang, S., Huang, W. and Zhang, C., 2019. Three-channel convolutional neural networks for vegetable leaf disease recognition. Cognitive Systems Research, 53, pp.31-41. https://doi.org/10.1016/j.cogsys.2018.04.006

[16] Johannes A, Picon A, Alvarez-Gila A, Echazarra J, Rodriguez – Vaamonde S, Navaias A.D, Ortiz-Barredo A, Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case, Computers and Electronics in Agriculture. 138, 200–209 (2017) https://doi.org/10.1016/j.compag.2017.01.013