JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Learning Approach for Herbal Plant Detection

Sumit Shinde[1], C.G.Patil[2], Ashwini Kalgunde[3],S.S.Bhosale[4], Nikhil Ganvir[5],S. N. Kamble[6]

"Dept. of E&TC Department of E&TC,SKNCOE,SPPU,Pune

[1]Sumitshinde3112@gmail.com
[2]chandrashekhar.patil_skncoe@sinhgad.edu
[3]shwinikalgunde27@gmail.com
[4]sonali.bhosale_skncoe@sinhgad.edu
[5] nikhilganvir@gmail.com
[6]sarika.kamble_skncoe@sinhgad.edu

Abstract- **These days, a large portion of the people groups don't know about their country and metropolitan restorative plants and their employments. Assuming we need to discover the plant subtleties for their restorative qualities, we must be presented to it beforehand. Nonetheless, certain individuals will see it to be a drawn-out measure since they don't know about these plants previously. Also, at certain occasions, we need to rely upon others for the distinguishing proof of restorative plants. (i.e.) an expert in or student of the scientific study of plants. So to stay away from this load of sorts of circumstances and with the accessibility of current figuring gadgets and innovation we will create a module (Profound Neural Organizations) for the distinguishing proof of restorative plants. To prepare the model we utilized around 1,834 pictures having a place with four various classes. We might get showed up with great response when testing with pictures taken from rural and urban areas. In this system, the herbal plant leaf recognition system is implemented using CNN, Vgg16 and Vgg19 algorithms. The training of the proposed system is performed using google colab cloud platform. The proposed system gives the promising results.**

Keywords- Herbal Plants, Classification, Image recognition, Convolutional Neural Network.

I.INTRODUCTION

India is rich with herbal plants that are not only useful in cooking but also is beneficial in medical.
Leaf is essential part of the plant which is used for recognition of plant
The manual recognition is difficult to the common peoples hence there is a need of automatic techniques which can recognize the Herbal or medical plant from leaf images.
The CNN is the best technique to the classification problem hence it is used along with transfer learning algorithm will be the strategy for development of this system.

II. LITERATURE REVIEW

Raisa Akter et al [1] proposed an automated system for the medicinal plant classification, which will help people identify useful plant species quickly. A new dataset of 10 medicinal plants of Bangladesh is introduced, collected from different regions across the country, and some state-of-the images collected from different sources. After that, a three-layer convolutional neural network is employed to extract the high-level features for the classification trained with the data augmentation technique. The training process was done on 34123 images, and the experimental result on another 3570 images proved that this method is quite feasible and effective, which gave by a 71.3% accuracy rate.
Izwan Asraf Md Zin et al [2] investigates the application of deep convolutional neural network (CNN) for herbal plant recognition through leaf identification. Traditional plant identification is often time-consuming due to varieties as well as similarities possessed within the plant species. This study shows that a deep CNN model can be created and enhanced using multiple parameters to boost recognition accuracy performance. This study also shows the significant effects of the multi-layer model on small sample sizes to achieve reasonable performance. Furthermore, data augmentation provides more significant benefits on the overall performance. Simple augmentations such as resize, flip and rotate will increase accuracy significantly by creating invariance and preventing the model from learning irrelevant features. A new dataset of the leaves of various herbal plants found in Malaysia has been constructed and the experimental results achieved 99% accuracy.

K. Priya at al. [3] in their study 30 Indonesian species of medicinal plants were used and there are 48 digital pictures of growing species.We used a questionnaire focused on heuristic assessment to assess user feedback for the program. The findings of the assessment indicate that MedPlant is effective in distinguishing medicinal plants. MedPlant can help to classify medicinal plants, explore new plant varieties and plant taxonomy in the botanic garden or in the maintenance of natural parks. This also lets individuals, organizations and societies discover their capacity to maximize the value of health plants without usage and without growth. The results would improve MedPlant's wealth, capital and economic richness.

Anh H. Vo et al. [4] deployed a computer vision aided herbal plan identification system to meet the demand of recognizing and identifying herbal plants rapidly. In this paper, the first herbal plant image dataset collected by mobile phone in natural scenes is presented, which contains 10,000 images of 10 herbal plant species in Vietnam. A VGG16-based deep learning model consisting of 5 residual building blocks is used to extract features from the images. A comparative evaluation of seven classification methods using the same deep convolutional feature extraction method is presented. Experiments on our collected dataset demonstrate that deep learning features worked well with Light GBM classification method for herbal plant recognition in the natural environment with a recognition rate of 93.6%.

C. Amuthalingeswaran rt al. [5] 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%.

Manojkumar P., Surya C et al. [6] collected 20 random Ayurvedic front and back side leaves of 40 different species. The Weka tool is used for identification of medicinal plants using machine learning algorithms. Color and texture features of leaves are extracted from color and binary images. Support Vector Machine (SVM) and Multilayer perceptron (MLP) classifiers are used to identify the leaves based on following features Geometric, centroid-radii (CR) distances, colour features, texture features, HU invariant moments and Zernike moments. MLP (94.5%) out formed than Support Vector Machine (SVM)

Herdiyeni and Wahyuni [7] used a fusion of fuzzy local binary pattern and fuzzy colour histogram and a probabilistic neural network (PNN) classifier on a dataset of 2448 leaf images (270 *240 pixels) obtained from medicinal plants from the Indonesian forests to achieve a classification accuracy of 74.5%.

III. PROPOSED METHODOLOGY-

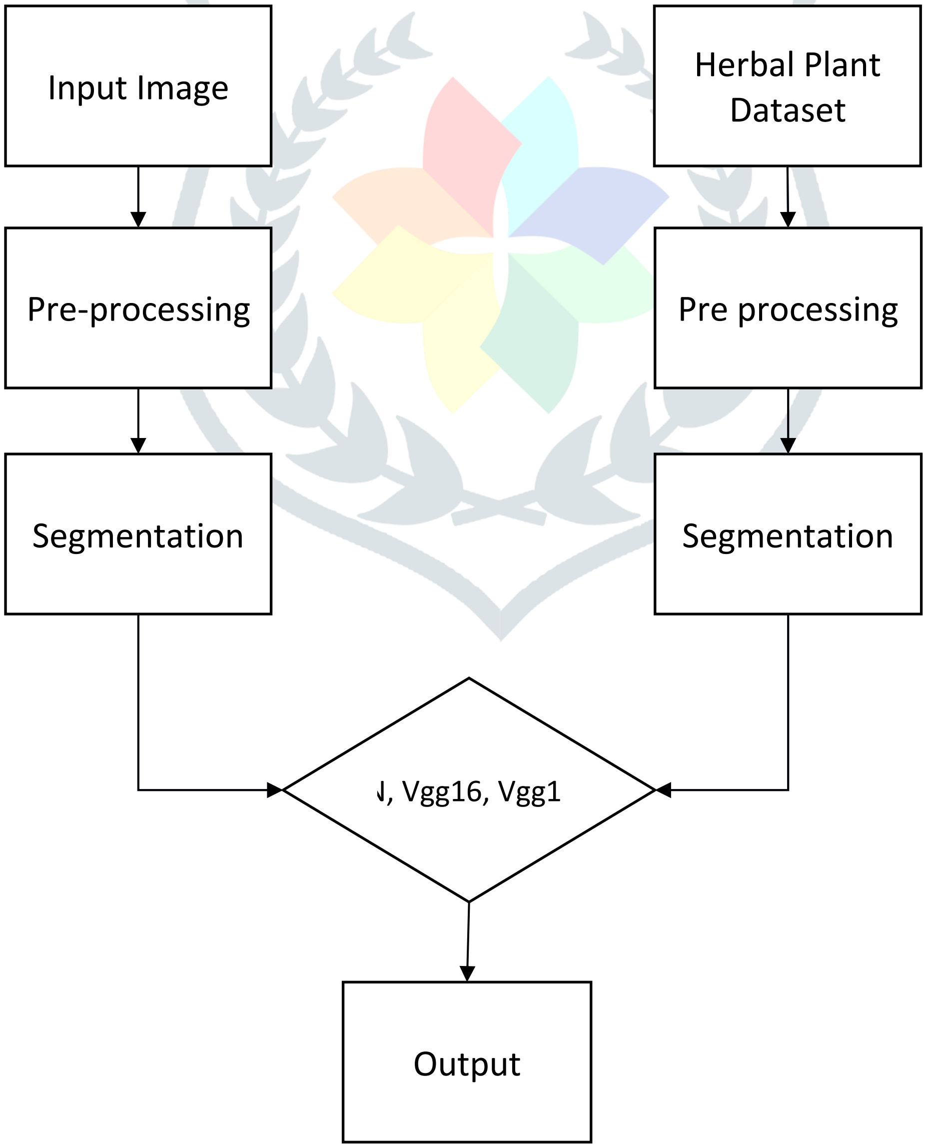


Figure 1. Block diagram Herbal Plant Detection

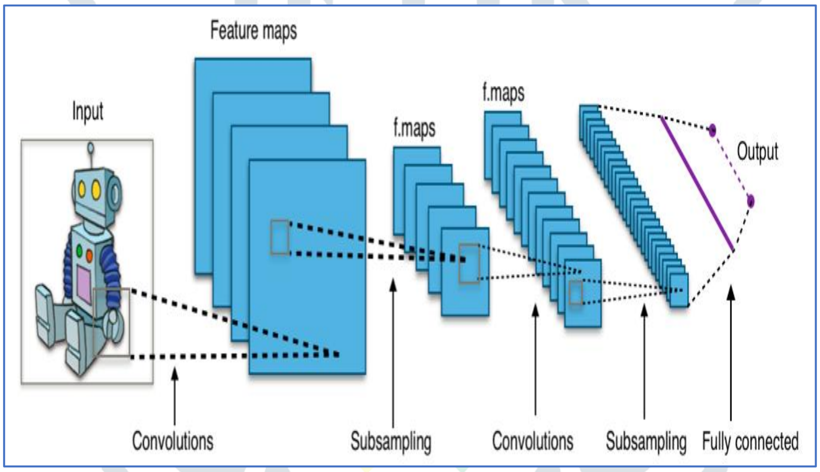
- This system capture the plant leaf image from camera or from the dataset.
- It preprocess the image including median filtering, segmentation of leaf
- The data is spitting into training (80%) and testing (20%)
- The training is performed using three algorithms
    - CNN
    - Vgg16:
        - VGG16 is a convolutional neural network architecture with 16 layers.
    - Vgg19:
        - There are 16 convolutional layers and three Fully Connected layers, which sum up to 19 weight layers.
- The performance is evaluated using accuracy parameter.

IV. IMPLEMENTATION

A. Model Description (CNN, VGG16, VGG19):

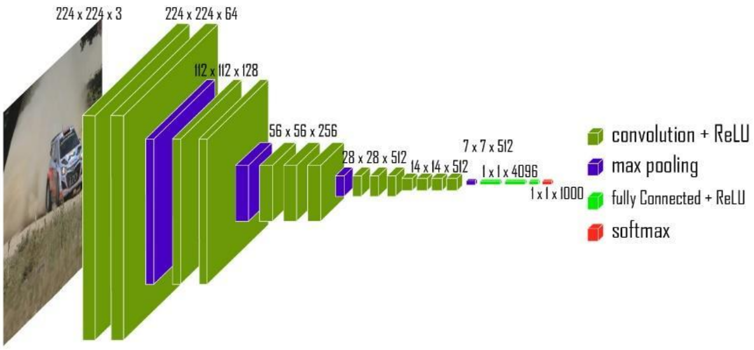
CNN:

CNNs are a category of Neural Networks that have proven very effective in areas such as image recognition and classification. CNNs are a type of feed-forward neural network made up of many layers. CNNs consist of filters or kernels or neurons that have learnable weights or parameters and biases. Each filter takes some inputs, performs convolution, and optionally follows it with a non-linearity



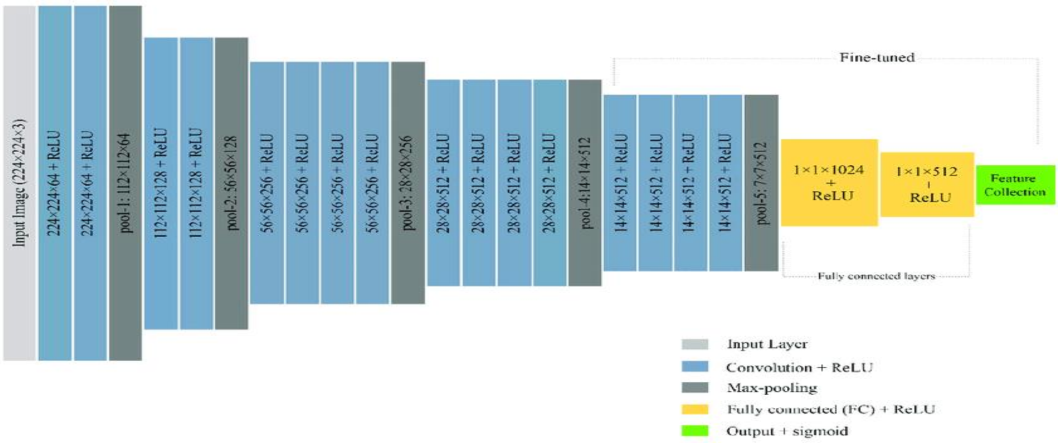
Vgg16:

The architecture of vgg16 model is as shown in Fig.3.8. 13 convolutional layers and 2 Fully connected layers and 1 SoftMax classifier VGG-16 - Karen Simonyan and Andrew Zisserman introduced VGG-16 architecture in 2014 in their paper Very Deep Convolutional Network for Large Scale Image Recognition. Karen and Andrew created a 16-layer network comprised of convolutional and fully connected layers



Vgg19:

VGG19 trained on over a million images from the ImageNet database. The depth of the configuration is 19 weight layers and may classify images into 1000 object classes. The number of parameters in VGG19 is 144 million. There are 16 convolutional layers and three Fully Connected layers, which sum up to 19 weight layers



I. Evaluating Models:

In general, a confusion matrix is an effective benchmark for analysing how well a classifier can recognize records of different classes . The confusion matrix is developed on the basis on the following terms:

1) True positives (TP): positive records that are correctly labelled by the classifier.
2) True negatives (TN): negative records that are correctly labelled by the classifier.
3) False positives (FP): negative records that are incorrectly labelled positive.
4) False negatives (FN): positive records that are mislabelled negative.

Below table shows the confusion matrix in terms of the TP, FN, FP, and TN values. Relying on the confusion matrix, the accuracy, sensitivity, and error rate metrics are derived. For a given models, the accuracy can be calculated by considering the recognition rate, which is the percentage of records in the test set that are correctly classified (Type of Herbal Plant). The accuracy is defined as

$$Accuracy = \frac{(TP+TN)}{number\ of\ all\ records\ in\ the\ testing\ set}$$

Actual	Predicted	
	Positive Class	Negative Class
Positive Class	True Positive(TP)	False Negative (FN)
Negative Class	False Positive (FP)	True Negative (TN)

Confusion Matrix

B. Graphical user Interface:

A graphical user interface (GUI) is an interface through which a user interconnect with electronic devices such as computers menus and other visual indicators or representations. GUIs graphically display guidance and related user commands, different text-based interfaces, where data and commands are strictly in text. GUI representations are managed by a electronic stylus such as a mouse.

Software applications use these and add additional GUIs of their own. All internet browsers, such as Chrome, Internet Explorer and Firefox use their own GUIs to allow the user to navigate through websites which may also have their own GUIs such as Facebook, Instagram, and amazon. For example, a video from a streaming video player inside a website, they will interact with different GUIs in total: the OS's, the browser's, the website's.

How interface with a computer is continuously being researched and modified. Human imagination has brought users from the keyboard to the mouse and touch screens. A visual language has evolved as GUI has become commonplace in both operating systems (Oss) and software applications.

V. EXPERIMENTATION

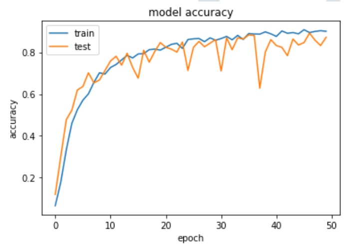
Model Training:

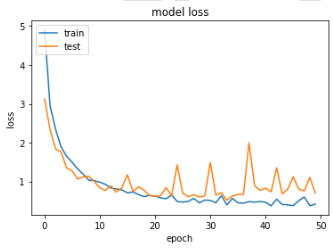
I. Importing the necessary libraries: All the required libraries are imported in one place — so that they can be modified quickly. For this Herbal Plant data keras is used.

II. Data Collection: The dataset for the proposed system has been collected from the "Medicinal Leaf Dataset" available online [15]. The dataset comprises of thirty species of healthy medicinal herbs such as Santalum album (Sandalwood), Muntingia calabura (Jamaica cherry), Plectranthus amboinicus / Coleus amboinicus (Indian Mint, Mexican mint), Brassica juncea (Oriental mustard), and many more. The dataset consists of 1500 images of forty species. Each species consist of 60 to 100 high-quality images. The folders are named as per the species botanical/scientific name

III.Pre-processing:
The captured image with cameras is noisy; hence pre-processing is required to remove the unwanted noise from the image. The proposed system utilizes a median filter to eradicate the salt and pepper noise. Median filtering is a valuable nonlinear process in reducing impulsive or salt-and-pepper noise. It is also helpful in preserving edges in an image while reducing random noise. Impulsive or salt-and-pepper noise (Is a form of noise sometimes seen on digital images. This noise can be caused by sharp and sudden disturbances in the image signal. It presents itself as sparsely occurring white and black pixels.) can occur due to a random bit error in a communication channel. In a median filter, a window slides along the image, and the median intensity value of the pixels within the window becomes the output intensity of the pixel being processed.

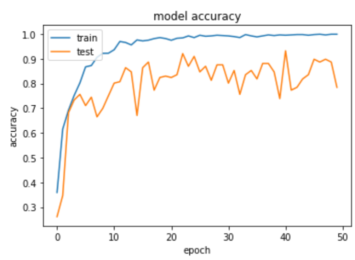
IV. CNN:
The CNN algorithms are trained with 11-layer architecture. The number of epochs is used are 50 to train the CNN network. The system achieved as training accuracy of 90.09%, training loss of 04172, validation accuracy of 87.22% while validation loss of 0.7043. The loss is more in the newly built CNN algorithm. Hence, we decided to built a transfer learning algorithm which uses weights of pretrained network.

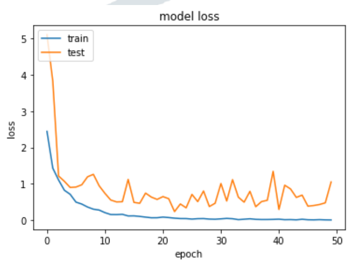




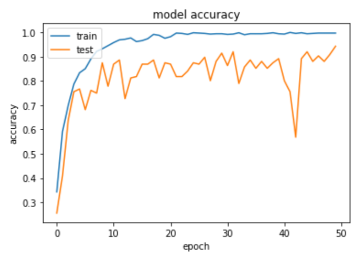
V. Vgg16:

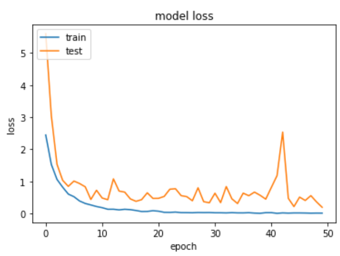
The Vgg16 algorithms are trained with 16-layer architecture. The number of epochs is used are 50 to train the Vgg16 network. The system achieved as training accuracy of 99.87%, training loss of 0.0081, validation accuracy of 78.41% while validation loss of 1.0521. The validation loss of Vgg16 network used in this approach is more hence there is a need on more layered architecture hence, in next phase we decided to trained the network with Vgg19 algorithm.





VI. Vgg19:
The Vgg19 algorithms are trained with 19-layered architecture. The number of epochs is used are 50 to train the Vgg16 network. The system achieved as training accuracy of 99.73%, training loss of 0.0124, validation accuracy of 94.32% while validation loss of 0.1946.





VI. RESULT

The comparative analysis of the proposed system for classification of medical plant leave using CNN, vgg16, vgg19 algorithm is tabulated in Table 6.1.

Table 6.1 Comparative analysis of performance of CNN, Vgg16 and Vgg19 algorithm

Algorithm	**Training**		**Validation**		**Execution Time in sec**
	Accuracy	**Loss**	**Accuracy**	**Loss**	
CNN	0.9009	0.4172	0.8722	0.7043	4053
Vgg16	0.9987	0.0081	0.7841	1.0521	3855
Vgg19	0.9973	0.0124	0.9432	0.1946	3329

From Table 6.1, it is observed that the vgg19 algorithm outperforms than the CNN and Vgg16 with the accuracy 99.73%. The execution time of the Vgg19 algorithm takes the minimum time than other two algorithm.

CONCLUSION:

In this project, the deep learning-based approach for medical herbal plant recognition system has been presented. The proposed system uses standard Medicinal Leaf Dataset is taken for evaluating the performance. The dataset consists of 30 different species of the medical plane.

In the phase I, the dataset is collected from the online source. First the dataset needs to preprocess and segment for further analysis, hence the images are converted into grayscale using weighted average method. Then images are segmented using thresholding algorithm, which produce binary image. The segmented part of the whole image is cropped to get a proper leaf part.

In phase II, the classification algorithms are implemented. We select CNN, Vgg16 and Vgg19 algorithm for classification strategy. Finally, the performance of the system evaluated using model accuracy and model loss. The CNN algorithms show the training and validation accuracy of 0.9009 and 0.8722, while Vgg16 achieved 0.9987 and 0.7841, and Vgg19 shows 0.9973 and 0.9432.

From the performance of the all the three algorithms, it is observed that the vgg19 algorithm outperforms than the CNN and Vgg16. The execution time of the Vgg19 algorithm takes the minimum time than other two algorithm. ACKNOWLEDGEMENT

It gives us great pleasure in presenting the preliminary project paper on ‘Herbal Plant Detection Using Deep Learning’.
We would like to take this opportunity to thank my internal guide Dr. C. G. Patil and co-guides Prof. S. S. Bhosale , Prof. S. N. Kamble for giving us all the help and guidance we needed. we are really grateful to them for their kind support. Their valuable suggestions were very helpful.
We are also grateful to Dr. S.K. Jagtap, Head of E&TC Engineering Department, SKNCOE for her indispensable support, suggestions.
We would also thank our Principal Dr. A.V. Deshpande for his great insight and motivation. Last but not least, we would like to thank my fellow colleagues for their valuable suggestions.

References

[1]. Nayana G. Gavhale, Dr. A. P. Thakare, “Identification of Medicinal Plant Using Machine Learning Approach,” International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 07, July 2020.

[2]. R. Akter and M. I. Hosen, "CNN-based Leaf Image Classification for Bangladeshi Medicinal Plant Recognition," 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE), 2020, pp. 1-6.

[3]. Izwan Asraf Md Zin, Zaidah Ibrahim, Dino Isa, Sharifah Aliman, Nurbaity Sabri, Nur Nabilah Abu Mangshor, “Herbal plant recognition using deep convolutional neural network,” Bulletin of Electrical Engineering and Informatics, Vol. 9, No. 5, October 2020, pp. 2198-2205

[4]. K. Priya, S. Ranjana, R. Manimegala, “Medicinal Plant Identification Using Andr Oid Application Based On Leaf Image,” European Journal of Molecular & Clinical Medicine, Volume 07, Issue 09, 2020, pp. 1496-1506.

[5]. Anh H. Vo, Hoa T. Dang, Bao T. Nguyen, and Van-Huy Pham, “Vietnamese Herbal Plant Recognition Using Deep Convolutional Features,” International Journal of Machine Learning and Computing, Vol. 9, No. 3, June 2019, pp.363-367.

[6]. C. Amuthalingeswaran, M. Sivakumar, P. Renuga, S. Alexpandi, J. Elamathi and S. S. Hari, "Identification of Medicinal Plant's and Their Usage by Using Deep Learning," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 886-890.

[7]. 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

[8]. 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

[9]. 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

[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.

[11]. 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

[12]. 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

[13]. 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

[14]. 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

[15]. "Medicinal Leaf Dataset," available online: https://data.mendeley.com/datasets/nnytj2v3n5/1