Identification of plants and determining its medicinal characteristics

[1] Brahma D [2] Fayaz Ahmed [3] Gnanesh M R [4] Bharath MB [5] Dilip B G
[6]Dr. Manjunath D R

[1][2][3][4][5] Final Year Students, Department of CSE, DSATM
[5] Associate Professor, Department of CSE, DSATM

Abstract

In Earth there large varities of planet consisting of number of many unknown and known species. Recognizing flora is ability to identify the plant species from the photographers and provide the medicinal information along with the diseases that can be cured by the plants it is an intelligent system .Creating a Tool that Identify the plant by data of certain characteristics features of Leaves. Comparison between standard data and recorded data is done based on predefined parameters. This will also help us to determine the medicinal values that particular plant has in by Classification based on the Characteristics. This identification should be automated as this process was done by human and every person could not identify accurately even if he identify correctly he could not be efficient.So with the help of the expert system can be designed .

Keywords: *Recognizing flora, medicinal information, predefined parameters, Automated*

I. INTRODUCTION

In India our ancestors were fully depended on Ayurveda for curing their diseases. Ayurveda is the oldest practice. The origins of Ayurveda has been traced to since 6000 BC when they originated from person to person as mouth words there was no document format. Several concepts of Ayurveda has existed seeing that in 330 BC during Indus Valley Civilization. The initial written forms of Ayurveda while medical work evolved from the Vedas that is in Atharvaveda. The ingredients for ayurvedic medicine is naturally available plants and their roots. In Ayurveda identification of the plants correctly plays a significant role. For making medicines the plants are collected by the people who do not have much knowledge of the plants and their use during collecting if they collect a plants that is mixed with other plant which is identical then that will lead to the adverse effect. To overcome this situation identification of correct medicinal plants an important task. In this system we implement a feature to get the description of medicinal value of the plant.

II. RELATED WORKS

Javed Hossain has used the approach in which the input image is converted to binary image with the basepoint and reference points selected by the user. Then the leaf image can be obtained utilizing a digital camera or Scanner. The noisy binary image is converted into enhanced binary image then it will create a angle of inclination The inclined leaf image will be turned into horizontally aligned binary image. The Horizontal leaf will have the Major axis l and minor axis m of a leaf. The following chore is to carry out classification of these lower dimensional features vectors. Then they apply the (PNN) for classification of leaf shape features for plantidentificatio

Taufik Hidyat used the approach in which the input image is converted to grayscale if the image is colored image. To separate background from the foreground of leaf, leaf Segmentation is done. Based on geometric properties such as area of leaf, roundness, slimness the feature characteristics in plant is identified. The training procedure done to recognize plants on the basis of the leaf image texture surface requires a parameters of the leaf image feature for training set. The method of backpropagation is used with 2 hidden layer utilizing the extraction of leaf feature of characteristics feature.

Sandeep Kumar E used the approach in which the procedure in this permits the recognition of medical plants on the basis of its edge features. The grey Scale equivalent image is obtained from the color image. The edge histogram is calculated using the grey scale image. Canny edge detection algorithm can be applied for this purpose This algorithms are applied to the test database image and image and area difference, edge color histogram and histogram is calculated.

A.D.A.D.S. Jayalatha have developed a leaf identification in the machine learning field. The supervised algorithm class is beyond by one of the wisely known method in Convolution Neural Network (CNN). It is utilized to detect what is present in the image and what is the image. During the process several features of images will be abandoned But CNN recognizes the image with greater accuracy. When the data collection was completed using CNN, creating a model to train each image of the flowers and leaves.

Stephen Gang Wu used the approach in which an RGB images is primarily reversed to a grayscale image. The extent of converting grayscale Images in to binary image is set on accordingly to the RGB histograms. frequently utilized "digital morphological features (DMFs)", derived from 5 basic feature, are taken so that a computer can procure feature values automatically and quickly on the basis of 5 basic features which was

introduced earlier, we defined 12 digital morphological feature which is used for leaf recognition. "An inter connected group of artificial neurons simulating the thinking process of human brain is called artificial neural network (ANN)". PNN is taken on for it has many benefits Its training speed is more times faster than a Backpropagation networks. As the running and training procedure can be carried out by matrix manipulation, the speed of PNN is very fast.

III. MACHINE LEARNING IN AYURVEDA FIELD

In Today's world Machine learning and artificial intelligence is playing a vital role. As days passes technology is becoming more advance and its new application is also growing with it. This technology helps in saving cost and many more. In using heartbeat diagnosis physical, Mental, emotional situation was predicted accurately. And the diseases was easily predicted by imbalance in the tridoshas. The coalition of Machine learning with Ayurveda made Ayurveda more hi-tech, comprehensive, trustworthy and more universal in recent times. As Ayurvedic system of medicine is affordable be common man and is more cost cutting with minimum side- effect, in Ayurveda e-commerce Artificial Intelligence can be used in order to stretch the market by notable improvement in user experiences. It can also be leveraged to improve economic applications that have a significant impact on cost reduction, revenue growth, and asset utilization.

IV.METHODS AND MATERIALS

A. Data Collection:

In present day, databases of Ayurvedic plants is unavailable. So the data set leaf detection is taken from the kaggle and is edited to our requirement. Medicinal plants images will be collected mainly from the Data set.

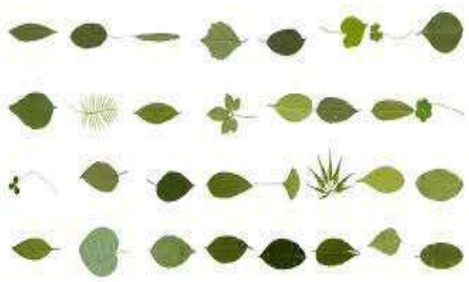


Fig. 1: A leaf image

B. Analysis phase:

In machine learning we are using Tensorflow for classification of data. It is open-source. It provides all resources libraries to perform the classification of data.

C. Convolution neural network

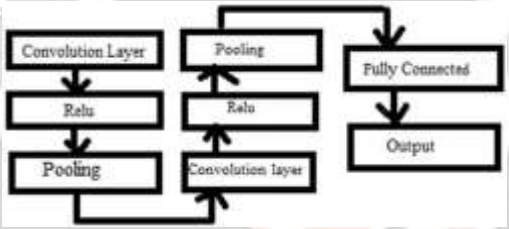


Fig.2 Convolutional Layers

Convolutional neural network is a kind of deep neural network, is using which analyzing of visual image is done. Neural network is utilized to recognise the image and what is present in the image. A CNN contains many convolution layers. It is a sequence of convolution relu and fully connected, grouped layers.

V. METHODOLOGY

The block diagram of the system that we are using is as below and we collected the image data set from the kaggle where we dint get specific data set for medicinal leaf itself so we manipulated the data set to our requirement.

After getting the data set we are performing the following tasks:

Image extraction: In this phase the actual leaf image data is being extracted eliminating the background data of the leaf image data.

Image Pre-processing: In this phase the image is converted into greyscale image and noise in the extracted leaf data is removed and if there is any distortion in the image if found then it is also removed by geometric tansformations of images such as translation, Rotation, Scaling.

Equalization of Histogram: In this phase Intensity distribution of the image takes place. So that classification will become easier.

Extraction of features: In this phase the feature of the image is extracted using which the classification is done after the classification the medicinal value of the leaf is determined.

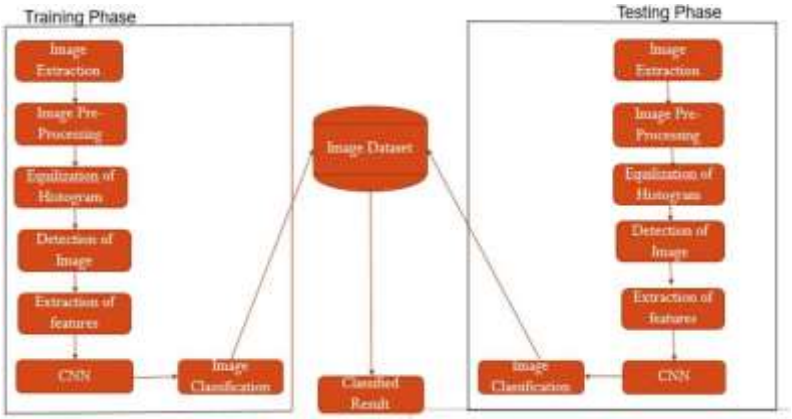


Fig. 3 Block Diagram

The System is trained with the image in the above phases and the accuracy of the trained data is tested

The good ratio of training that to testing is 7:1.5

VI. Module Algorithm

A. Image dataset is loaded to disk

In this the image data set downloaded from Kaggle is used for training the model.

B. Image processing.

Image is pre-processed , normalized and generating various possible image for the given image data.

C. Rescaling the image .

Image is rescaled to size (255,255). machine learning models train faster on smaller images. And many deep learning model ResNet50 architectures require that our images are the same size and our dataset images may vary in size.

D. Detection of Flower and its properties.

Various Classes of images are Identified. And Corresponding (Key : Value) pair will be created for the classes.

E. Model is Trained with RESNET50V2.

It is a Residual Network used in recognising images, the first layer may learn to detect edges, the second layer may learn to identify textures and similarly the third layer can learn to detect objects.

F. Prediction.

A function is defined which takes image as input, contains the medicinal properties of the classes defined, depending on the probability of classification of the image. The medicinal properties will be determined.

VII. Module Implementation.

A. Directories of the datasets:

In this we load the dataset, we assign a variable with the path of the data set.

B. Rescaling the Images :

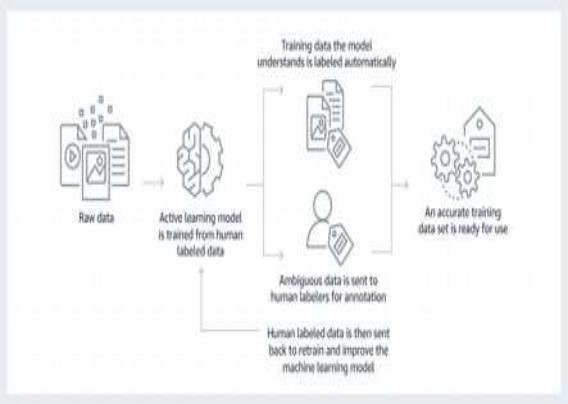
In this image is resized to (255,255). This is done because the image present in the dataset is of different size we need to rescale all images to standard size. And machine learning models train faster on smaller images.

C. Image Data Augmentation.

In This we Generate various possible images using ImageDataGenerator. With a batch size of 256. Image data augmentation is a technique that can be used to expand the size of a training dataset by creating modified versions of flower images in the dataset. improve the performance and ability of the model to generalize.

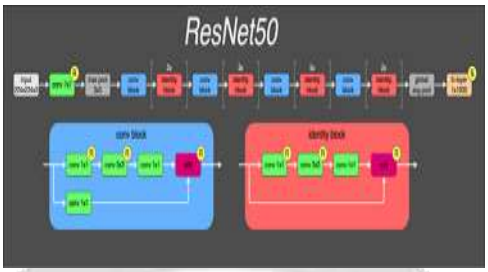
D. Understanding the Data.

Class_indicies dictionary containing the mapping from class names to class as part of .flow_from_directory() from its ImageDataGenerator class.
We will define a label , That performs labeling of classes. That is identifying raw data of flowers and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it.



E. Using Resnet50.

For ResNet, we call tf.keras.applications.resnet.preprocess_input on Flower inputs before passing them to the model. resnet.preprocess_input will convert the input images from RGB to BGR, then will zero-center each color channel with respect to the ImageNet dataset, without scaling.

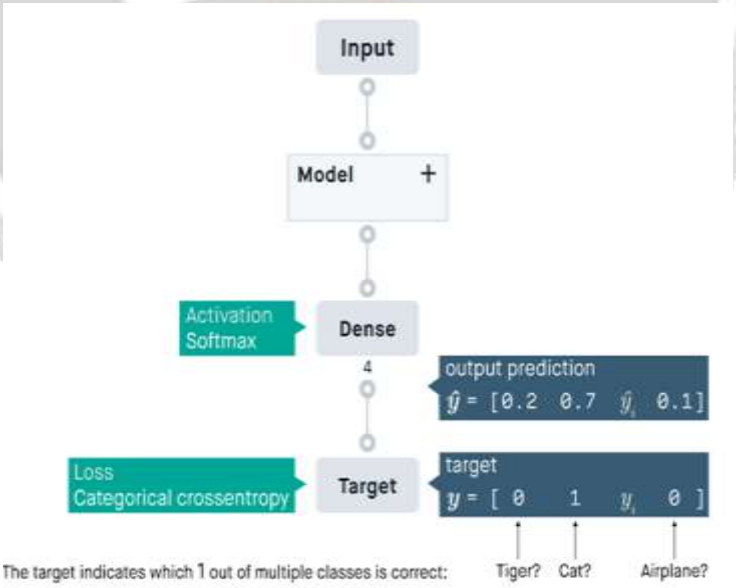


F. Adding the layer to the Dense Layers.

The output shape of the Dense layer will be affected by the number of neuron / units specified in the Dense layer. A Dense layer feeds all outputs from the previous layer to all its neurons, each neuron providing one output to the next layer.Having 2048 neurons

G. Training model Use a small learning rate for better accuracy.

Learning rate of 0.0001 is given along with RMSprop Optimization. This optimization balances the step size decreasing the step for large gradients to avoid exploding, and increasing the step for small gradients to avoid vanishing. We are training with more five classes of data, to set the lossfunction to "categorical_crossentropy" .These are tasks where an example can only belong to one



out of many possible categories, and the model must decide which one. We use accuracy as our metrics to train the model. Model fit Trains the model for a 50 epochs. That is we iterate 50 times on a dataset.
We also use the Generated various possible images using ImageDataGenerator during image augmentation phase.
Validation data: the validation loss of data provided using validation data which helps training model with the Valid Data.

H. Check for Your Own Image:

We are defining a function that takes the image path as a parameter. Loading this image into PIL format. This is done because the image which we use can be of JPG,PNG so to Standardize to work with all image formats we are using PIL format. We convert PIL Image instance to a Numpy array. And resize it to (255,255).We Will Insert a new axis that is we use the image array and perform comuptation in y-direction . We can predict the class for this new image using the predict_class() function of keras.
The output of this will be class indices. Depending on probablity of class indices the corresponding medicinal charactristics of a class will be predicted.
This medicinal value will be defined as a dictionary data type in which class index will be the Key and medicinal characteristics as the value for the key.

VIII. RESULT.

Accuracy and Loss of Existing System.

Fig.4: Accuracy graph on training set

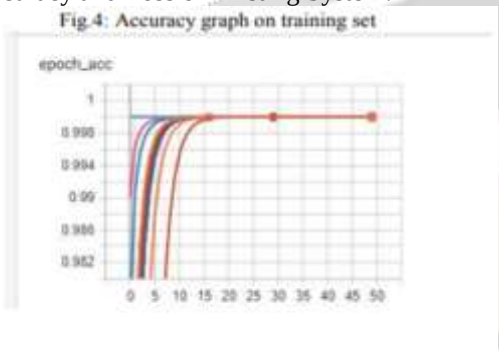


Fig.5: Validate-Accuracy graph on training set

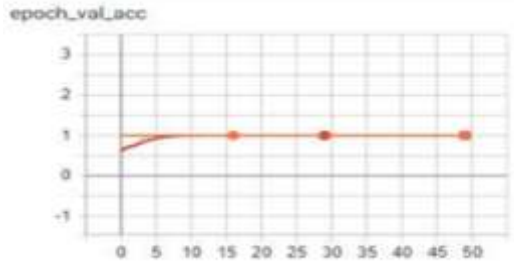
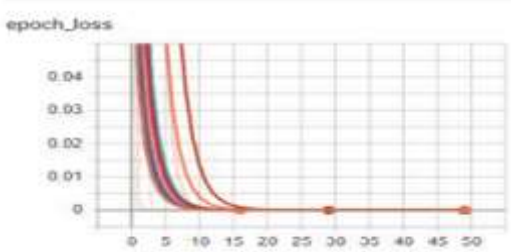
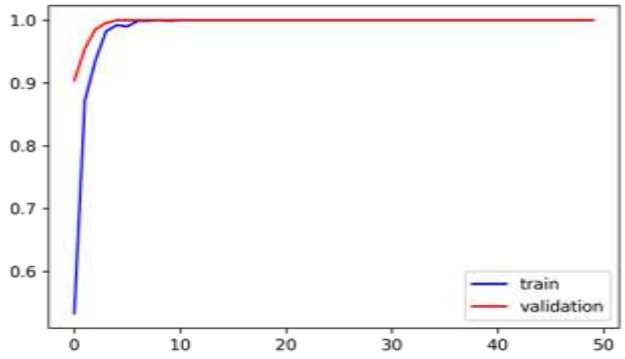


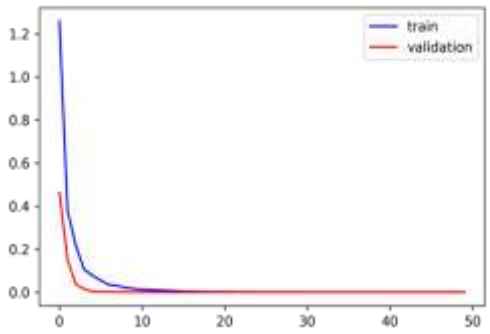
Fig.6: Loss graph on training set



In the existing system Epoch used is 60 and used with CNN . And the test set accuracies were between 95% - 97% when tested on 500 images each of 10 rare Species of flowers.



Accuracy-Graph



Loss-Graph

Accuracy-Graph: A plot of Training accuracy versus Validation accuracy
Loss-Graph: A plot of Training Loss versus Validation Loss.

In this we are using epoch of 50 and Residual network ResNet50 .And the test set accuracies were between 97% - 98% when tested on 906 images each of 5 Species of flowers

IX. CONCLUSION

In unit we have proposed a strong technique using CNN for the recognition of sparse medicinal plants. Tested accuracy will be high using the TensorFlow upon dataset which we created. This result will be aquired from the taking out the required characteristics of the leaf image. The accuracy increases with the number of training.

X. REFERENCES

[1]. L. Yang, F. Quan, and W. Shuzhi, ‘‘Lightweight CNN-based plant disease recognition method and mobile terminal application,’’ J. Agricult. Eng., vol. 35, no. 17, pp. 194–204, 2019.

[2]. Z. Shanwen, Z. Qingqing, and Q. Guohong, ‘‘Plant leaf recognition method based on the combination of Fourier descriptor and LBP,’’ Jiangsu Agricult. Sci., vol. 47, no. 14, pp. 273–276, 2019.

[3]. R. Shengnan, S. Yu, and Z. Haiyan, ‘‘Small sample plant disease recognition based on one-shot learning,’’ Jiangsu J. Agricult.Sci., vol. 47, no. 5, pp. 1061– 1067, 2019.

[4] A Complete Guide to Types of Neural Network. Available at: https://www.digitalvidya.com/blog/types-of-neuralnetworks

[5] Plant Identification Using Tensorflow. Available at: https://digitalcommons.calpoly.edu/cgi/viewco ntent.cgi?article=1271&context=cpesp

[6] Convolution Neural Networks (CNN) : Step 1 Convolutional Operation. Available at: https://www.superdatascience.com/blogs/convolution al-neural-networks-cnn-step 1-convolutionoperation.