CNN Based Approach to Identify Hibiscus Plant Species

PALLAVI SHETTY[1], DR. BALASUBRAMANI R[2]
[1] *Associate Professor, Department of MCA, NMAMIT, Nitte*
[2] *Professor, Department of ISE, NMAMIT, Nitte*

Abstract- ***Plants are vital to human survival. There are numerous rare plant species that humans need in their daily lives, whether for medicinal or other purposes. It is everyone's responsibility to recognize the need for such species and protect them for future generations. As a result, identifying plant species is primarily a research platform. Because leaf images are 2D in nature, while stem images and flower images are 3D, it has been proven in the technical plant identification process that identifying a plant using its leaf images is easier than identifying a plant using other components such as stem images, flower images, and so on. As a result of considering these factors, some Hibiscus species have been detected and classified using CNN algorithm technique.***

Indexed Terms- ***CNN, Hibiscus Species, Deep learning algorithm, leaf data set.***

I. INTRODUCTION

Hibiscus plant species comes under the family called Malvaceae. It has plenty subspecies, among which eight species identified for this research work. These species are used in medicine. Hibiscus Rosa Sinesis, Hibiscus Sabdariffa, Hibiscus Mutabilis, Hibiscus Schizopetalus, Hibiscus Syriacus, Hibiscus Trionum, Hibiscus Esculentus and Hibiscus Hispidissimus.

This process is accomplished by performing few important steps: creating dataset for predicting the species, creating labels, pre-processing (leaf images), undergoes trains-tests split stage, create CNN model for training the dataset to predict the type of species. Then Create dataset for classifying the species.

II. LITERATURE SURVEY

Identification of plant species has a wide range of socioeconomic implications. Pattern recognition is being used to identify plants, and there is a lot of research going on in this area. Rural medicine has the potential to return, as it did in previous decades, with the use of solid algorithms for leaf identification. CNN-based algorithm used for identifying Indian plant species using leaf image with a white background [1].

The proposed method recognizes a plant in pre-processing, feature extraction and classification. Morphological characteristics, Fourier descriptors and shape-defining features are extracted. The artificial neural network's input vector is made up of these features [2].

The Convolution Neural Network model is proposed for plant species identification. Convolution, maximum pooling, and global average pooling are used in the CNN procedure. Because of the irregular count in photographs, a model can be developed using over or under-sampling the data as a further improvement. Pre-processing techniques such as image segmentation, masking, and others may also help increase accuracy. Finally, if a mobile application based on the model is created, the model can be used by farmers for real-time plant identification [3].

Although there are numerous methods for detecting and classifying plant diseases using automatic or computer vision, research in this area is still lacking. Furthermore, no commercial solutions exist, with the exception of those that deal with plant species identification based on leaf images. In this paper, a novel approach to automatically classifying and detecting plant diseases from leaf images was investigated using deep learning methods. The established model was able to detect the presence of leaves and differentiate between healthy leaves and 13 diseases that can be visually diagnosed. The entire procedure was defined, from collecting the images used for training and validation to image pre-processing and augmentation, and finally training and fine-tuning the deep CNN. Various experiments were

carried out in order to assess the efficiency of the newly developed model [4].

In numerous tropical countries, this plant is used to heal wounds, inflammation, fever and coughs, diabetes, infections caused by bacteria and fungi, hair loss, and gastric ulcers. Hibiscus blossoms have traditionally been utilised as analgesics, antipyretics, anti-asthmatics, and anti-inflammatory medicines, as well as having anticancer effects [5],[6].

Hibiscus sabdariffa, also known as "red sorrel" or "roselle," belongs to the Malvaceae family. It is a well-known medicinal plant with over 300 varieties found in tropical and subtropical climates around the world. In a warmer, more humid area, Roselle may adapt to a variety of soil types. Many parts of the Roselle plant, including seeds, leaves, fruits, and roots, are used in cuisine and herbal medicine as a non-pharmacological treatment. Roselle extracts are used to treat a variety of medical conditions, including numerous cardiovascular ailments, helminthic disease, and cancer. The herb is also an anti-oxidant and is used to treat obesity [7],[8].

III. SYSTEM DESIGN AND ARCHITECTURE

1. Hibiscus species dataset: Collectively above 500 hibiscus images are used in this work which contains 8 type of distinct hibiscus species.
2. Label creation: Here labels are generated for each hibiscus leaf species present in dataset which contains information or data about their specifications.
3. Algorithm used in this research work is Convolution Neural Network [CNN]. It is a deep learning algorithm used to find samples, classifying data and also to learn features in images.

- Working of CNN: It has to execute 4 layers:

1. Convolution layer
   The aim of layer is to extract characteristics of the input images. It studies the characteristic of images and correlates with pixel values by using squares of input images.
2. Activation layer: Execution of this layer is carried out after every convolution layer. The operation of this layer is to introduce non linearity. The activation function used in this project is RELU (rectified linear unit).
3. Pooling layer:
   Pooling layer is executed to reduce the dimension or magnitude of the characteristic outline by keeping the necessary essential data. This project uses MAX pooling method which keeps large elements of characteristics outline.
4. Fully connected layer:
   This layer aggregates all the data from the last characteristics outline into a single vector value (i.e., it creates final classification). At this point activation function is changed to softmax because it functions on any random real valued vector and changes into a vector value to obtain the sum of 1.

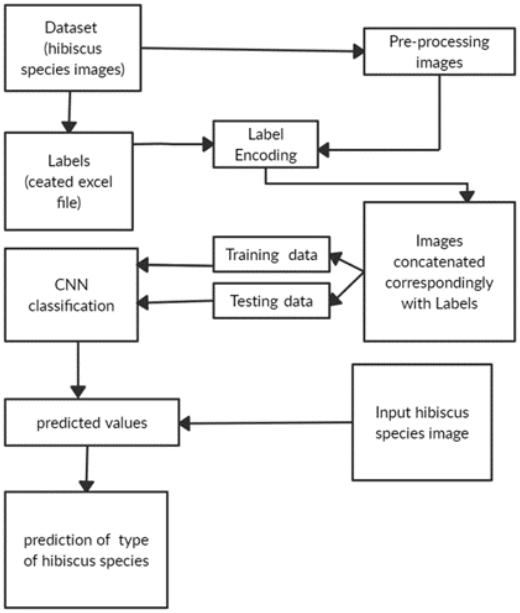

Fig 1: Block diagram that shows working of the Model.

IV. IMPLEMENTATION

Implementation using CNN algorithm to predict the type of hibiscus species.

Step1: Installation of the libraries that is required to implement the project. Libraries used are Tensorflow (Backend), Keras, Matplotlib, Pandas, Numpy, Sklearn.

Step2: Load the hibiscus species dataset that holds hibiscus leaf images. In this framework 8 types of hibiscus leaf images are used. Then upload the generated labels in the format of CVS file which contains labels of each image.

Step3: Pre-processes the dataset to make sure all the images are of equal size. If any images are not of same size then resize all images to same size. Original size of all images is 256x256. After pre-processing the dataset each hibiscus leaf image is resized into 350x350.

Step4: Once the process of resizing of all the hibiscus leaf images of dataset is completed. Then label encoding needs to be performed so that each hibiscus leaf image concatenates with its respective label.

Step5: After processing label encoding the hibiscus leaf species dataset has to be divided into two parts for training and testing. 80% of the data from dataset is passed for training to CNN model and 20% of the data is passed for testing to CNN model. This process is accomplished using a library called Sklearn.

Step6: Now develop a Convolutional Neural Network model. In this process model is built with 7 layers. Approach of running the model is sequential. This model is built using 16, 32, 64, 128, 128, 128 and 128 Conv2D filters. Size of the Kernel used in this model by all 7 layer is 3x3. Pooling filter for all the layers is 2x2. The 80% of the training dataset and the 20% testing dataset is now passed to CNN model. Then the CNN operates its four layers which are described in system design and architecture (section III). The input image is of dimension 350x350x3 that is passed to convolutional neural network to which kernel filter 3x3 is applied. A filter is moved one pixel at a time for stride value 1 to obtain components outline. Now to the components outline activation function ReLU is activated. Now BatchNormalization function is applied because it provides neural network to output more stable prediction. Once the components outline are normalized, to the output MaxPooling filter 2x2 is applied. MaxPooling filter is moved one pixel at a time for stride value 1 to obtain output components outline which contains only the essential information. Next dropout function is applied to pool components this will help neurons in the network to reduce the co-dependency amongst each other during training phase so that over-fitting problem is reduced. These same operations are performed by rest of the layers. Then the final output obtained from all the five layers are fattened to 1D or a vectored value which gives us the prediction values of the trained data. If accuracy and loss value is satisfied (accuracy value >= 90% and loss value <50%) then consider it is final model and save the weights so that the saved weights can be used to predict the user input image and show the prediction obtained for hibiscus leaf images.

V. RESULT AND DISCUSSION

Hibiscus species dataset is passed to CNN algorithm for training and testing to obtain the type of classification of hibiscus leaf species. For which dataset was separated into 2 parts that is 80 of the hibiscus leaf images is passed to CNN model for training with batch size 5 and epochs 10 and it resulted with accuracy value 0.920 and loss value 0.1905. The 20 of the hibiscus leaf images is passed to CNN model for testing with batch size 5 and epochs and it resulted with accuracy value 0.9020 and loss value 0.2017

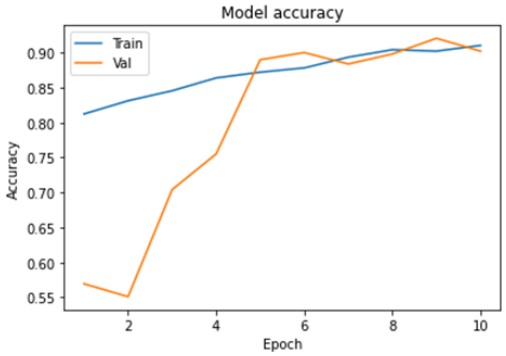

Fig 2: learning curve graph that shows predicted accuracy value for trained and tested data.

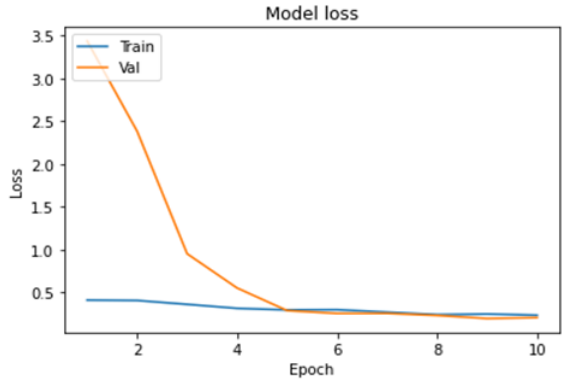

Fig 3: learning curve graph that shows predicted loss value for trained and tested data.

Prediction of tested input images for some types of hibiscus species classification are shown below.

```
Index(['Hibiscus Rosa Sinesis', 'Hibiscus Sabdariffa', 'Hibiscus Mutabilis',
       'Hibiscus Schizopetalus', 'Hibiscus Syriacus', 'Hibiscus Trionum',
       'Hibiscus Esculentus'],
      dtype='object')
prediction of Hibiscus Species type
Hibiscus Rosa Sinesis
array([0.29841143, 0.02943981, 0.1746229 , 0.09676033, 0.10633022,
       0.02987111, 0.01030636], dtype=float32)
```

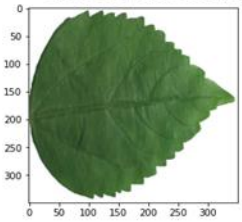

Fig 4: Predicted type of Hibiscus species is Hibiscus Rosa Sinesis.

```
Index(['Hibiscus Rosa Sinesis', 'Hibiscus Sabdariffa', 'Hibiscus Mutabilis',
       'Hibiscus Schizopetalus', 'Hibiscus Syriacus', 'Hibiscus Trionum',
       'Hibiscus Esculentus'],
      dtype='object')
prediction of Hibiscus Species type
Hibiscus Sabdariffa
array([0.00344664, 0.9818864 , 0.03896269, 0.01316845, 0.03213722,
       0.02014133, 0.01405019], dtype=float32)
```

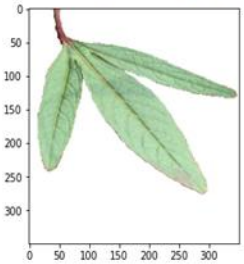

Fig 5: Predicted type of Hibiscus species is Hibiscus Sabdariffa.

CONCLUSION

This work provides information of different types of hibiscus species and methods used for classifying or predicting the different types of hibiscus species. The classification or prediction of hibiscus species is possible using deep learning algorithm called CNN, the CNN model is trained well using more than 500 images which resulted with good accuracy value above 90% and loss below 50%. Hence it satisfies all the terms and can successfully predict and classify all the 8 type of hibiscus species.

REFERENCES

[1] M. Vilasini, P. Ramamoorthy, CNN Approaches for Classification of Indian Leaf Species Using Smartphones, Computers, Materials & Continua CMC, vol.62, no.3, pp.1445-1472, 2020

[2] Aimen Aakif, Muhammad Faisal Khan, Automatic classification of plants based on their leaves, Biosystems Engineering, vol. 139, pp. 66-75

[3] M. Jaiganesh1, M. SathyaDevi2, K. Srinivasa Chakravarthy3, Ch. Sarada, Identification of Plant Species using CNN- Classifier, JOURNAL OF CRITICAL REVIEWS ISSN-2394-5125 VOL 7, ISSUE 3, 2020

[4] Srdjan Sladojevic,1 Marko Arsenovic,1 Andras Anderla,1 Dubravko Culibrk,2 and Darko Stefanovic, Deep Neural Networks Based and Neuroscience Volume 2016, Article ID 3289801, 11 pages.

[5] Asmaa Missoum, An update review on Hibiscus rosa sinensis phytochemistry and medicinal uses, Journal of Ayurvedic and Herbal Medicine 2018; 4(3): 135-146.

[6] Ghosh A, Dutta A. GC-MS analysis and study of potential antioxidant activity of the crude ethanolic flower extract of Hibiscus rosa sinensis L (wild variety) by hydrogen peroxide scavenging assay. International Journal of Current Trends in Science and Technology. 2017;7(11):20405–20410.

[7] Pragya Singh,1 Mahejibin Khan, Hailu Hailemariam, Nutritional and health importance of Hibiscus sabdariffa: a review and indication for research needs, Journal of Nutritional Health & Food Engineering, eISSN: 2373-4310

[8] Okereke CN, Iroka FC, Chukwuma MO. Phytochemical analysis and medicinal uses of Hibiscus sabdariffa. International Journal of Herbal Medicine. 2015;2(6):16–19.