International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 12 | Dec 2020 www.irjet.net p-ISSN: 2395-0072

Leaf Classifying Model in Crop Identification using Machine Learning Algorithm

Dr.G.Muneeswari[1]

[1]*Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST (Deemed to be University), Bangalore*

Abstract – *In the current scenario, one of the problem faced by the farmer is the identification of crops which is one of the biodiversity plantation fields. Many farmers have different names to the crops sharing the same features and same name to the crops sharing different features, by using image processing and deep learning through convolutional neural networks, the identifying of crops through images is possible. This paper introduces a basic approach in classifying and identifying crops of the biodiversity beds using Anaconda's Jupyter notebook and various other python libraries for the purpose of pre-processing, feature extraction and creation of a deep learning model using Random Forest Algorithm.*

Key Words: Leaf Identification, Machine Learning, Random Forest, Deep Learning, Feature Extraction

1. INTRODUCTION

Image Classification and Recognition have been can be used for a countless number of applications and has been a trending domain under Artificial Intelligence for android users [7]. This paper is a model which is capable of classifying and identifying plants by simply taking a picture of the leaf, taking into account that the leaf of a plant holds a unique identification to it and is a reliable biometric for the classification of plants. The model is trained to identify the features of the leaves which make it unique and uses these features to classify and identify leaves with similarity in these features when put together [8].

Since the identification of a plant as a whole could have many challenges with respect to extracting relevant features, removing noise around the image, identifying the unique objects within the image of a plant, etc can be a tedious and difficult task, the image processing of leaves was chosen as a better and more efficient way of identifying plants. Leaves differ in shape in accordance with their conditions such as their age after budding, their presence or absence of insect/disease damage, absence of lobes, etc [9]. The most common leaf shapes are oval, truncate, elliptical, lanceolate and linear. Leaf tips, bases and arrangement may also be unique. The margin of the leaf also exhibit differentiating features. The major classifications of leaf margins are entire leaf, toothed/serrated leaf, lobed leaf and parted leaf[10]. In terms of texture, leaves have unique structures called veins which are responsible for transportation of water and nutrients to leaf cells. The two main classifications of venation patterns are Pinnate venation and Palmate venation. All these distinct features of leaves along with many other make it the ideal candidate for classification and identification of different plant species.

The paper is organized as follows: Section 2 describes about literature survey, Section 3 elaborates on working methodology and section 4, 5 explains about evaluation results and conclusion.

2. RELATED WORK

For the Android [1] detection and reorganization of plant diseases are provided with a comparative study on results acquired by various results of different features such as shape, colour and texture which were extracted and the accuracies secured from various classifiers such as kNN, PNN, ANN, RF, etc. In [2], the methodology talks about using image processing for the purpose of classifying leaves. For this, plants with similar properties into the same class (Eg: medicinal, non-medicinal and poisonous plants). In [3] initially. The training set consisted of 27 leaves with 9 pictures of each leaf with different noise types. The paper proved that neural networks give high accuracy in leaf recognition and that ANNs prove a good method of classification of the leaf images.

The paper [4] provided information on the pre-processing of binary leaf images so as to remove noise and understanding the spatial parameters of a leaf using image processing. A paper on a method of Feature extraction from Leaf Architecture [5] talks about important features of leaves which are independent of leaf growth. Image processing which deals with translation, rotation and scaling are also studied. The paper provided insights of creating a Global Feature Vector for the purpose of classification, which has been adopted in the proposed work. Various pre-processing techniques mentioned in the paper were experimented upon and chosen according to the requirements of the proposed work. The paper mentions 5 Basic Geometrical features of leaves which include Longest Diameter, Physiological Length , Physiological Width, Leaf Perimeter and Leaf Area, as well as Digital Morphological Features such as Smooth factor, Aspect ratio, Vein features, Perimeter ratio, etc. Plant Species Identification Using Colour and texture analysis of rice leaves [6] proposed by M. RavindraNaik et al provides a thorough analysis and comparison of primary studies on computer vision approaches for plant identification. The main features which were studied in this paper were shape, texture, colour, texture and margin, all of which have been adopted

as the primary features used to identify leaves in the paper. The paper also provides the generic steps of an image-based plant classification process, which has been taken as the route architecture of the work.

3. PROPOSED METHODOLOGY

In the proposed method, the leaf classifying model takes into account a number of features of the leaf and put them in a Global Feature Vector for the purpose of classification. The features which have been taken into account are shape (margin), colour, texture (venation pattern), and slimness (length/breadth). For the purpose of training, a dataset was first made with pictures of a few types of leaves inside the campus as a dataset of leaves which suited to the requirements of the project could not be acquired. The dataset prepared for testing is then use to check accuracy of the model. The model makes use of a Random Forrest classifier to classify image features against the Global Feature Vector. Upon classification, the highest probable value is labelled onto the im- age and produced as the predicted output.

Whereas for the leaf classification, identify different kinds of leaf so that when the dataset for different kinds of crops arrived it would be easier to implement the same methodology. Different crops have different names and characteristics, having a module to help with the image identification is quick and versatile. The leaf of a plant holds a unique identity and can be used to classify and identify plants. Since the number of plant species in the world are countless, it is impossible to acquire the entire dataset of plant species. So, a small dataset of leaves from within the Mindtree campus has been created for the purpose of creating the model. The model must be reliable and efficient in identifying the leaf by extracting various features from it, which make it unique to any other leaf of a different plant species. The various steps to be carried over the classifying model are given as follows:

1. Making a dataset of leaf images of different plant species.
2. Pre-process the leaf images to eliminate noise, crop, scale, rotate and filter the image to get best results.
3. Construct a model which is trained to understand the relevant features of leaves of various plant species.
4. Test the model to classify any leaf image within the plant species available in the dataset and acquire accurate results
5. Display relevant information about the plant which has been identified

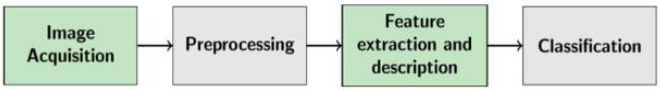

Fig -1: System Architecture

The entire system architecture is shown in Fig.1. Various research papers on feature extraction of leaves were referred to find the best features to acquire optimal results on classifying leaf images. Moreover, a thorough learning of image processing with Open-CV, which involved thresholding, adaptive thresholding, contouring, masking, cropping, etc. were used to pre-process the leaf images so as to increase accuracy of the model. A complete study on the anatomy of leaves, the distinguishing features and the dimensions of various leaves were studied to get a thorough understanding of the relevant features to be extracted and methods to extract them. The features to be considered should be independent of the age of the leaf to get a reliable model.

4. EXPERIMENTAL RESULTS

A training dataset of leaves is first loaded onto the model for feature extraction after pre-processing of the images. The images are pre-processed to remove noise, normalize, scale, crop, rotate the images to extract the features as effectively as possible. Specific features of the leaves are extracted from the leaf and the values extracted are stored into a Global Feature Vector for each leaf.

Fig -2: Angular Analysis of Leaf Set-1

The leaves are classified by a Random Forrest Classifier and the highest probable prediction is displayed as the result.

Fig -3: Angular Analysis of Leaf Set-2

Fig -4: Angular Analysis of Leaf Set-3

Fig -5: Angular Analysis of Leaf Set-4

In Fig.2, Angular Analysis of Leaf Set-1, In Fig.3, Angular Analysis of Leaf Set-2, In Fig.4, Angular Analysis of Leaf Set-3, In Fig.5, Angular Analysis of Leaf Set-4 are depicted appropriately. In Fig.6, the confusion matrix is shown for the predicted label vs true label and the results show 95% accracy for the random forest method.

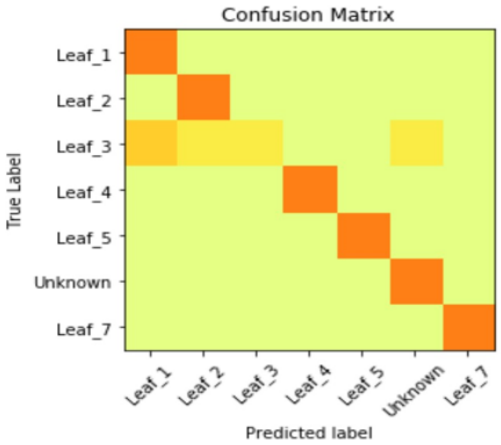

Fig -6: Confusion Matrix with predicted label and true label

The features of relevance which have been extracted in the model are the shape, colour, texture and slimness of the leaf as they have been identified as the most essential features which needed to be extracted from prior study and research on the anatomy of a leaf. The performance analysis shown in Fig.7, focuses on different feature selection algorithms and our current work expects to give promising results in terms of detection rate and accuracy. In an aim to increase the accuracy and detection rate, we used two stage process for data selection.

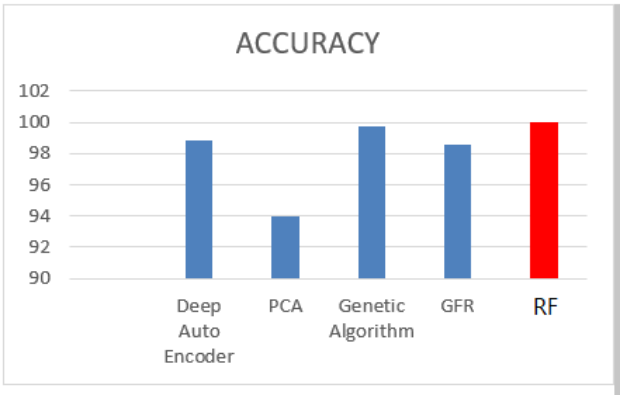

Fig -7: Performance Analysis with Random Forest Classifier

Our proposed model will show a very high accuracy rate comparing with the above-mentioned models. Since we are using an Random forest classification for angular analysis the time taken to model the leaf classification also will be less. In this current paper, we provided a solution to reduce the time taken to train the model with the reduced features. From the above graph it is clear that RF shows the highest accuracy.

4. CONCLUSION

Image Classification and Recognition have been can be used for a countless number of applications and has been a trending domain under Artificial Intelligence. The project is a model which is capable of classifying and identifying plants by simply taking a picture of the leaf, taking into account that the leaf of a plant holds a unique identifi- cation to it and is a reliable biometric for the classification of plants. From the dataset which was created, the images were pre-processed using Open-CV to remove noise, normalize, scale, crop, rotate the images so that they can be processed more effectively and accuracy can be increased. The dataset prepared for testing is then use to check accuracy of the model. The model makes use of a Random Forrest classifier to clas-sify image features against the Global Feature Vector. Upon classification, the highest probable value is labelled onto the image and produced as the predicted output.

REFERENCES

[1]Renuka Rajendra Kajale, "Detection and Reorganization of Plant Leaf Disease Using Image Processing and Android O.S." March-April 2015.
[2] Sachin D. Khirade and A. B. Patil. "Plant Disease Detection Using Image Processing." International Conference on Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on,pp. 768-771. IEEE, 2015.
[3] Harshad Shetye, Tejas Rane, Tanmay Pawar, Prof. Anuradha Dandwate" An Analysis of Methodologies For Leaf Disease Detection Techniques " February 2016.

[4] Tzionas, P., Papadakis, E.S., Manolakis, D.: Plant leaves classification based on morphological features and a fuzzy surface selection technique. In: Fifth International Conference on Technology and Automation, Thessaloniki, Greece, pp. 365–370 (2005)
[5] M. RavindraNaik, Chandra Mohan Reddy Sivappagari "Plant Leaf And Disease Detection By Using HSV Features and SVM Classifier" December 2016
[6] Sanyal, P., Bhattacharya, U., Parui, S.K., Bandyopadhyay, S.K.: Color Texture Analysis of Rice Leaves for Detection of Blast Disease. In: Proceedings of the 20th CSI Conference, pp. 45–48 (2006)
[7] A.S.Deokar, Akshay Pophale, Swapnil Patil, Prajakta Nazarkar, Sukanya Mungase, "Plant Disease Identification using Content Based Image Retrieval Techniques Based on Android System" 2016.
[8]I. EI Massi, Y. Es-saady, M. EI Yassa, D. Mammass and A. Benazoun. "Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers." 2nd International Conference on Electrical and Information Technologies (ICEIT), IEEE, 2016.
[9] Aakanksha Rastogi, Ritika Arora and Shanu Sharma. "Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic." 2nd International Conference on signal processing & Integrated Network (SPIN), IEEE, 2015.