International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Ayurvedic Plant Identification

Kshitija Pagare[1], Siddhi Nagare[2], Siddhesh Shigwan[3], Prof. M. R. Gorbal[4].

[1]B.E. Student, [2]B.E. Student, [3]B.E. Student, [4]Associate Professor of Information Technology Department
Department of Information Technology Engineering,
Shivajirao S. Jondhale College of Engineering, Dombivli East, Thane, Maharashtra 421201, India.

Abstract - *Ayurveda, the ancient form of medicinal practice is one of the most revered medicinal approach in the world. Ayurveda is mainly built on herbs, plants, flowers, fruits, vegetables and all plants that grows around us. Medicinal plants have been known to be an important possible source of curative aids. This involves the use of medicinal plants not only for the treatment of diseases but also as possible material for maintaining good health. Identifying an Ayurvedic plant in any huge vegetation is a monotonous work for us. We recognize a plant based on its size, leaves, flowers, fruits, etc. A leaf is a part of the plant which can be found on plants almost in all seasons and most of the time we have to identify plants on the basis of its leaf. While dealing with a leaf of a plant, it is important to consider the minute details of the curves representing the shape of the leaf. In our project, we are trying to build a system which has a database of 3 species of leaves. Our system involves mainly three procedures which includes resizing the image, feature extraction, and classification to get the result. A database of 95 leaves of 3 species i.e. Ajwain, Neem, Panfuti has been created. Each leaf image is resized first to decrease the time complexity of the classifier. Then for feature extraction we have used a scale-invariant feature transform (SIFT) algorithm. The SIFT is a feature detection algorithm to identify and describe local features in images. By using this algorithm we have extracted local features of plants like texture, edges, and values. And to classify the given leaf image we have used the K-Nearest Neighbor (K-NN) classifier which gives results on the basis of distance metrics.*

1. INTRODUCTION

Plants are a fundamental part of life on the Earth. Life on planet Earth cannot be imagined without plants as they are an essential component of the ecosystem and play an important role in the lives of animals as well as in humans by maintaining its balance. Plants are considered critical assets because in the many ways they support life on Earth. They release oxygen into the atmosphere, absorb carbon dioxide, provide habitat and food for animals and human beings, and regulate the water cycle. Many plants are used for medicinal purposes. Aloe Vera, Tulsi, Neem are a few examples. To implement the plant identification system based on feature extraction, the local features involved are shape, vein, and texture, edges. Once these features are extracted the key points are extracted from those features and value of that key point is calculated. This features are used as input to classifier for classification of leaf.

2. PROBLEM DEFINITION

To develop an efficient system which will identify medicinal plants from the leaf image (provided by the user), using dataset of leaves by making use of Feature extraction technique and machine learning algorithm, which is user friendly and easy to use

3. LITERATURE SURVEY

The paper titled "Contour based retrieval for plant species" by Asrani K, Jain R .In this paper for recognition of plant they have used Tangential Angle Approach. It is used for feature extraction invariant to scale, rotation and translation. The main advantage of this approach is that it is rotation, scale and position invariant. The image processing is done by Scanning of leaf images against plain background is done using flat bed scanner. They are processed to extract the feature vectors. There is no restriction on the dimensions of the image while scanning. Once the images are scanned, they are scaled maintaining the aspect ratio, as the property of aspect ratio is very important for identifying the leaf. After scanning the images next step is to extract the features. Feature extraction is done through edges. For extracting the edges, the colored leaf image is converted to binary image and then edges are detected. And the last step i.e. classification. For that the approach presented in this paper (tangential angle approach) enables to measure the similarity based on shape boundary. Tangential angle approach is used to extract the features. The main advantage of this approach is that it is scale, rotation and position invariant. Also the computational complexity is too low. The drawback of this paper is, only angle details cannot be considered as a criterion to represent the complete image. To improve the retrieval performance, the approach needs to be refined by considering more details of the boundary like magnitude& parameters representing the shape need to be identified so that the feature vector corresponding to the shape [1].The paper titled "Multiple Classifier System for Plant Leaf Recognition" by Araujo, V., Britto, A. S., Brun, A. L., Koerich, A. L., & Palate used Multiple Classification System for identifying the plants on the bases of color & shape features. A Static Selection method is used to search sets which will maximize the score. Image clef 2011 and 2012 datasets are used which shows different types of classifiers. The classifiers are mainly based upon shape and texture feature. The main focus of image processing is to remove unwanted structures from the leaf image. After converting the image to greyscale, the Otsu's method is used to segment

the leaf from the background. Next the feature is extracted based on the texture and shape. For classification SVM and Neural Network classifiers is trained on four different feature sets, namely, Local Binary Pattern (LBP), Speed of Robust Features (SURF) ,Histogram of Gradients(HOG), and Zernike Moments (ZM)This paper can be concluded as MCS approach also overcomes the performance of the monolithic solutions. Further work can be done on evaluating deep learning for feature learning as well as the use of dissimilarity feature space since the number of plant species is huge [2].The paper titled " Plant Identification System Using its Leaf Feature "by Dr. Pradeep Nijalingappa Madhumathi V. J. used multiclass SVM as (MSVM) classifier for identifying the plants based on Features of a leaf. A Grey Level Co-Occurrence Matrices (GLCM) is used to take texture features along with that authors have used morphological, vein features. Also mentioned FLAVIA and local datasets are used to store images of plant. Pre-processing is done by converting RGB images to grey level images improve the accuracy. Wiener filter and Median filters are used in order to remove the noise present in the picture. In this paper, morphological, vein and texture features are extracted. Grey Level Co-occurrence Matrices (GLCM) is used to take out texture features. Morphological Features extracted by using Physiological length, Physiological width, diameter, leaf area, leaf perimeter etc. Multiclass SVM (MSVM) is used as the classifier tool for classification. Conclusion can be drawn from this paper as or feature extraction they have used Morphological features which is a little complex because, they are also using other feature for extraction which increases the computational complexity[3]. The paper titled" Leaf Analysis for Plant Recognition "by Sahay, A., & Min Chen mentioned three components. Reprocessing operation is performed for noise reduction of plant. Then it identifies respective features and also computes scale invariant feature descriptor. Then matching of plant species, identification and returned are done by using weighted K-Nearest Neighbor search algorithm. In this paper to recognize various plants by the gray-level image of plant leaf, while ignoring the color information because all plants have same color i.e. green .To remove noise from the image, a morphological operation called erode-dilation is applied on the gray scale image Erosion method erodes away the boundary of the foreground object. It is useful for removing small noises from the image. After erosion method, noise is removed. For feature extraction, instead of using general image shape features such as contour and edge, local features are used. In leaf analysis application, local feature of an image was computed using Scale Invariant Feature Transform (SIF1). SIFT feature is invariant to rotation and scale. Classification is done by running a weighted K nearest neighbors search on database images. The conclusion can be drawn from this paper is instead of using general image shape feature they are using local features. Which represent highly localized information from small areas of an image and are defined around interest points [4].

4. DESIGN AND IMPLEMENTATION METHOD

4.1. WORKING

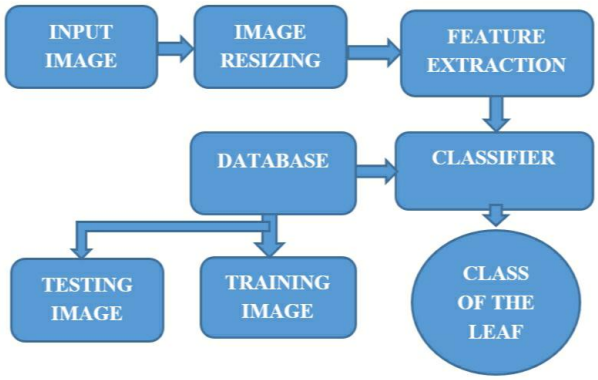
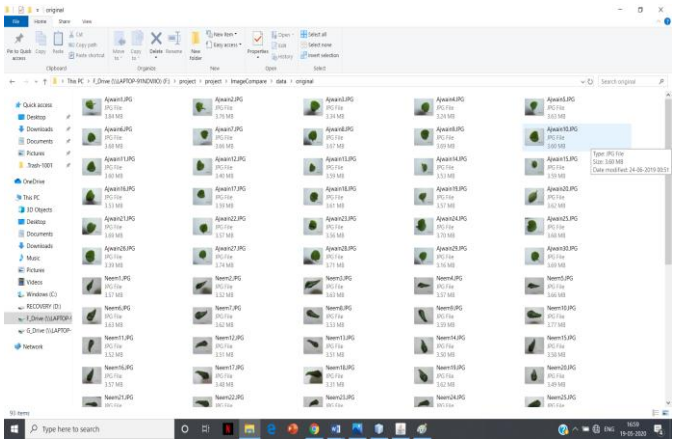


Fig -1: Working

In our project we have taken 90 images as a database on that we have done training and testing. Firstly we have resized the images just to decrease time complexity. In order to obtain the output at faster rate, the images are resized before giving to the classifier otherwise it will take more time to give the output. We have given 80 images to train and 10 images to test. By using sift algorithm the values of each and every pixel of trained database is calculated first. Then the remaining 10 images are given for testing to check the working of the classifier. Now the input leaf image which should be identified is checked for its image size. If the image is more than 3000kb then we have to resize that image. The value of each pixel of that image is calculated and feature is extracted by using SIFT algorithm. It is then given to classifier that works on K-NN algorithm to identify the class of the input plant leaf.

4.2. LEAF DATASET

The dataset is the collection of huge data. These data can be image, words, sentences, etc. In our system we have used the local dataset of 3 species of ayurvedic plants. The species are Ajwain (scientific name- Trachyspermum ammi), Neem (scientific name-Azadirachta indica), Panfuti (scientific name-Bryophyllum pinnatum) we have created a dataset with 30 images of each species. The images are taken by placing the leaf on white paper for white background. Each image is of 3-4 Mb hence for better performance of the system while training the dataset, we reduced the size of every image to 1mb or less. We have 150 images of three species, out of which 90 images (30 of each species) we choose for training purpose and 18 for testing purpose



Fi

Fig-2: Dataset of original image

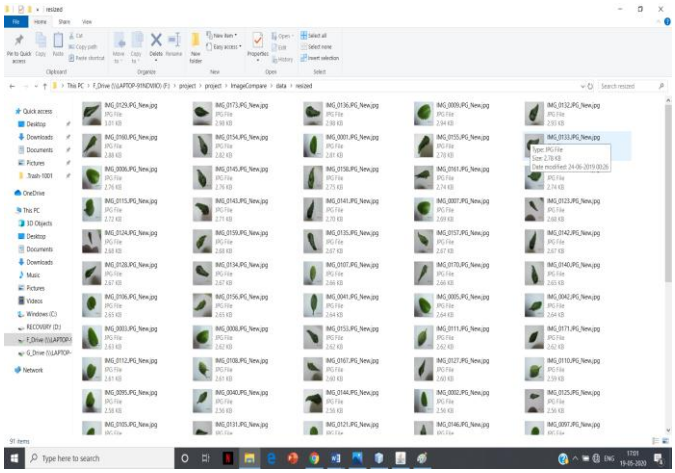


Fig-3: Dataset of resized image

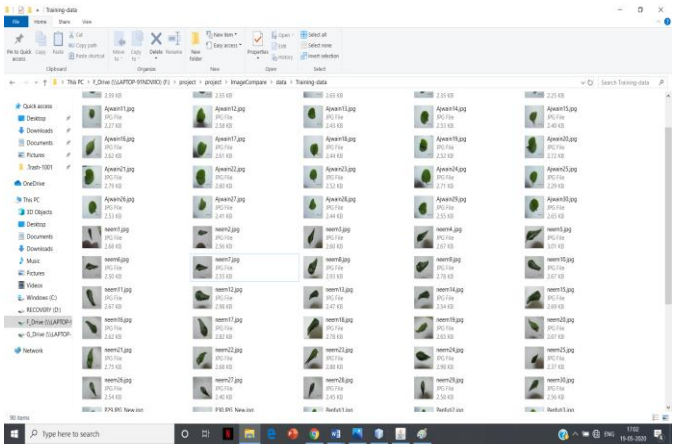


Fig -4: Training dataset

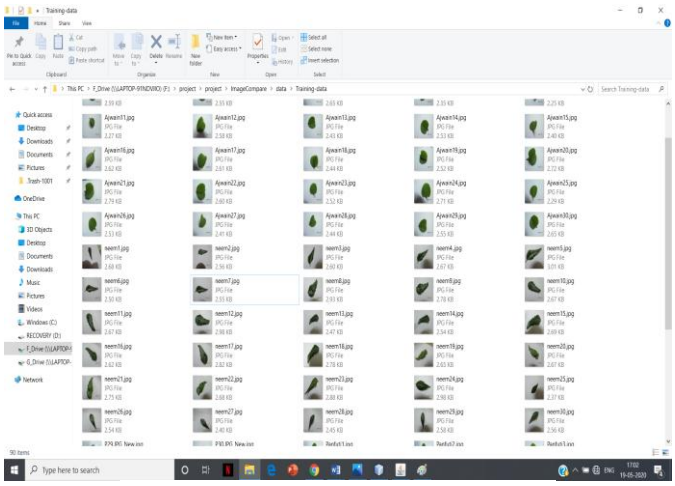


Fig-5: Testing dataset

4.3. FEATURE EXTRACTION

Feature extraction initiates from a uniformed data and creates obtained values and calculated to be informative, facilitating the succeeding learning and initiation steps. In our project we gave done feature extraction by using SIFT algorithm i.e. Scale Invariant Feature Extraction. The scale-invariant feature transform (SIFT) is a feature detection algorithm to discover and determine local features in images. By using this algorithm we have extracted local features of plants like texture, edges and values of training database and testing database. After extracting the features the value of each and every pixel of extracted feature is calculated by using SIFT algorithm and is stored in CSV file i.e. comma-separated value file

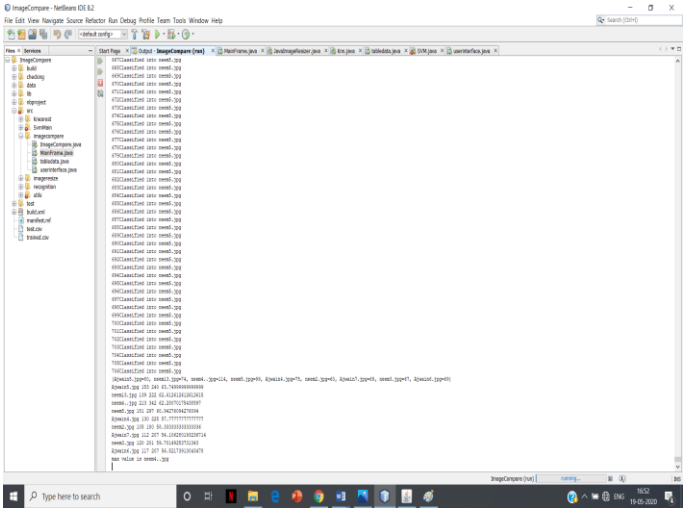


Fig-6: Values to form CSV file of training dataset

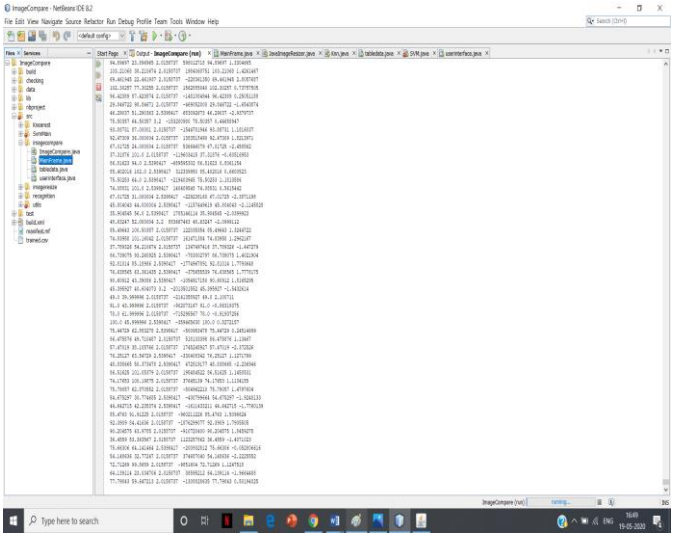


Fig-7: Values to form CSV file of testing dataset

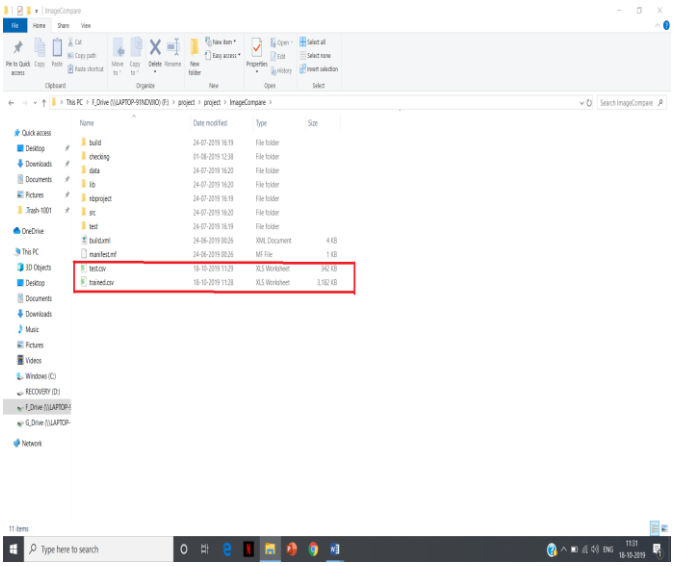


Fig-8: CSV file of train and test dataset

4.4. CLASSIFICATION

For classification we have used k-nearest neighbor method which in non-parametric. KNN classifies data based on the distance metric. In our system there are more than two categories of plant. Hence here KNN classifier will classify the output on the basis of feature matrix of input image obtained by training the data-set. The image data-set consist of 90 images. K-Nearest Neighbour is used to find the best candidate point for each SIFT point in query image among all the key-points descriptor from training image. A matrix which is created by training the model with data-set of 90 images will have the row wise value of each pixel of image in the data-set. The generated .CSV file of values obtained by training a model KNN to classify the input into specific category

5. RESULTS

In our system, system administrator can insert more images in data set and can retrain the system. Fig. 9 shows the UI for adding training dataset. Output of the training dataset is .CSV file gets created in root folder of project. Test data section can be used by both: user and system admin. System admin can simply use this for checking the accuracy of the model

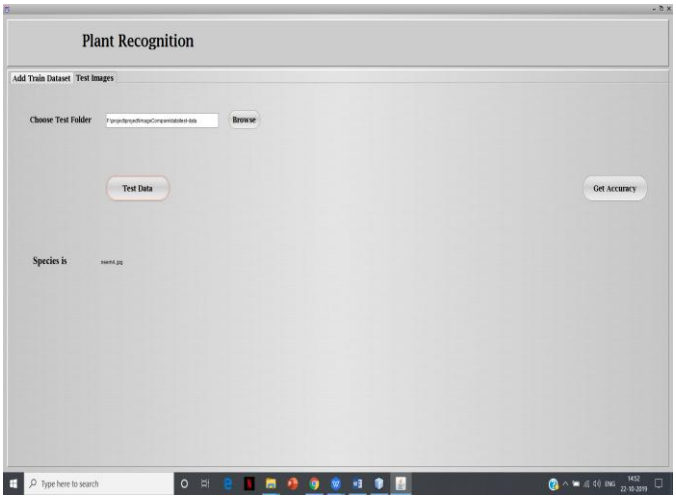


Fig-9: User interface

After this we have made one folder in our root folder as one image. In one image folder we have to copy the image of leaf from our database, Then SIFT algorithm will extract the local features and will calculate the values of those feature and with the help of K-NN classifier class of the leaf can be identified. System also have the option of getting the accuracy. In this we can get the output which shows the accuracy of each file of testing data comparing with trained data and shows how much accurately the classifier has classified the files.

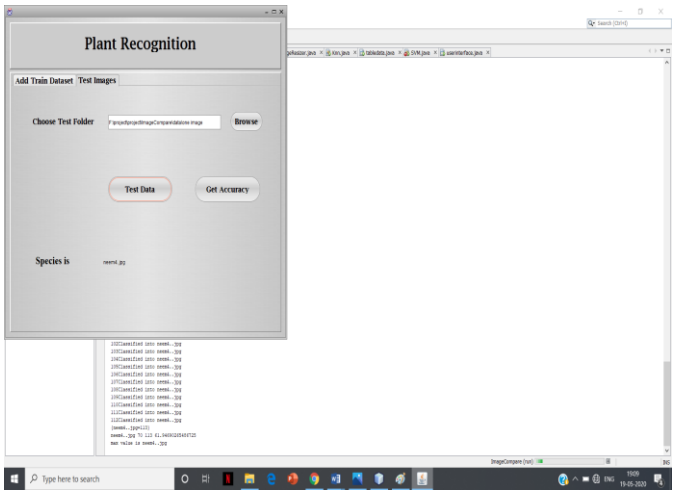


Fig-10: Identified species

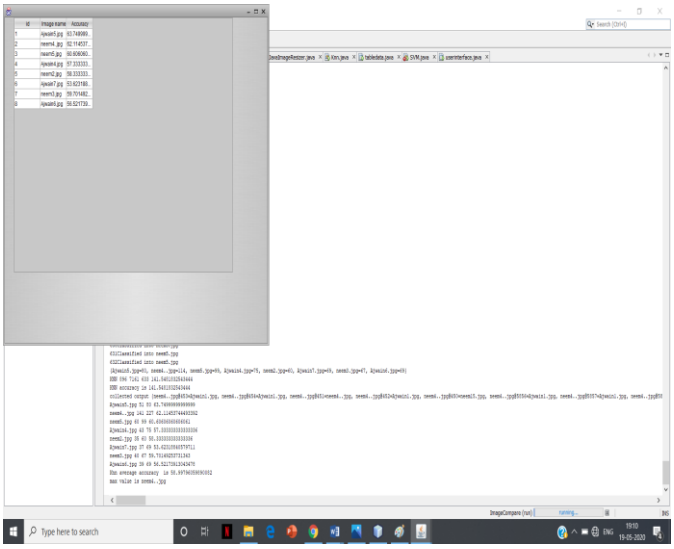


Fig -11: Accuracy of result

6. CONCLUSION

Using extracted features, KNN classifies the input data into its proper species using SIFT feature extraction technique. The accuracy of our model can be further increased by increasing the dataset size. Our system takes more time for computations and classification, this time can be decreased by training our model or else by modifying our method. The current ayurvedic plant identification system that we have developed is desktop application written in java using feature extraction and classification technique. The main motto of our project is to build the system for people living in rural areas who can use the application for the means of identification of medicinal plants.

REFERENCES

[1] Asrani K, Jain R (2013) Contour based retrieval for plant species..

[2] Voncarlos Araujo, Alceu S. Britto Jr, Andre L. Brun , Alessandro L. Koerich and Rosane Falate “Multiple Classifier System for Plant Leaf Recognition” October 2017.

[3] Dr. Pradeep Nijalingappa Madhumathi V. J. “Plant Identification System Using its Leaf Feature” 2015.

[4] Aparajita Sahay, Min Chen “Leaf Analysis for Plant Recognition ”2016.

[5] C. Mallah, J. Cope, and J. Orwell ,"Plant leaf classification using probabilistic integration of shape, texture and margin features," Signal Processing, Pattern Recognition and Applications, 2013

[6] S. Zhang, and K. W. Chau, "Dimension reduction using semi-supervised locally linear embedding for plant leaf classification, "Emerging Intelligent Computing Technology and Applications, Springer Berlin Heidelberg, 2009, pp. 948-955

[7] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Leaf classification using shape, color, and texture features , arXiv preprintarXiv:1401.4447, 2013

[8] A. H. M. Amin, and A. I. Khan, "One-shot Classification of 2-D Leaf Shapes Using Distributed Hierarchical Graph Neuron (DHGN)Scheme with k-NN Classifier," Procedia Computer Science, 24, 2013,pp.84-96

[9] J. Chaki, and R. Parekh, "Plant leaf recognition using shape based features and neural network classifiers," International Journal of Advanced