Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 76 - 81
Received 05 March 2021; Accepted 01 April 2021.

Plant Classification Using Leaf Images Processing

J. Vishalini[1*],T.N. Kirubalini[1], Dr. R. Subhashini[2*]
[1*,1] UG Student, Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
[2]Professor, Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
[3]Head, GIS and Remote Sensing, M.S. Swaminathan Research Foundation, Chennai
vishalij99@gmail.com, kirubalini14@gmail.com,
*subhaagopi@gmail.com,

Abstract. Plant exists everywhere. The plants are very much useful in balancing the environment. Usage of plants in the field of medical is wide developing. Rather than classifying plants based on molecules and biological methods,classification of plants based on leaf is found to be the first step. Our project aims at developing an image-based classification of plants with high accuracy. The features of the leaf such as, area, perimeter, width, height, aspect ratio, etc., are extracted, PCA is applied for dimensionality reduction and classified using SVM classifier.

Key words: Principal component analysis, Support Vector Machine, Feature extraction, plant classification.

1 INTRODUCTION

Most of the plants we observe are multicellular. They have different types of cells which help them to perform different functions[8,9]. These cells are not visible to our naked eyes. Currently, phylogenetic classification method is followed to classify them. Which mainly focuses on the evolutionary relationship of the plants[12,13]. At the earlier stage, plants where classified only based on their morphological features i.e. colour, shape of leaves, etc. Plants are majorly classified into two major groups: Vascular and Non-Vascular plants. Non-Vascular plants are otherwise called as Bryophytes. Vascular plants are also called as tracheophytes[10,11,18].

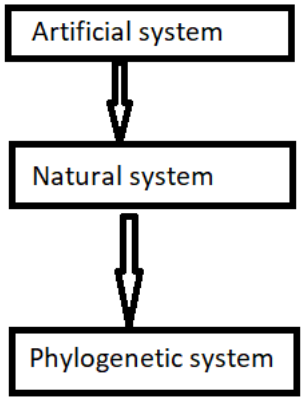

Fig. 1. Flow diagram of proposed scheme

Botanist use various methods to identify the plant species (mostly based on molecular

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 76 - 81
Received 05 March 2021; Accepted 01 April 2021.

observations). Comparative to the other methods, [6] plant classification using the leaf image was considered to be the initial and perfect solution for general public [14,15,19]. Sometimes, even the botanist themselves would find difficult to identify plants based their scientific approaches. However, determination of plant species was not the ultimate aim of a user other than biologist [21,22]. Their aim will be just to find the name of the plant or the plant's use without having knowledge about it[6,7].

The rich development of digital processing techniques and also the development of information technology such as portable devices, digital cameras, etc[23,24]. allow us to do more hand in hand operations[16,17]. In case of digital processing techniques, there are various algorithms that deal with Image processing, image analysis, image enhancements and mapping. These methods are helpful in various areas of applications such as: Medical field, Chemotherapy, industrial automation, etc.(7) Image based classifications serve as much convincing approach for species identification and classifications. The user can simply take a picture in either a camera or a mobile phone to identify its name[18-20]. Rather than taking a sample, applying preservatives and identifying their species name, this method will be specific and optimising solution[4,5].

Image acquisition- In this step, the image of either the part of the plant or the whole plant based on the need is collected. Here the images of a leaf are obtained and the data set is created[1-3].

2Image Pre-processing

Pre-processing is done in order to improve the quality of the image, reduce the distortions and obtain the relevant features of the image.[3] The various steps involved in image pre-processing includes:

a. Back-ground subtraction: The technique used in computer vision and image processing to retrieve the foreground image only without the background noises.

b. BGR to Grey scale conversion: for extracting the features such as texture, shape, edges, etc. Coloured images are harder to progress. Hence, they are converted to gray scale.

c. Low pass filtering: smoothing the image helps us to adjust the pixel values to the chosen cut-off value.

d. Thresholding: Various thresholding methods are applied. Otsu thresholding performed so as to close any holes present in the leaf.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 76 - 81
Received 05 March 2021; Accepted 01 April 2021.

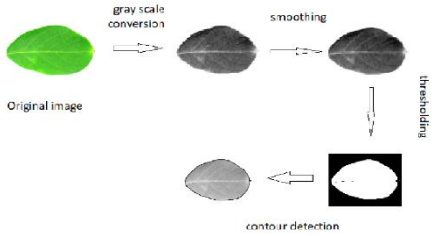

Fig. 2. Pre-processing

3Feature Extraction

Feature extraction mainly refers to the geometric measurements of the leaf image obtained. The three features such as colour, texture and shape-based features are extracted from the image [5].

a) Shape based features: Boundary extraction is done in order to extract the shape-based features such as aspect ratio, diameter of leaf, etc. Table 1 Shows features of shaped leaf

area	perimeter	physiological_length	physiological_width	aspect_ratio	rectangularity	circularity
197484	3479.036038	1416	759	1.865612648	5.442182658	61.28948044
101248	2490.381812	1190	130	9.153846154	1.527931416	61.25554646
86570.5	2290.683327	1095	119	9.201680672	1.505189412	60.612219
190214	2856.479353	1318	254	5.188976378	1.759975606	42.89628677
227727	2917.248904	1324	286	4.629370629	1.662797999	37.37080438

Table 1. Shape based features

b) Texture based features: Thetexture-based features such as contrast, correlation and entropy are calculated using Haralick moments. Texture based features is showed in Table 2.

contrast	correlation
12.63965607	0.997665773
8.137424083	0.997191135
8.553729373	0.996610015
8.440040958	0.998418772
8.641438458	0.998567725

Table 2. Texture based features

c) Colour based features: the mean and standard deviation of the RGB channels are considered as the colour-based features for extraction. Table 3 indicate the features of the colour.

mean_r	mean_g	mean_b	stddev_r	stddev_g	stddev_b
6.395667	13.64341	4.388007	24.02533	40.20093	21.44841
7.049316	9.232018	10.87607	33.81621	37.38222	46.47923
3.434303	6.371511	2.644757	19.9757	29.05737	19.27505
7.670415	13.3036	6.049157	28.82289	40.22185	26.9486
8.992028	16.67117	6.294281	30.96716	45.04019	28.59533

Table 3. Colour based features

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 76 - 81
Received 05 March 2021; Accepted 01 April 2021.

4Proposed Method

In this paper, the main aim of PCA is to reduce the dimensions of the dataset containing many variables [2].

(i)The first step in PCA corresponds to normalizing the data i.e. consider two dimensions X and Y, the data set whose mean is found to be zero is obtained.

(ii)Covariance matrix is calculated.

(iii)Calculation of eigen value and eigen vectors is performed.

(iv)The dimension is reduced by turning out the eigen vector corresponding to the highest eigen value.

$$\text{Feature vector} = (\text{eig1}, \text{eig2})$$

$$\text{Principle component} = (\text{feature vector})^{T} * (\text{Scaled data})^{T}$$

Support Vector Machine:

Support Vector Machine is used to classify the data. The goal of SVM is to create the clear line or decision boundary so that we can easily categorise the data elements. SVM is generally classified into two types: a. linear and b. non-linear. When the data relies into exactly two classes, it is classified using linear SVM. Those which cannot be classified by linear SVM are classified using non-linear SVM. (Fig:3)

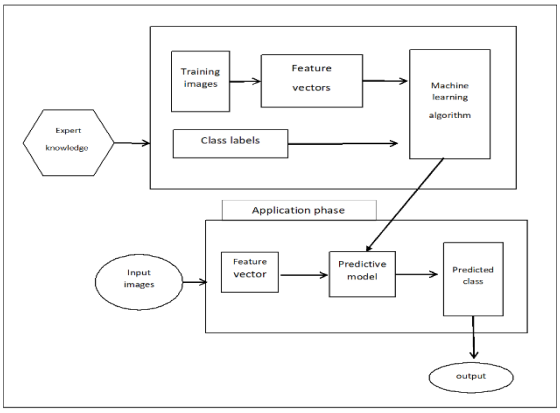

Fig.3. Flow diagram

5 Results and Discussion

Lower accuracy is achieved with shape-based classification. While the combination of shape and colour-based classification gave a high accuracy of 90%. 17 features from the leaf is extracted and over 1600 images are used to train and test the performance. Due to the simplicity of the frame work, we can apply more concepts to improve the accuracy.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 76 - 81
Received 05 March 2021; Accepted 01 April 2021.

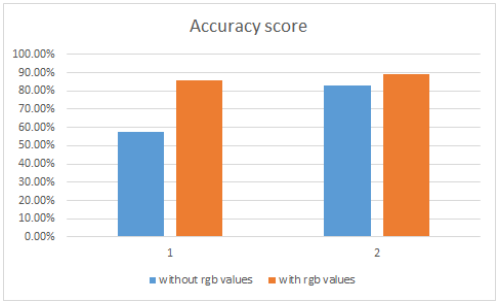

Fig. 4. Comparing accuracy scores of two different samples with and without considering rgb features.

6. Future Works & Conclusion

Our future works is to increase the number of species that is classified and also identifying the stress crop among the plant distribution over an area. This in turn helps to increase the productivity with lowering the expenditure.

The understanding of vegetation and classification of plants play a major role on describing the environment. This method helps in classification with better accuracy than other methods there by knowing the vegetation of a region. Our future work is in progress to achieve a better result.

References

1. Singh.R.S(2002),“PlantDiseaseManagement",Oxford and IBH,Newdelhi.
2. Sanjay b,patil et.al (2011)," leaf disease severity measurement using image proccesing",Intjourodeng and Tech.
3. kadir.A, L.E.Nugroho,A.Susanto,p.Insapsantosa (2011),"leaf classification using shape,color,and texture features",international journal of computer Trends and technology.
4. chaki.J,Parekh.R,Bhattacharya.s (2015),"plant leaf recognition using texture and shape fetures with neural classifiers".
5. Wu,G.S,Bao,F.s, et al(2007),"Recognition Algorithm for Plant Classification Using Probabilistic Neural Network".
6. T.J.Jassmann,R.Tashakkori,R.m.Parry (2015),"leaf classification utilizing a convolution Neural Network",southeast conference.
7. A.krizhevsky,I.sutskever,G.E.Hinton, (2012),"ImageNet Classification with Deep Convolutional Neural Network.
8. Shahina, K.,BevishJinila (2016), Y.,”An improved model for load balancing and dynamic channel allocation in cluster based manets”, ARPN Journal of Engineering and Applied Sciences, Vol. 11, Issue. 13,pp. 8094
9. Revathy, S., B. Parvathavarthini, and S. Shiny Caroline. "Decision Theory, an Unprecedented Validation Scheme for Rough-Fuzzy Clustering." International Journal on Artificial Intelligence Tools 25, no. 02 (2016): 1650003
10. V.VijeyaKaveriansS.Natarajan (2017), ” An Efficient Rekeying Framework For Group Key Management With Video Streaming (2017) ”, Jour of Adv Research in Dynamical & Control Systems, 06-Special Issue, July 2017,PP.No.32-41.
11. Vamsi, Manikanta, Anandhi, T.(2019),Survey on enhancing drainage maintenance system

using IOT,International Journal of Recent Technology and Engineering 7(5), pp. 130-135.
12. Jenifer, M., & Bharathi, B. (2016, March). A method of reducing the skew in reducer phase—Block chain algorithm. In 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT) (pp. 1-4). IEEE.
13. "Mohana Prasad K, Sabitha R, Oviya (2016), 'An Efficient Clustering Formulation From Resemblance In Extant Algorithms', ARPN Journalof Engineering and Applied Sciences, Vol. 11, No.13, ISSN 1819- 6608 , 8278 – 8283."
14. Subhashini, D. R., Sethuraman, R., & Milani, V. Reinforcing Telemedicine Through an Interactive Voice Response Service for Rural Indians. International Journal of Engineering and Technology, ISSN, 0975-4024.
15. Kumar, V., Vasudevan, S., &Posonia, M. (2006). URBAN MODE OF DISPATCHING STUDENTS FROM HOSTEL.
16. Hema Prasanna, K., MurariDevakannan, K.(2015),Performance of support vector machine in predicting the relative risk of diabetes mellitus with the help of association rule mining,International Journal of Applied Engineering Research,10(2), pp. 2257-2264.
17. Franklin, R.G.(2015),Prevention of XML based dos attacks for a secure web service,Global Journal of Pure and Applied Mathematics,11(5), pp. 2889-2896."
18. Nagarajan, G., & Minu, R. I. (2018). Wireless soil monitoring sensor for sprinkler irrigation automation system. Wireless Personal Communications, 98(2), 1835-1851.
19. Nagarajan, G., Minu, R. I., & Devi, A. J. (2020). Optimal Nonparametric Bayesian Model-Based Multimodal BoVW Creation Using Multilayer pLSA. Circuits, Systems, and Signal Processing, 39(2), 1123-1132.
20. Jesudoss, A., Vybhavi, R., & Anusha, B. (2019, April). Design of Smart Helmet for Accident Avoidance. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0774-0778). IEEE.
21. Sheela, A. S., & KUMAR, C. (2014). DUPLICATE WEB PAGES DETECTION WITH THE SUPPORT OF 2D TABLE APPROACH. Journal of Theoretical & Applied Information Technology, 67(1).
22. Prince Mary, S., Usha Nandini, D., Ankayarkanni, B., & Sathyabama Krishna, R. (2019). Big Data Deployment for an Efficient Resource Prerequisite Job. Journal of Computational and Theoretical Nanoscience, 16(8), 3211-3215.
23. Nagarajan, G., and K. K. Thyagharajan. "A machine learning technique for semantic search engine." Procedia engineering 38 (2012): 2164-2171.
24. Nandini, D. U., &Divya, S. (2017, January). A literature survey on various watermarking techniques. In 2017 International Conference on Inventive Systems and Control (ICISC) (pp. 1-4). IEEE.
25. Aravind, K.R.N.V.V.D., PraylaShyry, S., Felix, Y.(2019),Classification of healthy and rot leaves of apple using gradient boosting and support vector classifier,International Journal of Innovative Technology and Exploring Engineering,8(12), pp. 2868-2872.