Classification of plant based on leaf images

Mayank Arya Chandra[1], S S Bedi[2]

[1,2]MJP Rohilkhand University
Bareilly, UP, India
[1]machandra100@gmail.com
[2]erbedi@yahoo.com

Abstract-- Plant identification has been important and complex task. Leaves are key part of the plant that distinguish, characterize and classify. Every plant has unique leaves and each leaf has a set of features that differentiate it from the other, inspire these features of plant leaves, here, this paper utilizes the idea of plant leaves for plant classification. Leaves are unique in relation to one another by attributes; for example, shape, shading, surface and different other characteristics. This paper employ, an algorithm for plant classification is proposed based on plants leaves image features data through Linear Norms Decision Directed Acyclic Graph Least Square Twin Support Vector Machine (LN-DDAG-LSTSVM) classifiers. This proposed algorithm is demonstrate on leaf images from standard benchmarks database and compared with other methods where experimental results deliver higher accuracy.

Keywords— Plant, Leaf, Classification, Linear norm, Recognition, Least Square, SVM, Twin SVM, DDAG.

1. INTRODUCTION

Plants are the backbone of life on earth and give oxygen, nourishment, medicine, fuel and considerably more. Plants are the major source of oxygen, it is key factor of ecosystems, can decrease carbon dioxide by the process of photosynthesis [1]. Apart from this, plant a play vital role in production of food. Plant uses the energy of sunlight with water and carbon dioxide gas to make a food [2]. Today the thousands of plant species are in danger of extinction or endangered caused by different anthropogenic activities. Thus, conserving of plant biodiversity, classification of plants is followed and the newly discovered species are distinguished and arranged. According to news report of nature journal approximately 14% land species named and classified rest 86% remain undiscovered [3]. Scientist and botanist have named and classified millions of plants. It is not an easy task for any botanist to remember in excess of a minor division of the aggregate number of named plant species. The identification of the recently discovered species requires a specialist who has vast knowledge about it. The conventional techniques require significant amount of effort and time. So, it is demand of time to build up a system of automatic recognition of plant species.

Plant species identification and classification based on plant leaves is the fastest and facile approach to recognize a plant [4]. There are a few reasons why leafs are normally selected for classifying the plant species. Leaf are easily found and collected in any seasons. Leaf is 2-D shapes as compare to other parts of plant species (as blooms, seeds and so on.)[5]. A lots of plants have remarkable leaves that are unique in relation to one another dependent on various attributes, eg. shape, color, texture, and the margin [6-7]. Some researchers utilize more features for more appropriate strategies [8-9]. Shape features are deal with geometrical structure of leaf (aspect ratio, compactness, roundness, etc…) [10]. Color features are utilized by a few schemes [11] because most of the plant

species have comparative leaf colors property, this highlights is not of key consequentiality. Textures features can be utilized to describe overall portray structure of the leaf or venation [12].

SVM is a successful mathematical supervised learning tool to resolve the classification and regression problems [1]. In the last decades, various modifications emerged in SVM like Lagrangian SVM [14], Least square SVM [15, 16], Proximal SVM [17]. Twin SVM (TWSVM) proposed by Jayadeva [18]. Twin SVM is mathematical modification of traditional SVM and it solves pair of relative reduced QPP rather than large one [19]. TSVM computational cost is approximately four times faster than original SVM. A new version of twin SVM, known as least square Twin SVM was proposed in which QPP problem solve by equality constraints rather than inequality constraints of TSVM [20]. LS-TSVM has better performance than TSVM. Current scenario LS-TSVM was also modified by different researcher in ordered to achieve the better generalization capability and computational performance [21-24]. Motivation of their improved classification accuracy in classification domain, here, this paper utilizes the idea of least square twin support vector machine for classifying the different plant species. Here we extend the LSTSVM by using decision directed acyclic graph with linear Norm concept. It is analyzed that Linear Norm Decision Directed Graph Acyclic Graph Least Square Twin Support Vector Machine (LN-DDAG-LSTSVM) has present the better execution result with good prediction accuracy and running time complexity. This paper propose a scheme that recognize the different plants through their leaf images and their characteristics like aspect ratio, eccentricity, Elongation Maximal Indentation Depth etc. and texture feature like Intensity, moment, Contrast, Entropy etc. The scheme LN-DDAG-LSTSVM classifier for plant leaf recognition system is tested and compared to other machine learning approaches.

This paper is formatted as follows. The next section deals with review of previous works. Section 3 applied methodology and formulation of LN-DDAG-LSTSVM. In the fourth section, experiment results and their comparison. Finally conclude the paper at the end.

2. PREVIOUS WORKS

Plant recognition and classification are depending on images of leaf. Plant leaf images contain visual data and their visual qualities like its shape, its color, its surface structure, its vein, etc., can be acclimated to describe plant species. C. Im et al. design a system to identifying the plant using leaf shape features [25]. Wang use the concept of fuzzy on different leaf shape features [26]. Concept of hyper-sphere classifiers introduce to recognition leaf image based on shape features [27]. A new method is proposed for leaf image that have complex background where a special feature Hu geometric and Zernike moments calculated from segmented binary images. Moving center hyper-sphere classifier is used to classification leaf images [28]. The concept of histogram of oriented gradients is used to identify the shape feature of leaf image and a new dimensional reduction technique are utilize in this paper, known as Maximum Margin Criterion (MMC). HOG play a crucial role with MMC in order to achieve a better performance on leaf classification [29]. Morphological features and Zernike moments are best choice for recognize and categorize the plant leaves. The process of feature extraction is not depends upon the leaf growth as well geometric image transformation. These all terms build up an enhanced methodology that creates the best classification approach [30].

A new scheme has been proposed for identifying and characterizing plant leaves using texture and shape features. Texture descriptor of the leaf is calculated by Gabor filter and

Gray Level Co-occurrence Matrix (GLCM) while shape descriptor is determined by curvelet transform with Invariant Moments [31]. A system application known as LeaVes is buildup for plant classification dependent on the leaf's shape and venation. In this paper the different types of machine learning procedures and different types of image filter are discussed like as canny, moment invariance centroid and ANN [32]. A new method which automatically identifies the different plant leaves has been proposed in which texture features are used for identifying and recognize the plant leaf [33]. An approach which is used the different color, shape, texture and vein descriptor for the purpose of classification and recognizing plant leaf images [34-35]. The concept of Convolutional Neural Network are used to design a new algorithm known as AlexNet, leaf in which different feature was extracted by different CNN based models [35].

Since plants are the major source of oxygen and its ability to absorb carbon forms around 66%. Without plants it is difficult to think of existence of human life on the earth. So Classifying plants helps at ascertaining the protection survival of all regular life. Plants can be classified on the basis of different parameters. This paper proposes a new scheme for plant classification based on leaf image.

3. PROPOSED METHOD

This proposed plant leaf classification scheme comprises of 4 principles as appeared in fig. 1. First are image acquisition, preprocessing, segmentation, feature extraction and classification.

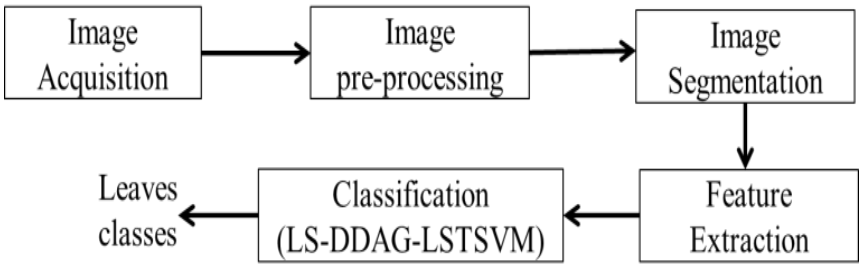

Fig.1 Proposed Architecture of plant classification approach

Image Acquisition: In this paper database of plant leaf image are used. Total 40 leaf sample images from four different plant species selected from open source database [36]. Overall description of each plant with scientific name and total number of leaf samples are presented in table 1 [36].

Table 1. Description Selected Plant Leaf images

Class	Scientific Name	Instance
1	Quercus suber	12
2	Salix Atrocinera	10
3	Populus Nigra	10
4	Alnus Sp	8
5.	Quercus robur	12

Image preprocessing: Image pre-processing is a preliminary step to reduce noise and upgrading the appearance of the leaf images which increase the overall performance of leaf identification. RGB images are converted into black and white (BW) images.

Image Segmentation: The image segmentation is a method of partitioning a leaf images into region of interest (ROI). Here color slicing is used with Otsu's algorithm.

Feature Extraction: In this paper public image leaf data base are used and here we select only 4 plant species with all features.

Classification: we proposed a technique refer as linear norm decision directed acyclic graph least square twin support vector machine (LN-DDAG-LSTSVM) for classification and identification of plant leaf.

CLASSIFICATION ALGORITHM:

We proposed a technique refer as linear norm decision directed acyclic graph least square Twin Support Vector Machine (LN-DDAG-LSTSVM). In decision DAG, it has been observed that upper nodes of decision DAG have more effective over the performance of classifier. Therefore build a predicted decision DAG in such a way that each node should be classify most separate group of classes from the other one, i.e., split selection strategy at the root node is lowest error estimation between the two classes. DDAG always start to compare between first and the last class on the basis of decision function values. Repeat this process until remain one class which have minimum criteria value. DDAG is generalizing the class of decision trees which allow repetition that occurs in different branch of tree.

These data values normalized at the same time by using linear norm as given formula Eq. (1).

$$\mathrm{x} = (\mathrm{a} - \mathrm{e} * \min)/(\mathrm{e} * (\max - \min)) \tag{1}$$

Where a data vector, min is data vector minimum values, max is data vector maximum values and e is unit vector.

1. Let's input data set $I = \{(x_1, y_1)(x_2, y_2) \ldots (x_n, y_n)\}$.

2. Compute the partition of the respective class Vs rest at each node by using LS-TSVM as:

   a) Let's two matrixes A_i & B_j and Select penalty parameter c_1, c_2
   Solve given equation (2) and (3)

$$\min_{w_1 b_1 \xi_1} \quad \frac{1}{2}(Aw_1 + eb_1)^T(Aw_1 + eb_1) + \frac{c_1}{2}{\xi_1}^T\xi_1$$

$$Subject \qquad -(Bw_1 + eb_1) + \xi_1 = e, \quad \xi_1 \geq 0 \tag{2}$$

$$\min_{w_2 b_2 \xi_2} \quad \frac{1}{2}(Aw_2 + eb_2)^T(Aw_2 + eb_2) + \frac{c_2}{2}{\xi_2}^T\xi_2$$

$$Subject \qquad -(Bw_2 + eb_2) + \xi_2 = e, \ \xi_2 \geq 0 \tag{3}$$

   b) Calculate hyper-plane parameters according to equations (2) and (3)

$$\begin{bmatrix} w_1 \\ b_1 \end{bmatrix} = -\left(B^T B + \frac{1}{c_1} A^T A\right)^{-1} B^T e \quad (4)$$

$$\begin{bmatrix} w_2 \\ b_2 \end{bmatrix} = -\left(A^T A + \frac{1}{c_2} B^T B\right)^{-1} A^T e \quad (5)$$

3. Repeat Step 2, until all nodes have been trained by LS-TSVM.
4. Save the decision DAG and return.

Decision of the class 'k' for each new data sample decided by the given decision function based on respective perpendicular distance from the two planes

$$\text{Class } k = \min|x^T w_k + b_k| \qquad \text{for } k = 1, 2 \quad (6)$$

4. EXPERIMENT RESULT

For validating and examination of proposed method we select the 40 leaf sample images from 4 different plant species, as table no.1, taken from open source data base provided leave recognition system [36]. The plant classification system implemented under the environment of MATLAB R2013a on window 10 with Intel Core i3 1.7GHz, RAM 4GB. The optimal values for penalty parameters are selected from the range $V_i \in (10^{-8}, 10^{-7} \ldots 10^{8})$.

Table 2. Quantitative measure for performance of different classifiers.

Class	NN Acc(%)	SVM Acc(%)	DAG-SVM Acc(%)	Proposed Method Acc(%) (Running time (s))
2	70	92.2	93.82	95 (0.000369)
3	76.48	89.7	91.59	94.16 (0.000739)
4	83.64	78.61	81.27	85.50 (0.001672)
5	82.89	76.47	78.93	85.71(0.002408)
Avg	78.2525	84.245	86.4025	90.0925 (0.001297)

The proposed algorithm experimentally tested on binary class, 3 class and 4 class problem. The performance in terms of accuracy has been compare with Neural Network (NN), SVM, and DAG-SVM. From table 2 and fig. 2, the fact can be extracted that proposed algorithm yields an average accuracy more than 90%. The quantitative comparison of the results shows that the proposed method has greater classification accuracy than existing systems.

Equation for calculating the accuracy is defined as

$$\text{Accuracy} = \frac{TP+TN}{FP+FN+TP+TN} \quad (7)$$

Where TP, TN, FP, FN are number of 'true positive', 'true negative', 'false positive', 'false negative' respectively.

Fig.2. Accuracy comparison chart of proposed system and other classifiers.

5. CONCLUSION

In this paper, we have proposed a LN-DDAG-LSTSVM method for classifying the plant leaf based on combining leaf shape and texture feature. The proposed plant leaf classification method is robust and computationally efficient, which takes into consideration 16 shapes and texture feature leaf. LN-DDAG-LSTSVM classifier trained with these 16 leaf features. The proposed method was tested and evaluated on open source leaf image database.

Performance comparison chart display in fig.2 shows that LN-DDAG-LSTSVM is improved classification accuracy. Experimental results demonstrate that the proposed LN-DDAG-LSTSVM classifier based plant recognition system achieves the best computation efficiency as compared to ANN, SVM, DAG-SVM. As the number of leaves relating to each plant is small, the computational accuracy of the proposed method could be upgraded with large amount of leaves. This technique is very useful tackling difficulties in traditional SVM, TSVM and other classifiers. It is interesting to extract different features in future to improve the classification accuracy.

REFERENCES

1. Govindjee .: A sixty-year tryst with photosynthesis and related processes: an informal personal perspective. Photosynthesis Research, PP 1-29, (2018).https://doi.org/10.1007/s11120-018-0590-0.
2. Beghin, T., Cope, J. S., Remagnino, P. and Barman, S.: Shape and Texture Based Plant Leaf Classification. International Conference on Advanced Concepts for Intelligent Vision Systems (ACVIS), Sydney, Australia, December 13-16, pp. 345-353, 2010, Springer.
3. Sweetlove, L.: Number of species on Earth tagged at 8.7 million. Published online 23 August 2011, Nature, doi:10.1038/news.2011.498.
4. Man, Q. K., Zheng, C.H., Wang, X.F. and Lin, F. Y.: Recognition of plant leaves using support vector machine. Advanced Intelligent Computing Theories and Applications with Aspects of Contemporary Intelligent Computing Techniques (Springer, 2008), pp. 192–199.

5. Bhardwaj, A., Kaur, M., and Kumar, A., : Recognition of plants by leaf image using moment invariant and texture analysis. International Journal of Innovation and Applied Studies, vol. 3, no. 1, pp. 237–248, 2013.
6. Kadir, A., Nugroho, L. E., Susanto, A., and Santosa, P. I. : Leaf classification using shape, color, and texture features. International Journal of Engineering Trends and Technology, 2(1), pp. 225–230, 2011.
7. Du, J. X., Wang, X. F. and Zhang, G. J.: Leaf shape based plant species recognition. Applied mathematics and computation, vol. 185, no. 2, pp. 883–893, 2007.
8. Novotn`y, P., and Suk, T.: Leaf recognition of woody species in central Europe. Biosystems Engineering, vol. 115, no. 4, pp. 444–452, 2013.
9. Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F. and Xiang, Q.-L.: A leaf recognition algorithm for plant classification using probabilistic neural network. IEEE international symposium on signal processing and information technology. Giza, IEEE, 2007, pp.11-16.
10. Chaki, J., Parekh, R., Bhattacharya., S.: Plant Leaf Recognition using Texture and Shape features with Neural Classifiers. Pattern Recognition Letters. vol. 58 issue C. Pages 61-68, 2015.
11. Caglayan, A., Oguzhan,G., Can, A. B.: A plant recognition approach using shape and color features in leaf images. ICIAP 2013, Lecture Notes in Computer Science, vol. 8157 (2013) 161-170.
12. Sabu, A. and Sreekumar, K.: Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach. 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2017, pp. 145-149, 2017.
13. Cortes, C., and Vapnik V.: Support vector network. Machine Learning, vol. 20, no. 3, pp 273–297, 1995.
14. Mangasarian O. L. and Musicant, D.R. : Lagrangian Support Vector Machines. Journal of Machine Learning Research 1, pp. 161-177, 2001.
15. Suykens, J. A. K. and Vandewalle , J. :Least squares support vector machine classifiers. Neural Process, Lett. 9, pp. 293-300, 1999.
16. Suykens, J. A. , Gestel, T.V. , Brabanter, J. D., Moor B.D. and Vandewalle, J.: Least Squares Support Vector Machines. World Scientific publishing Co., Singapore 2002.
17. Fung G. and Mangasarian O.L.: Proximal Support Vector Machine Classifiers. KDD 01 (2001) San Francisco CA USA, ACM.
18. Jayadeva, Khemchandani, R., Chandra, S.: Twin support vector machines for pattern classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29 no.5, pp. 905–910 ,2007.
19. Chandra, M.A. & Bedi, S. S.: Survey on SVM and their application in image classification. International Journal of Information Technology pp 1–11 (2018). https://doi.org/10.1007/s41870-017-0080-1.
20. Kumar, M. A., Gopal, M.: Least squares twin support vector machines for pattern classification. Expert Systems with Applications Vol. 36, Issue 4,pp. 7535–7543, 2009.
21. Xu, Y., Pan, X., Zhou, Z. et al.: Structural least square twin support vector machine for classification. Applied Intelligence, Volume 42, Issue 3, pp 527–536, April 2015.
22. Mozafari, K., Nasiri, J. A., Charkari, N. M. and Jalili, S: Hierarchical Least Square Twin Support Vector Machines Based Framework for Human Action Recognition," 2011 7th Iranian Conference on Machine Vision and Image Processing, pp. 1-5, Tehran, 2011.
23. Richhariya,B., Tanveer, M.: A robust fuzzy least squares twin support vector machine for class imbalance learning. Applied Soft Computing, Volume 71, Pages 418-432, October 2018.
24. Chandra, M. A. and Bedi. S. S.: Benchmarking Tree based Least Squares Twin Support Vector Machine Classifiers. International Journal of Business Intelligence and Data Mining. DOI:10.1504/IJBIDM.2018.10009883.

25. Im, C., Nishida, H. and Kunii, T. L. : Recognizing plant species by leaf shapes-a case study of the Acer family. Proceedings. Fourteenth International Conference on Pattern Recognition, pp. 1171-1173 vol.2, Brisbane, Queensland, Australia, 1998.
26. Wang, Z., Chi, Z., and Feng, D.: Fuzzy integral for leaf image retrieval. 002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02, Honolulu, HI, USA vol. 1, pp. 372-377 , USA, 2002,.
27. Wang, X.F., Du, J.X., Zhang, G.J.: Recognition of Leaf Images Based on Shape Features Using a Hypersphere Classifier. Advances in Intelligent Computing. vol 3644. pp 87-96 ICIC 2005, Springer.
28. Wang, X.F., Huang, D.S., Dua, J.X., Xu, H., Heutte, L.: Classification of plant leaf images with complicated background. Applied Mathematics and Computation 205 (2008) 916–926.
29. Xiao, X.Y. Hu, R., Zhang, S.W., Wang, X.F.: HOG-Based Approach for Leaf Classification. Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence, ICIC 2010, pp 149-155 Springer.
30. Harish, B. S., Hedge, A., Venkatesh, O., Spoorthy D. G. and Sushma, D.: Classification of plant leaves using Morphological features and Zernike moments. 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, 2013, pp. 1827-1831.
31. Chaki,J. Parekh, R, Bhattacharya, S.: Plant Leaf Recognition using Texture and Shape features with Neural Classifiers. Pattern Recognition Letters, Vol. 58, Issue C pp. 61-68, 2015.
32. Codizar, A. L. and Solano, G.: Plant leaf recognition by venation and shape using artificial neural networks. 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA), Chalkidiki, 2016, pp. 1-4.
33. Elnemr, H. A. : Feature selection for texture-based plant leaves classification. 2017 Intl Conf on Advanced Control Circuits Systems (ACCS & 2017 Intl Conf on New Paradigms in Electronics & Information Technology (PEIT), Alexandria, 2017, pp. 91-97.
34. Mittal, P., Kansal, M. and Jhajj, H. K. : Combined Classifier for Plant Classification and Identification from Leaf Image Based on Visual Attributes. 2018 International Conference on Intelligent Circuits and Systems (ICICS), Phagwara, 2018, pp. 184-187.
35. Tomar, D., Agarwal, S.: Leaf Recognition for Plant Classification Using Direct Acyclic Graph Based Multi-Class Least Squares Twin Support Vector Machine. International Journal of Image and Graphics Vol. 16, No. 3, 1650012 (2016)
36. Tan, J. W., Chang, S., Binti S., Kareem A., Yap, H. J. and Yong, K.: Deep Learning for Plant Species Classification using Leaf Vein Morphometric. in IEEE/ACM Transactions on Computational Biology and Bioinformatics.
37. Silva, P.F.B., Marcal, A.R.S., de Silva R.M.A.: Evaluation of Features for Leaf Discrimination. Springer Lecture Notes in Computer Science, Vol.7950, 197-204, 2013.