

Domain Specific Knowledge based Machine Learning for Flower Classification using Soft Computing

Rajesh S. Sarkate
Dept. of CS & IT
MGM's College of CS & IT,
Nanded

Prakash B. Khanale, PhD
Dept. of Computer Science
DSM College
Parbhani

ABSTRACT

Machine learning is ability of the machine to capture the data, analyze it and make decision as like human being perform in everyday life. Sometimes it also referred as pattern recognition or classification. With use of latest computing technology and soft computing, it is now possible for machines to act with intelligence as human. Various industries like automobile, medical diagnosis are using machines for fast and accurate data capturing and analyzing.

Our study is conducted to justify these machine abilities in the field of floriculture. A knowledge base in flower domain is used for the Intra-class sorting purpose. The flower sorter is designed to capture the flower images and with artificial neural network classifier, the images are classified in four flower color classes. In the study, supervised learning algorithm is used for machine learning. The classification accuracy of flower sorter is found to be 98%.

General Terms

Pattern Recognition, Horticulture, Neural Network.

Keywords

Machine Learning; Classification; Flower; Sorting; ANN; Knowledge base.

1. INTRODUCTION

There was a time when experts had doubts in human being replaced by machines as the capabilities and performance especially in the sagacity of machines were debatable. Nevertheless, studies conducted in this domain have proved not only the capabilities but also the perfection of machines over human limitations.

In the current study has investigated an approach of machine learning, specifically, with the superior computational ability of segmentation and for the judgment of the color wise discrimination in flowers. This is possible with superior computer storage and computation for extracting statistical features and for locating objects in a high-dimensional feature space. In past, flowers from different varieties were classified by researchers based on color, texture and shape as features. The present study concentrates of Intra-class separation of single flower specie i.e. *Gerbera jamesonii* with limited feature i.e. color as shape and texture remain same. It is possible that in the domain of flower recognition, current study could help in commercial development of human unaided machine automated flower sorter system. The knowledge base that uses priory information of specific domain guides to interpret and learn the machine about the objects of interest i.e. flowers. In addition, the proposed

method can be used in applications that require human free interface to other agriculture products.

Computer vision technology is widely used in Agriculture[1] for pest or disease detection[2][3][4], lesion estimation[5], grading[6][7]. In this flower recognition system, a domain-specific knowledge is base for machine learning. The key to efficient classification is the flower colors that allow the, like human, machine to use the features throughout the classification process. [8]

Gerbera (*Gerbera jamesonii*) is top in the list of most popular commercial cut flowers in the world and according to the global trends in Floriculture; it comes in top ten among cut flowers[9]. Throughout the year, *Gerbera* flower are in demand in both domestic and export markets due to attractive blooms that are suitable for any type of floral arrangements and are available in different shades and hues. The cut flowers have long vase-life reasoning premium market prices. Due to changes in social and cultural life style of people, cut flowers have found an important place in various social functions and daily activities.

In many plants, flowering is conspicuous in the field. Flowers are displayed prominently at the top of the canopy as in crops such as *Gerbera* and they appear perpendicular to the stalk. While flowering is an important developmental stage in most crops, enumerating flowers is labor intensive, especially when flowers need to be counted on a daily basis. Human intervention in counting process may lead to the physical contact with the plants resulting in damage to plants. The development of low-cost color digital cameras that use charge coupled device (CCD) arrays to capture images offer a potential method for measuring flowering in plants with appropriate architecture and flower color.[10]

2. MATERIALS AND TOOLS

The different flowers samples used in the present work are collected from main flower market in Nanded district of Maharashtra state in India for the production year 2013. The images are acquired with a color Digital Camera, and then transferred to computer system with core2deo @2.4 GHz. The camera is mounted on a stand with a facility for vertical movement to fine tune the orthogonal distance of the camera from the flower samples in a properly illuminated chamber. The images are illuminated with light source of 40W, 230 V fit to the test table at an angle of 45° from the camera. Camera is calibrated to get petal part that decides the flower color.

A database of 50 flowers from 4 classes with the help of digital camera is developed for the study. The pictures are deliberately taken with different resolutions i.e. 320 x 240 pixels, 1600 x 1200 and 2592 x 1944, to check resolution

factor on classification. There are four color classes of flowers used for the study; Orange, yellow, red and pink. Humans have natural ability to recognize such variety of flowers, but it is thorny task for machine to classify them accordingly. The view from the camera is represented with R, G, and B digital values for each pixel in the digital image and stored in joint photographic experts group (JPEG) file format. Images are then transferred from the camera to a personal computer as JPEG files.

3. METHODOLOGY

After capturing the flower images, the pre-processing and post processing on the images are carried out in Matlab 7.0 with the help of image processing toolbox. Considering the memory and processing requirement, all the images are resized to 500 x 500 pixels. For flower discrimination purpose, the neural network classifier is designed using neural network toolbox provided by Matlab.

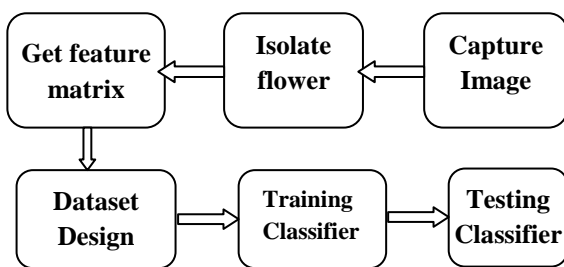
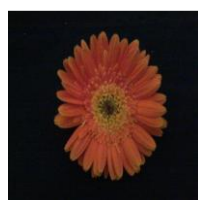


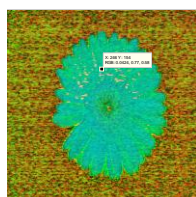
Figure 1: System Flowchart

3.1 Feature Extraction

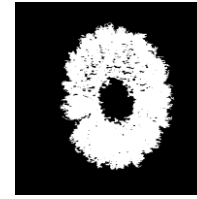
Features are the attributes of different information present in the image. Features are extracted by applying chain of image processing operations like resizing, filtering[11]. For further processing, the images are transformed from RGB color space to HSV color space where every pixel is represented with three components i.e. Hue, Saturation and Value. The HSV color space is considered to be closer to the human eye perception. Hue, representing the color of pixel in the image, Saturation, and value are counted between 0 -1 for every pixel[12]. Thus, this process results in three 500 x 500 matrix with values ranging 0-1 providing information present in image. This information is analyzed and used, by the machines, in learning and classifying the flowers



a.



b.



c.

a. RGB flower b. HSV image c. Flower region (Binary)

Figure 2 : Classifier training set generation

Color space transformation and threshold segmentation isolate the flower region from the background (figure 2). As most of the pixels representing flower are segmented, the accuracy of the segmentation does not affect the learning process. After performing morphological operations on region of interest, the pixel matrix representing the flower region is derived. Mapping of these pixels are done from binary image to the HSV image and Hue, Saturation and Value components are extracted and finally stored in three matrixes H, S and V.

3.2 Classifier Design

Artificial Neural Networks are universal estimators to map the input space values to the desired output space. ANN is massively parallel-distributed processor that has inbuilt ability to store trained knowledge in the form of weights in the interneuron connections and use it for testing phase to generate the desired output. This work has used supervised learning with back propagation neural network. For pattern classification, Back propagation neural networks architecture is most commonly used neural networks [13][14] [15]. Another reason to chose back propagation due to its ability to perform pattern classification on data where the input and the output had no linear relationship, as in the case of this application [16].

In ANN, the number of inputs of the network is equal to the number of features used for classification, whereas the number of outputs is equal to the number of classes to be separated[17]. The network architecture is depicted in figure 2. The input layer is made up of three neurons used to provide each feature vector of H, S and V components. The hidden layer is set, after various passes of training, to 30 neurons with log sigmoid activation function. The output layer, with four log sigmoid neurons, generates the values in 0 to 1.

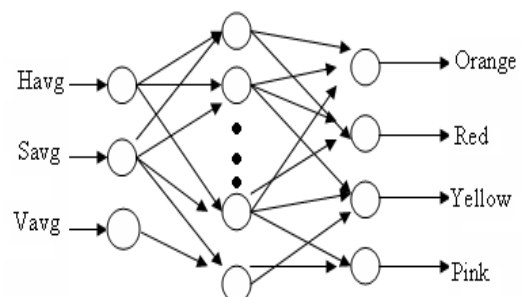


Figure 3 : Classifier design

3.3 Training dataset

The feature extraction phase generates the three matrixes H, S and V for the pixels that represent the flower. To reduce the input dimension for training, we have deliberated mean values for each matrix H, S and V. In the field of computation, mean is considered to be most popular and useful when comparing sets of data. So now every flower is featured by only three values in lieu of thousands of individual pixel values. The output vector is designed for four preset neurons such that value around 1 is expected to corresponding output neuron of presented flower and rest of the neurons will generate values around 0.

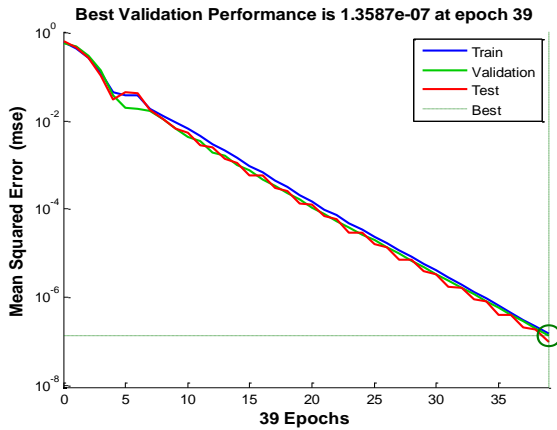


Figure 4 : Network Performance

Network is trained with input vectors with their known class identifier as the desired output vector with error back propagation. The training process continuously updates interneuron weights and biases until the mean- squared error converges to a desired threshold.

All Confusion Matrix

	1	2	3	4	
1	20 25.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
2	0 0.0%	20 25.0%	0 0.0%	0 0.0%	100% 0.0%
3	0 0.0%	0 0.0%	20 25.0%	0 0.0%	100% 0.0%
4	0 0.0%	0 0.0%	0 0.0%	20 25.0%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%
	1	2	3	4	

Target Class

Figure 5 : Confusion Matrix

4. RESULT AND DISCUSSION

The flower sorter system works well for the given domain and knowledgebase. The results of the system are compared to manual classification. The algorithm accuracy is given in the table 1.

Table 1: Accuracy of Proposed algorithm for sorting

No	Flower Color	Total Sample	Manual Sorting	System Sorting	Accuracy %
1	Orange	15	15	14	93.33
2	Red	15	15	15	100.00
3	Pink	10	10	10	100.00
4	Yellow	10	10	10	100.00
Average success %					98.33

As shown in table, proposed algorithm has given accuracy performance of 98.33% with 1.77 % lacuna. Result may get affected possibly due to the luminance or uncorrected noise present in the image. Fig 4 shows the Matlab GUI for input images, extracted feature set and the result of the ANN classifier.

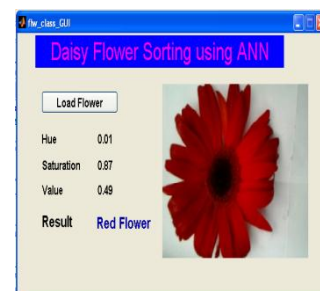


Figure 6 : GUI for Flower Sorter System

5. CONCLUSION AND FUTURE WORK

Machine learning has advanced in many areas of day to day life like face recognition, biometric etc. This study has shown another application of machine learning for sorting flowers that require subtle handling of the material. The algorithm can be implemented in commercial flower industry for separation and packaging without human interface resulting in flowers quality preservation. Future study will address the problem of determining the quality of flower based related parameters and grading as like floriculture experts.

6. ACKNOWLEDGEMENT

The authors like to thanks to the Principal of DSM College of Science, Arts and Commerce, Parbhani and Head of Department of Computer Science for providing computer research laboratory including image processing facilities and journal access and also for their valuable guidance.

7. REFERENCES

- [1] P. Boissard, V. Martin, and S. Moisan, "A Cognitive Vision Approach to Early Pest Detection in Greenhouse Crops," vol. 2, no. November 2007, pp. 81–93, 2008.
- [2] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," vol. 15, no. 1, pp. 211–217, 2013.
- [3] J. Of and C. Technologies, "Imaging Technique to Measure Leaf Area , Disease Severity and Chlorophyll Content : A Survey Paper," 2012.
- [4] Z. Qin and M. Zhang, "Detection of rice sheath blight for in-season disease management using multispectral remote sensing," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 7, no. 2, pp. 115–128, Aug. 2005.
- [5] R. Sarkate and P. Khanale, "Lesion Sternness Estimation in Crop using Color Image Segmentation," pp. 1–4, 2013.
- [6] D. E. Guyer, "Apple Grading Using Fuzzy Logic," vol. 27, pp. 375–382, 2003.
- [7] S. S. Sannakki, V. S. Rajpurohit, V. B. Nargund, A. K. R, and P. S. Yallur, "Leaf Disease Grading by Machine Vision and Fuzzy Logic," vol. 2, no. 5, pp. 1709–1716, 2011.
- [8] J. Zou, G. Nagy, and N. York, "Evaluation of Model-Based Interactive Flower Recognition," pp. 8–11.
- [9] P. Y. Paduchuri, G. V Deogirkar, S. R. Kamdi, M. C. Kale, and M. D. Rajurkar, "INVITRO CALLUS INDUCTION AND ROOT REGENERATION STUDIES IN *Gerbera jamesonii*," vol. 1, no. 2, pp. 87–90, 2010.
- [10] F. J. Adamsen, T. A. Coffelt, J. M. Nelson, E. M. Barnes, and R. C. Rice, "CROP ECOLOGY , MANAGEMENT & QUALITY Method for Using Images from a Color Digital Camera to Estimate Flower Number," pp. 704–709, 1997.
- [11] K. Glocer, D. Eads, and J. Theiler, "Online feature selection for pixel classification," *Proc. 22nd Int. Conf. Mach. Learn. - ICML '05*, pp. 249–256, 2005.
- [12] R. S. Sarkate, N. V. Kalyankar, and P. B. Khanale, "Application of computer vision and color image segmentation for yield prediction precision," *2013 Int. Conf. Inf. Syst. Comput. Networks*, pp. 9–13, Mar. 2013.
- [13] Z. Effendi, R. Ramli, and J. A. Ghani, "A Back Propagation Neural Networks for Grading *Jatropha curcas* Fruits Maturity," vol. 7, no. 3, pp. 390–394, 2010.
- [14] H. Demuth, "Neural Network Toolbox."
- [15] R. S. Sarkate, P. B. Khanale, and S. B. Thorat, "BPNN Approach in Pixel Classification based Precision Segmentation for Agriculture Images," *IJCA*, no. IJCA Proceedings on National Conference on Recent Advances in Information Technology, pp. 25–27, 2014.
- [16] H. Gassoumi, N. R. Prasad, and J. J. Ellington, "Neural Network-Based Approach For Insect Classification In Cotton Ecosystems," pp. 1–7, 1994.
- [17] N. Wang, K. State, F. Dowell, G. Marketing, N. Zhang, and H. R. Chicago, "Determining Wheat Vitreousness Using Image Processing and a Neural Network," vol. 0300, no. 026089, 2002.