



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 13, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

9940 572 462

6381 907 438

ijrset@gmail.com

www.ijrset.com



Design and Analysis of Image Recognition System using Machine Learning Algorithms for Fragrance Industrial applications

S.Hariprasath¹, R.Deivanai², A.Aarthy³, M.Hemavathi⁴, R.Jeyalakshmi⁵, M.Santhi⁶

¹Assistant Professor, Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

^{2,3,4,5} U.G Scholars, Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

⁶ Professor, Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

ABSTRACT: India is a land of olfactory indulgence, where aromatherapy, incense and ittar have existed since ancient times. Essential oil perfume were a part of the ancient royal lore. The earliest scents used were the healing scents suggested by Ayurveda School of medicine. The suggestions were for healing mental disorders, treating ailments, hygiene and age-control, etc....India is one of the fore-running countries in the world with rich diversity in flora and fauna with its numerous Geo- climatic zones. India can produce organic fragrant raw materials which have great demand in the world. Considering its close linkages with grass-root economics, it can reboot Indian economy from ground level. There is a huge number of flower species in the world. Many different types of flowers and new species of flowers have been discovered continuously. Developing a system for classification of flowers is a difficult task because of considerable similarities among different classes. A digital flower classification and identification system can be used for automatic recognition of flowers without requiring the expertise of a floriculturist. The floriculture industry comprises flower trade, nursery and potted plants, seed and pulp production, micro propagation, and extraction of essential oil from flowers. In such cases, automation of flower classification is essential. Since these activities are carried out manually and are laborious tasks, automation of the classification of flower images is a necessary task. Flower identification research has been growing since few years and it is mainly used to recognize the flower by extracting its features. In this automatic classification of flowers are carried out using ANN method. Colour, Textural and morphological features are extracted and with the help of distance based classifiers, the flower images are classified. The accuracy obtained is 84%. The reported accuracy could be improved by applying CNN methods.

KEYWORDS: Artificial Neural Networks, Flower Pattern Recognition System (FPRS), Morphological features, Textural Classifiers.

I. INTRODUCTION

India is a land of olfactory indulgence, where aromatherapy, incense and ittar have existed since ancient times. Essential oil perfume were a part of the ancient royal lore. The earliest scents used were the healing scents suggested by Ayurveda School of medicine. The suggestions were for healing mental disorders, treating ailments, hygiene and age-control, etc.... Most of these applications well relevant till today. This traditional fragrance industry in India has undergone abundant of changes in the recent years with the introduction of technology. The Indian fragrance industry is one of the largest in terms of production, consumption. India is one of the fore-running countries in the world with rich diversity in flora and fauna with its numerous Geo- climatic zones. India can produce organic fragrant raw materials which have great demand in the world. Considering its close linkages with grass-root economics, it can reboot Indian economy from ground level. Global fragrance & flavour industry is worth \$24.10 Billion and India contributes approximately \$500 million [1]. The growth rate of perfume industry in India is approximately 11% in the last few years and an exponential growth is expected in the upcoming years.

There is a huge number of flower species in the world. Many different types of flowers and new species of flowers have been discovered continuously. Developing a system for classification of flowers is a difficult task because of considerable similarities among different classes. In a real environment, images of flowers are often taken in natural outdoor scenes where the lighting condition varies with the weather and time. Also, there is a lot more variation in viewpoint of flower images and the background also makes the problem difficult as a flower has to be segmented

automatically. There are about 250,000 named species of flowering plants in the world. As there is a similarity between the flowers, the automated classification of flowers is an interesting and on demand task.

A digital flower classification and identification system can be used for automatic recognition of flowers without requiring the expertise of a floriculturist. The floriculture industry comprises flower trade, nursery and potted plants, seed and pulp production, micro propagation, and extraction of essential oil from flowers. In such cases, automation of flower classification is essential. Since these activities are carried out manually and are laborious tasks, automation of the classification of flower images is a necessary task. Flower identification research has been growing since few years and it is mainly used to recognize the flower by extracting its features.

The organization of the research work is as follows. The description and discussion of successfully implemented system are reviewed in section II. The proposed methodology is described in section III. The results and inferences are given in section IV. The future scope is listed out in Section V.

II. LITERATURE SURVEY

Isha patel and Sanskruti patel [1] noted a methodology for extraction of basic and morphology features using computer vision techniques including pre-processing and segmentation. Dataset accumulation that contains 25000 images with 102 species. Building a prediction model using MKL – SVM with multi-label classification that identifies and classifies the newly input flower image with optimum accuracy. The flower is divided using a threshold-based technique, and texture choices, specifically the colour texture moments (CTMs), gray level co-occurrence matrix (GLCM), and scientist responses, unit of measurement are extracted. These choices unit of measurement used for work and classification is applies using a probabilistic neural network. A measuring of foliage colours victimization principal component analysis. A semi-automatic plant identification is based on digital leaf and flower images. They used edge and color features of flower pictures for the KNN to classify flowers. In this paper, the KNN, Random Forest and ANN have not worked well while SVM achieved good accuracy with 72% accuracy. However, the hybrid MKL + SVM achieved the highest accuracy among the other classifiers with 78% accuracy.

The MKL-SVM classifier has given a recognition rate of 78.3%. Warisara Pardee et al.[2]. In this paper they use the dataset of 15 different types of flowers. They used the Grabcut in the segmentation phase to segment the foreground from the background. Feature Extraction had done by RGB Histogram. Random Forest algorithm was used as classification process. Performance testing of the system shown that it was able to separate image foreground from an image background with 86.14% accuracy. With 15 types of input flowers, the system was able to classify flowers with 80.67% accuracy.

H. Salome Hema Chitra¹, S. Suguna¹ and S. Naganandini Sujatha [3] They use the images of three types of products the image was resizing, RGB to Grayscale conversion and gray to binary conversion, closing and opening operations to obtain the pre-processed image and segmented an image by using edge function, Laplacian Of Gaussian (LoG) method. They extracted the features as color, shape and texture. work is useful to know about various issues in seed identification in agricultural sector. This work provides efficient methods to overcome the manual identification of seeds. We can also extend this work with an efficient identification and classification methodology in image analysis techniques in agricultural sector.

Anelia Angelova et al.[4] They had a large dataset containing 578 flower species and 250,000 images. For simplicity we use the super-pixel segmentation method by Felzenszwalb and Huttenlocher to over-segment the image into small coherent regions. Segmented the flower feature by using Laplacian-based segmentation.

Asma Najjar and Ezzeddine Zagrouba [5] evaluation was performed using a flower dataset provided by Oxford University and used to achieve the work of Maria-Elena Nilsback and Andrew Zisserman. There were 17 classes in this dataset and 80 images in each class. Segmentation had done by OTSU thresholding, Lab color space, RGB to Lab transformation. They proposed a fast flower image segmentation schema based on OTSU thresholding technique and Lab color space.

III. METHODOLOGY

The system developed through this research work consists of a sequence of image pattern recognition tasks generally followed in any pattern recognitions system. The various modules of the system are shown in figure 1. The functions of each module is described below.

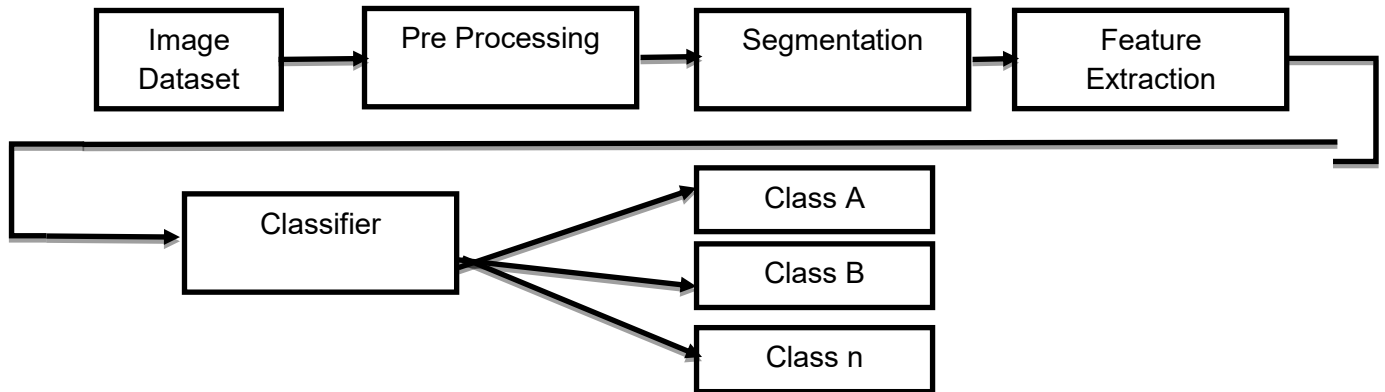


Figure 1. Block Diagram of Flower Pattern Recognition [FPR] System

IMAGE DATASET

The name of the dataset is FLDataset. The dataset is extracted from Kaggle. This dataset contains 4242 images of flowers. The data collection is based on the data flickr, google images, yandex images. one can use this dataset to recognize plants from the photo. The pictures are divided into five classes: chamomile, tulip, rose, sunflower, dandelion.

For each class there are about 800 photos. Photos are not high resolution, about 320x240 pixels. Photos are not reduced to a single size, they have different proportions.

A.PREPROCESSING

The first step towards flower image processing is the pre-processing of an image. As image pre-processing is the fundamental step, it is applied to increase the quality of images and remove the inappropriate noises presented in images. The aim of pre-processing techniques is image enhancement and image restoration. Image enhancement makes a picture look better, without regard to how it really truly should look. Image restoration tries to fix the image to get back to the real, true image. An image may contain various types of noises including Gaussian, Salt and Pepper, Poison, Speckle etc. The images are resized from 320x240 to 400x400. Then noises are removed by applying an appropriate filter that gives an optimum result for specific noise removal. For removal of blur, Gaussian filtering technique has been used. Image enhancement is done by Dilution. Image restoration is done by Erosion.

B.SEGMENTATION

Segmentation is applied on the image for further processing. Removing the undesirable background in the image is the second step in flower identification. Images that contain flowers also contain parts of plants, leaves or grass in the background. In order to extract the correct features, it is required to separate the flower image from its background. To remove the background of images and improve the quality of flower image foreground, segmentation techniques are used. There are many methods available for image segmentation: split and merge, region-based method, watershed-based segmentation, threshold-based, etc. Based on the experiment carried out, machine-driven threshold-based segmentation methodology cited as Otsu’s Threshold methodology was applied.

C.FEATURE EXTRACTION

In the field of computer vision and image processing, feature describes information that is used to identify or detect some object related to certain application. Feature extraction is the process applied to an image to convert its visual information into vector space.

These features include color by using HSV color spacing and LAB color spacing and texture by using HARALICK technique. Morphological features like Corona, Disk and Aestivation of petals.

1) COLOR:

Various color model including RGB, HSV, LAB, etc. are available to extract color features. To extract the color feature from a flower image two color models are used i.e. HSV color space and LAB color space.

1.a) HSV Color space :

HSV stands for Hue, Saturation and Value, which are the three components that make up the HSV color model. In Image processing, HSV is a color model used to represent colors in an image. Hue (H) represents the color in terms of a 360-degree circle, where red is at 0 degrees, green is at 120 degrees, and blue is at 240 degrees. Other colors fall at points in between. Saturation (S) is the intensity of the color, ranging from 0 (gray) to 100 (full color). Value (V) represents the brightness of the color, ranging from 0 (black) to 100 (white). By using the HSV color model, we can separate the color information from the brightness information in an image, which can be useful for image segmentation and color detection tasks in image processing.

2) LAB Color Space :

The LAB color model, also known as the CIELAB color space, is a color model designed to be more perceptually uniform than other color models like RGB or CMYK. It is based on the opponent color theory, which states that color perception is based on three opposing channels: lightness vs. darkness, redness vs. greenness, and yellowness vs. blueness. In the LAB color model, color is represented by three values: L, a, and b.

L represents the lightness or darkness of a color, and ranges from 0 (black) to 100 (white).

The a and b values represent the redness vs. greenness and yellowness vs. blueness of a color, respectively.

The a value ranges from -128 (green) to 127 (red)

The b value ranges from -128 (blue) to 127 (yellow)..

2) TEXTURE:

A wealth of textures are observed on both artificial and natural objects such as those on wood, plants, materials and skin. In a general sense, the word texture refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. A texture is usually described as smooth or rough, soft or hard, coarse or fine, matt or glossy, and etc.

Textures might be divided into two categories, namely, *tactile* and *visual* textures. Tactile textures refer to the immediate tangible feel of a surface. Visual textures refer to the visual impression that textures produce to human observer, which are related to local spatial variations of simple stimuli like colour, orientation and intensity in an image. This thesis focuses only on visual textures, so the term 'texture' thereafter is exclusively referred to 'visual texture' unless mentioned otherwise.

Texture feature is extracted using GLCM:

GLCM stands for Gray-Level Co-occurrence Matrix. It is a statistical method used in image processing and computer vision to analyze the spatial relationships between pixels in an image. The GLCM is a matrix that describes the frequency at which different combinations of gray levels occur in an image. The GLCM is created by analyzing pairs of pixels in an image that are located at a certain distance and orientation from each other. The matrix is then used to calculate various texture features that can be used to characterize the image. These texture features include contrast, correlation, energy, and homogeneity.

1. Contrast: measures the local variations in gray-level intensity between adjacent pixels in an image. A high contrast value indicates that adjacent pixels have very different gray-level intensities.

2. Correlation: measures the degree of linear dependence between gray-level intensities of neighbouring pixel. A high correlation value indicates that adjacent pixels have a similar gray-level pattern.

3. Homogeneity: measures the closeness of the distribution of gray-level intensities to the diagonal of the GLCM. A high homogeneity value indicates that the gray-level intensities of neighbouring pixels are very similar.

4. Energy: also known as angular second moment, measures the sum of squared elements in the GLCM. A high energy value indicates that there are many co-occurring gray-level pairs in the image.

These texture features can be used to distinguish between different types of textures or materials in an image. For example, a rough texture may have high contrast and low homogeneity values, while a smooth texture may have low contrast and high homogeneity values.



D.CLASSIFIER

The classifier used here is that Distance based classifier. Distance based algorithms are methods of classification that each item that is mapped to the same class may be thought of as more similar to other items in the class than it is to the items found in the other classes. In Distance based classifier there are two types:

1. Similarity based distance measurement
2. Dissimilarity based distance measurement.

Similarity measure is a way of measuring how data samples are related or closed to each other. The similarity measure is usually expressed as a numerical value: It gets higher when the data samples are more alike. It is often expressed as a number between zero and one by conversion: zero means low similarity (the data objects are dissimilar). One means high similarity (the data objects are very similar). Dissimilarity measure is to tell how much the data objects are distinct.

Hamming distance:

The hamming distance is equal to the number of digits where two code words of the same length differ. In the binary world, it is equal to the number of different bits between two binary messages. The hamming distance between two messages can be calculated using:

$$H(P, Q) = \sum_{i=1}^n |p_i - q_i|$$

Manhattan Distance:

This metric is very useful in measuring the distance between two streets in a given city, where the distance can be measured in terms of the number of blocks that separate two different places. This method was created to solve computing the distance between source and destination in a given city where it is nearly impossible to move in a straight line because of the buildings grouped into a grid that blocks the straight pathway.

$$d(P, Q) = \|P - Q\| = \sum_{i=1}^n |p_i - q_i|$$

RANDOM FOREST CLASSIFIER:

A random forest classifier is a machine learning algorithm that is used for classification tasks. It is an ensemble learning method that combines multiple decision trees to make a final prediction.

IV. EXPERIMENTAL RESULTS

TABEL I: Colour Features

Flower Name	HSV Colour Feature			LAB colour		
Flower Name	H	S	V	L	A	B
daisy	61	145	182	166	108	129
dandelion	95	118	133	122	122	120
lotus	89	150	234	105	158	94
rose	108	64	230	201	127	109



sunflower	98	248	225	167	112	90
tulip	79	103	217	185	134	106
water lily	62	127	148	89	127	148

TABEL II. Haralick Features

Flower	Homogeneity	Energy	Contrast	Correlation
daisy	0.645	0.470	21534.769	32121.271
dandelion	0.471	0.403	32833.383	52474.822
lotus	0.607	0.463	24607.871	36501.066
rose	0.608	0.477	24073.804	24170.654
sunflower	0.664	0.605	42928.456	64434.281
tulip	0.551	0.356	26839.408	45158.853
water lily	0.695	0.496	17665.021	28497.067

TABEL III. Morphological feature

Flower	Corona	Disk	Aestivation of flower
daisy	0	1	1
dandelion	0	1	1
lotus	1	0	1
rose	0	1	1
sunflower	0	1	1
tulip	0	1	1
water lily	1	0	1



TABEL IV: Hamming Distance Classifier

Flower name	Threshold value	Accuracy
daisy_1	0.65	0.410
daisy_2	0.65	0.414
daisy_3	0.65	0.417
dandelion_1	0.65	0.605
dandelion_2	0.65	0.601
dandelion_3	0.65	0.621
lotus_1	0.65	0.572
lotus_2	0.65	0.500
lotus_3	0.65	0.553
rose_1	0.65	0.781
rose_2	0.65	0.794
rose_3	0.65	0.726
sunflower_1	0.65	0.468
sunflower_2	0.65	0.470
sunflower_3	0.65	0.443
tulip_1	0.65	0.345
tulip_2	0.65	0.383
tulip_3	0.65	0.318
water lily_1	0.65	0.455
water lily_2	0.65	0.456
water lily_3	0.65	0.434

TABEL V: Manhattan Classifier

flower name	threshold value	accuracy
daisy_1	0.78	0.668
daisy_2	0.78	0.643
dandelion_1	0.78	0.564
dandelion_2	0.78	0.523
lotus_1	0.78	0.401
lotus_2	0.78	0.441
rose_1	0.78	0.601
rose_2	0.78	0.671



Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, India

sunflower_1	0.78	0.786
sunflower_2	0.78	0.731
tulip_1	0.78	0.654
tulip_2	0.78	0.681
tulip_3	0.78	0.611
water lily_1	0.78	0.881
water lily_2	0.78	0.875

Table VI. Classification Accuracy

Classifier	Threshold Value	Accuracy
Hamming Distance	0.65	76%
Manhattan Classifier	0.78	85%
Random Forest Classifier	—	90%

REFERENCES

1. Isha Patel and Sanskruti Patel, “Flower Identification and Classification using Computer Vision and Machine Learning Techniques”, International Journal of Engineering and Advanced Technology (IJEAT), Volume-8 Issue-6, August 2019
2. Dr.S.M.Mukane and Ms.J.A.Kendule, “Flower Classification Using Neural Network Based Image Processing” IOSR-JECE Volume 7, Issue 3, Sep.-Oct. 2013
3. Warisara Pardee, Prawaran Yusungnern, Peeraya Sripiyan, “Flower Identification System by Image Processing”, Conference Paper · August 2015.
4. Isha Patel, Sanskruti Patel, Atul Patel, “Analysis of Various Image Preprocessing Techniques for Denoising of Flower Images”, International Journal of Computer Sciences and Engineering, Vol.-6, Issue-5, May 2018, pp. 1012-1018.
5. H. Salome, Hema Chitra, S. Suguna and S. Naganandini Sujatha, “A Survey on Image Analysis Techniques in Agricultural Product”, Indian Journal of Science and Technology, Vol 9(12), March 2016, pp. 1-13.
6. Anelia Angelova, Shenghuo Zhu, Yuanqing Lin, “Image segmentation for large-scale subcategory flower recognition”, IEEE, 2013.
7. Asma Najjar, Ezzeddine Zagrouba, “Flower image segmentation based on color analysis and a supervised evaluation”, ResearchGate, IEEE, 22 May 2014, pp. 402-406.
8. Tzu-Hsiang Hsu & Chang-Hsing Lee & Ling-Hwei Chen, “An interactive flower image recognition system”, Springer, Multimed Tools, Appl (2011), pp. 53–73.
9. Aswini Kumar Mohanty, Amalendu Bag, “Image Mining for Flower Classification by Genetic Association Rule Mining Using GLCM features”, International Journal of Advanced Engineering, Management and Science (IJAEMS), Vol-3, Issue-5, May- 2017, pp. 516-521.
10. Basavanna M., S. S. Gornale, “Identification and Classification of Flowers Images using Colour models”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 6, June 2015, pp. 1003-1007.



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details