International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Flower Identification and Classification using Computer Vision and Machine Learning Techniques

Isha Patel, Sanskruti Patel

Abstract*: Computer vision techniques plays an important role in extracting meaningful information from images. A process of extraction, analysis, and understanding of information from images may accomplished by an automated process using computer vision and machine learning techniques. The paper proposed a hybrid methodology using MKL – SVM with multi-label classification that is experimented on a dataset contained 25000 flower images of 102 different spices. Basic and morphology features including color, size, texture, petal type, petal count, disk flower, corona, aestivation of flower and flower class are extracted to increase the classification accuracy. Various classifiers are applied on extracted feature set and their performance are discussed. The result of MKL – SVM with multi-label classification is very promising with 76.92% as an accuracy rate. In brief, this paper attempts to explore a novel morphology for feature extraction and the applicability of symbolic representation schemes along with different classification strategies for effective multi-label classification of flower spices.

Index Terms*: Flower morphology features, Image processing, Machine learning, MKL-SVM, Multi-label

I. INTRODUCTION

There are almost 250,000 named species of flowering plants within the world. Several blooming flowers are often discovered within the garden, park, roadside, and plenty of alternative locations, and identification of flowers are often done solely by taxonomists or botanists. The general public does not have any data regarding to these flowers and so as they mainly use flowers guide books or use the relevant websites to have the related information.

Flowers are a type of plants that have many categories; many of those categories or species have very similar features and looks, while one can find dissimilarity among the same flower species. This similarity and dissimilarity make the flowers recognition process with highly accurate result a very hard challenge. Identifying unknown plants depends a lot on the inherent data of a skilled plant scientist. However, this method of manual recognition is usually arduous and long. However, flowers area unit most typically used for plant identification, the stem, leaves, petals, seeds and even the full plant may be employed in an automatic method.

Recently, taxonomists started attempting to search out efficient strategies to satisfy species identification wishes, like developing a digital image methodology and pattern recognition techniques. Digital image methodology refers to the employment of algorithms and procedures for operations like image improvement, compression, image analysis, mapping, and dereferencing. The approaches available in the field of computer vision focused on the problem of how to make possible for computers to see. Basically, computer vision uses the observed image data to infer something about the world. Flowers generally having certain morphological structures, such as Calyx, Corolla, Stamen or Pistil [1]. Machine learning is a field of artificial intelligence (AI). It provides a set of algorithms that make computer system capable to learn automatically. Moreover, it makes systems able to improve gradually without being explicitly programmed based on experience. From a machine learning perspective, images are made of with pixels and contained the information that can describe that image. As these information is too extensive and mixed-up therefore it is not possible to be directly used by a machine learning algorithm. Image Processing techniques are used for feature extraction and Machine Learning algorithms are deploys these features to build a model. As data increases in a tremendous way it is necessary to organized them to extract the meaningful information within. Moreover, it has been observed that, prediction of multiple labels are highly required by today's classification problems and it is needed to integrate the approaches related to multi-label classification. In this paper, a novel methodology is represented which is the combination of two classification algorithms, MKL (Multiple Kernel Labeling) and SVM (Support Vector Machine), with 14 different sets of features. The features extracted are color, size, texture (GLCM), disk flower, no. of petals, type of petals, flower aestivation, flower class, and their combinations. The experiment is carried out on flower dataset that contained 25,000 flower images that are categorized with 102 classes of flower species. In this research, an innovative approach that is multi-label classification using MKL-SVM is proposed.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Isha Patel*, Faculty of Computer Science and Applications, Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat 388421, India.
Sanskruti Patel, Faculty of Computer Science and Applications, Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat 388421, India.

In order to use of the multi-label, i.e. label powerset, a problem is transformed into a multi-class problem where with training data set, one multi-class classifier is trained on all distinctive label groupings [2]. Determine the acceptable kernel throughout the training method is troublesome; therefore, several researchers are attempting to develop a lot of versatile kernel learning which is referred to as Multiple Kernel Learning (MKL) [3]. Variety of studies have shown that MKL could also be an excellent tool for visual perception of flower images. Here, a flower image is represented by multiple sets of choices and MKL is applied to combine these different feature sets. Multiple kernel learning (MKL) could also be a scrupulous approach for selection and composition of kernels for a given recognition task. The contribution of the paper are: (i) Efficient methodology for extraction of basic and morphology flower features using computer vision techniques including pre-processing and segmentation. (ii) Dataset accumulation that contains 25000 images with 102 species (iii) Building a prediction model using MKL – SVM with multi-label classification that identifies and classifies the newly input flower image with optimum accuracy.

The content of the paper is organized as follows. The summary of the related work is provided in Section II. Section III contains the proposed algorithm for flower identification and classification. The experimental result and discussion of the performance analysis are available in Section IV. Comparison with previous work is presented in Section V.

II. Related Review

Nilsback et al. [4] noted that color and type unit of measurement are the key choices in flower classification. 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. Y. Yoshioka et al. [5] performed a measuring of foliage colours victimization principal component analysis. They thought-about the first five principal components (PCs) of the foremost sq. on the petals. In their work, T. Saitoh et al. [6] described the associate automatic technique for recognizing a blooming flower supported some way for extracting the lower regions. T. Saitoh et al. [7] used a piecewise linear discriminate analysis for recognition. The system for classification of Malaysian blooming flower is made available by Fadzilah Siraj et al. [8]. In this paper, they showed the usage of NN and a picture method for flower image options. The prediction accuracy of provision regression is 26.8%. A semi-automatic plant identification is represented by Pavan Kumar Mishra et al. that is based on digital leaf and flower images [9]. By taking the base of color, volume of a shape and features of a cell, a multiclass classification is proposed. Tanakorn Tiay et al. projected flower recognition system supported image process [10]. They used edge and color features of flower pictures for the KNN to classify flowers. The accuracy of this method is around eighty percentage. Dr. S.M. Mukane et al. planned Flower Classification exploitation using ANN based method [11]. The planned technique relied on textural options like gray level co-occurrence matrix and distinct rippling rework (DWT).

III. A Hybrid Methodology For Flower Identification And Classification

The proposed methodology for flower identification and classification is represented in figure 1. The process is started with image pre-processing step with image resizing, noise filtering and RGB and Grayscale conversion. Afterwards, image segmentation is completed with Otsu thresholding technique. After performing segmentation, the features are take out from segmented image including basic and morphology features of flowers. The outcome of data pre-processing, segmentation and feature extraction is the final training data set that is used during model building phase [12]. In model building phase, a model builds and tests by applying appropriate machine learning method on training data set. The dataset containing training examples is considered as an input for the model building phase. It is generally divided into three categories: training, validation, and testing. 80% of the data is contained by training data set including 10% of the validation data set and 10% of the data is contained by the test data set. During experimental work, MKL – SVM is emerged as a most efficient algorithm and it has been tested over a dataset containing 25000 flower images. After a model has been trained and tested, it is able to identify and classify a flower image and generates the predicted output.

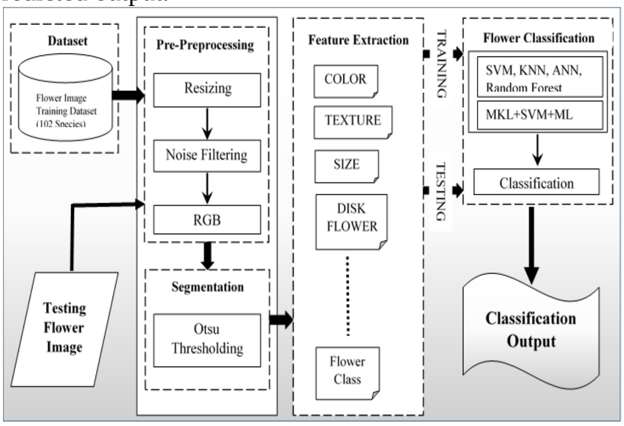

Fig. 1. An Architectural Diagram of Proposed Methodology for Flower Identification and Classification.

A. Flower 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 removing the inappropriate noises presented in images. The aim of pre-processing techniques is image enhancement and image restoration [13]. Image Enhancement is a significant process that aimed to recover the visual look of an image. It is provided for the better transform representation for the next phases of image detection [14]. Image enhancement is characterized in three types: Image scaling, Color space transformation, Contrast Enhancement [15].

For image restoration, first noise level resides in the image is measured. Normally, an image may contain various types of noises including Gaussian, Salt and Pepper, Poison, Speckle etc. [16]. These noises are removed by applying appropriate filter that gives an optimum result for specific noise removal. The applied filters are average (5x5), median (5x5), and min (5x5) and max (5x5) [17]. For removal of blur, the filtering technique like, Inverse filter, Wiener filter, Lucy Richardson filters have been used [18] (figure 2 (a) and (b)). After applying noise removal, RGB to grayscale transformation is applied to an image.

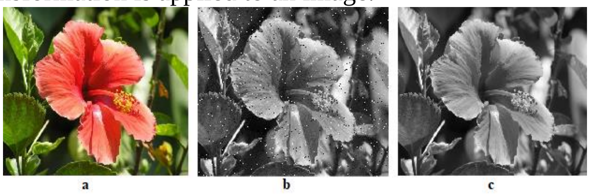

Fig. 2 (a) Input image; (b) Salt and pepper noise image; (c) Filtered image by median filtering

B. Flower Segmentation

After applying necessary image pre-processing techniques, segmentation is applied on image for further processing. Removing the undesirable background in the image is the second step in flower identification. Images that contains flowers are too contain parts of plant, 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. [19, 20 and 21]. Based on the experiment carried out, machine-driven threshold-based segmentation methodology cited as Otsu's methodology [22] was applied. It transforms the flower image into a grayscale to binary for reducing the intricacy of the data. Further, to store the segmented images, feature extraction operations. The figure 3 contains the outcome of the Otsu technique applied on a flower image.

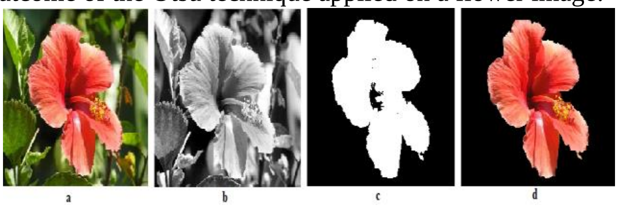

Fig. 1 (a) Filtered image after preprocessing; (b) Convert image into grayscale; (c) Otsu thresholding; (d) Final result of segmentation

C. Flower Feature Extraction and Selection

After applying image segmentation, features are extracted from the segmented image. In the field of computer vision and image processing, feature describes an 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. A flower has two types of features: basic and morphology. Selection of feature is a crucial step as it removes irrelevant attributes for the purpose of increasing prediction accuracy as high dimensionality and high volume may become a challenge to the learning task [23, 24]. Within identical species, flowers could look completely different and generally flowers from different species contains high similarity. Moreover, some flowers area unit distinguishable by their colors, whereas others have special types of texture. The key challenge of classification is to figure out acceptable choices to infer the visual information of flower image and to produce a classifier in such a way, that it is able to differentiate between different species. In this work, total 14 distinctive features pertaining to flower image are extracted using various methods. These features include color, texture, size, and morphology features like disc flower, petal type, no. of petals, aestivation of the petal, corona, and flower class [25]. The following subsections depicts the process of extracting a specific feature from a flower image.

D. Basic Flower Features

1) Color: Various color models including RGB, HSV, Lab, etc. are available to extract color feature. To extract color feature from a flower image, two color models are used i.e. HSV and Lab color space [26]. HSV color space normally having better representations about how people relate to color than RGB color model and it also generates high quality graphics. In HSV color space, Hue, Saturation and Value refer to tint, shade and tone respectively. To approximate human vision, normally lab color model is used. To add more precision for interpreting color feature value, another prominent model, Lab color space is also used. Lab color model represents value for L as "lightness", value for a as "green to red" and value for b as "blue to yellow" [27]. The figure 4 depicts the values of both of the color models.

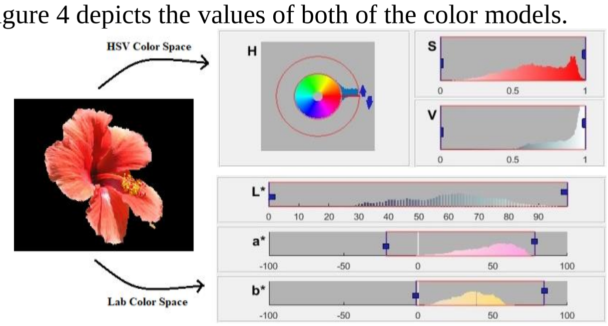

Fig. 4 Color histogram for HSV and Lab color space

2) Size: Plants take completely different times to maneuver from germination to maturity once flowers develop all their procreative elements and therefore the life cycle begins once more. Width and height ratio is a ratio between the blooming of a flower maturity [28]. Based on the height and width function measures, it is identified that flower is in blooming or bud stage. Size feature is captured by the centroid point and boundary box methods [29 and 30]. The flower centroid and boundary tracing algorithm usually starts from the beginning point. When it reaches to end, it stops. By this way, it gets the four flower boundaries that are partial and then it will be combined them to form the whole flower boundary. Using this techniques, shape of flower image can properly have identified and extracted with width and height measurements. The outcome of applying this method is shown in following figure 5.

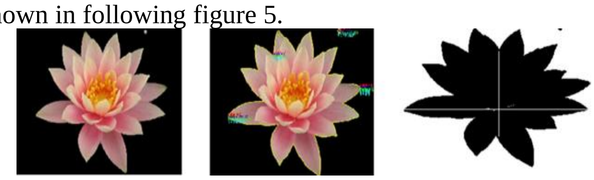

Fig. 5 Height and Width Measurements of a Flower Image

3) Texture: Texture feature is extracted using Gray level co-occurrence matrix, abbreviated as GLCM [31]. GLCM feature extraction is a mixture of gray levels co-occurs in a picture section or image. A quantitative description of a spatial pattern can efficiently provide by applying GLCMs. The four GLCM-derived options extracted from a flower image are homogeneity, energy, contrast, and correlation [32]. The following Table 1. Describes the formulas used for calculating GLCM feature values.

TABLE 1
Formula to calculate texture features from GLCM

Sr. No.	GLCM Features	Formula
1.	Homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$
2.	Energy	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
3.	Contrast	$\sum_{i,j=0}^{N-1} P_{i,j}\,(i-j)^2$
4.	Correlation	$\sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$

E. Morphology Features of Flower

The study of varied external options of an entity is understood as morphology. To identify an image with optimum accuracy, it is highly essential to extract flower morphology features [33]. Morphology provides physical form and external structure of one flower which make it distinguishable with another flower. In this work, the morphology features extracted from a flower image are Disc flower, Corona, Petal type, No. of petals, Aestivation of a petal and flower class.

1) Disc Flower: Various color models including RGB, HSV, Lab, etc. are available to extract color feature. To extract color feature from a flower image, two color models are used i.e. HSV and Lab color space [26]. HSV color space normally having better representations about how people relate to color than RGB color model and it also generates high quality graphics.

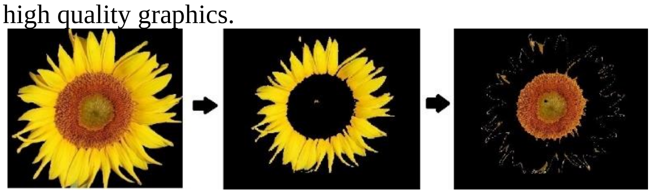

Fig. 6 Disk Flower Structure extracted from a Sunflower Image

2) Corona: One of the other prominent morphology feature of a flower image is the corona (Crown). It is a characteristic that possesses androecial derived structure and found in Asclepiadaceous flowers with a variable form, color and dimension, which is incorporated in the gynostemium. The corona typically consists of inner and outer lobes. Corona feature is captured by the k-means clustering method [34]. It helps to identify flowers like water lily, lotus etc.

Fig. 7 Corona Extraction from a Water Lily Flower Image

3) Petal Type: Modified and normally colored leaves that surround the generative components of flower area unit are known as petals. They are formed to draw an attention of pollinators and they are usually brilliantly colored. Together, all the leaves of a flower referred to as a Corolla. Petal type can be free or united. The free petal type structure is shown in figure 8. Applying the convex hull method and ripple image algorithm [33], the feature of flower petal-type which is free or united is extracted. Using the convex hull method, as it is used to represent description of region shape of flower image, the petal of flower boundary or edges can be converted in binary image and easily detected. After applying convex hull method, the ripple algorithm is applied for representation of the target in different view angles and sizes [36]. The features of concentric circles are obtained to conclude whether flower petals are free or united.

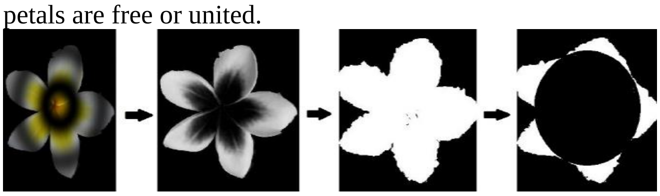

Fig. 8 Free Petal Type of Plumeria Flower

4) No. of Petals: It is also essential to calculate number of petals remain present in a flower as the quantity of petals in flower has significantly helps to classify a plant. A composed form of all the petals remains in a flower image is represented by a Corolla. Corolla of different species of plants varies greatly in color and pattern. This feature works well where the roundness of the flower is high. Sample of a number of petals remain present in a flower image is shown in figure 9. The Skeleton with DCE (Discrete Curve Evolution) technique is applied for counting the flower petals [37]. Discrete Skeleton Evolution (DSE) mainly applied to reduce a morphological or topological skeleton [38]. It is a form of pruning that removes noisy or redundant branches generated by the skeletonization process. Moreover, it preserves information-rich trunk segments. The sample results of how to get number of petals using DSE method is shown in figure 9.

Fig. 9 No of Petals Generated by Applying DSE Method

5) Aestivation of Petal: The arrangement of petals or sepals within the flower is denoted as Aestivation. Different aestivation of petals are available includes (i) Valvate: The floral petals are settled side by side. (ii) Twisted: The petals are settled regular intersecting in one direction. (iii) Imbricate: The flower petals are completely in, one completely out and others are intersecting in one direction. (iv) Quincuncial and Vexillary: It found only in pentamerous flower and it is rate one. Aestivation of petal feature is captured by the Laplacian filter method [39]. The Laplacian of an image highlights regions of rapid intensity change and is therefore often used for edge detection. This operation in result produces such images which have greyish edge lines and other discontinuities on a dark background. This produces inward and outward edges in an image as per shown in figure 10.

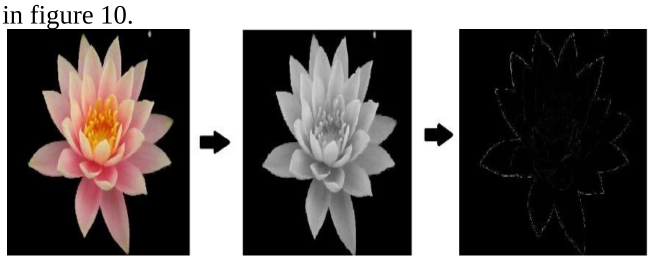

Fig. 10 Sample Result of Lotus Flower Aestivation of Petal

6) Flower Class: Monocots or Dicots are types of flower classes which can be easily identified by dividing number of petals [40]. Where Monocots Petals are appeared in combinations of three, example is a Lily flower, Dicots Petals are seemed in combos of five or four. The example of a dicots flower is a Rose. To obtain the flower class, the number of petals is divide by 3 or 4 or 5. It is one of most required morphology feature which facilitates the classification of a flower. The figure 11 shows the flower class types.

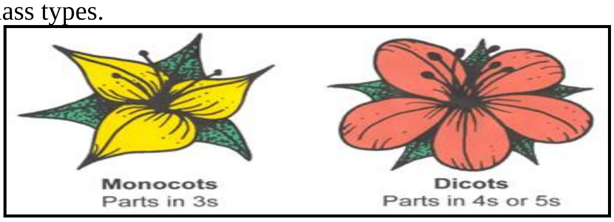

Fig. 11 Types of flower classes (Monocots and Dicots)

F. Flower Features and Its Probable Values

The following Table 2 represents all fourteen features for a specific flower image with feature name, method used to extract that feature and its probable value.

TABLE 2

Flower Features and its Probable Value

Flower Image	Use Method for Feature Extraction	Features Name	Probable Values
	HSV converter	HSV	1.75
	LAB converter	Lab	86.76
	Contrast	Contrast	0.11298
	Homogeneity	Homogeneity	0.97478
	Energy	Energy	0.41411
	Correlation	Correlation	0.988819
	Height	Height	178 (in cm)
	Width	Width	563 (in cm)
	K-means Clustering	Disk flower	0
	K-means Clustering	Corona	1
	Skeleton with DCE (Discrete Curve Evolution)	No of petals	25
	Convex hull and ripple image algorithm	Petal type	0.8570
	Laplacian Filter	Aestivation of petal	0.0043
	No of petals and divide by 3 or 4 or 5	Flower Class	5

G. Machine Learning Algorithms for Flower Species Identification

Machine Learning is an emerging field of Artificial Intelligence and it offers the set of algorithms that are used to diagnose and uncover the hidden patterns in data to perform useful inference using those patterns that have been learned [41]. Machine learning algorithms are categorized into supervised, unsupervised and reinforcement learning. Classification algorithms are part of supervised learning approach where algorithm learns from the pre-labelled data [42]. The algorithm learns and discovers the patterns from the data and assigns a new label automatically to the new and unseen data based on the learning. In classification, based on the dataset, a model is build and trained to predict the outcomes. There are variety of algorithms are available for classification and it is not possible to conclude the best among others. Classification is a technique to categorize the information into a definite and desired number of categories with label to each class. The most commonly used classification methods are: 1) k-nearest-neighbors; 2) Random Forest; 3) Artificial Neural Network and 4) SVM [43, 44, 45, 46]. Applying these methods to classify flower image from 102 species dataset, containing 25,000 images of flowers, this dataset has been split into a training set containing 80% images, validation set containing 10% images and other 10% for test set images.

H. A Proposed Hybrid Method using MKL – SVM with Multi-Label Classification

The dataset generated after applying various computer vision techniques contains different features of flowers that are linear and nonlinear. A single machine learning algorithm is unable to produce better time and accuracy performance. The reason behind use of MKL (Multiple Kernel Learning) algorithm is to add an extra parameter to the minimization problem of the learning algorithm. As the combination of data from different features have different notions of similarity therefore they require different kernels. The ability to select for an optimal kernel and parameters from a larger set of kernels, reducing bias due to kernel selection while allowing for more automated machine learning methods [47]. In this paper, we focused on two challenges of flower image identification and classification; and proposed a method to resolve them. The first challenge is that representing flower images by different combination of basic and morphology features, such as color, shape, texture, disk flower, flower class, etc. which helps to provide better information and identification of flower.

And the second challenge is that the huge collection of flower classification using the multi-labeling.

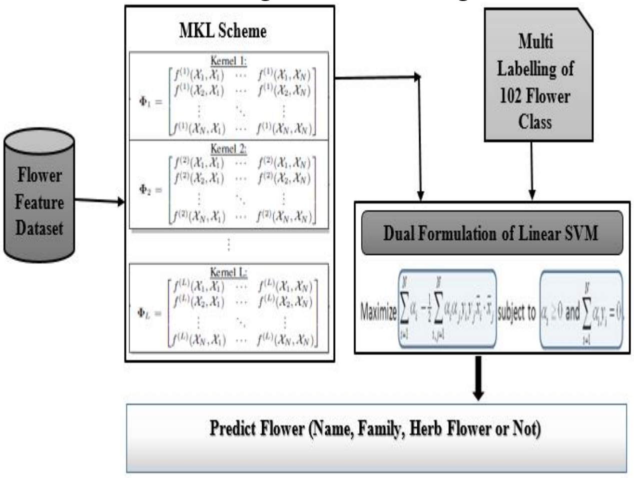

Fig. 12 Graphical representation of MKL-SVM Approach

In this section, the hybrid approach (Multiple Kernel Learning –Support Vector Machine) of classification with Multi-Labelling is explained. The above fig. 12 represents the flow of hybrid approach (MKL-SVM with Multi-lableling). The dataset of extracted features is assigned to the MKL model [48]. To classify data, instead of a single kernal, MKL considers a linearly weighted sum of kernels. Different kernals are combined using MKL including Linear, Quadratic, MLP, RBF and Polynomial. Each kernel is computed based on a discrete feature. Therefore, for an example, to describe the color information of a flower image, a color based kernel is used. In this way, the flower dataset is assigned to SVM algorithm which analysed (flower) data for flower classification. After applying MKL classifier for large datset set; SVM is needed as it separated by hyperplane that has the largest distance to adjoining training data point of any class.

Here MKL enables to estimate optimal weights to combine image features for each category. First a training set is assigned with labeled samples to MKL model, the following equation is derived N no of training set with labeled sample;

$$\mathbf{D} = \{(X_i, Y_i)\}_{i=1}^{N};$$

Where $X_i \in R^n,\ Y_i \in \{-1,\ 1\}$

As we know MKL is the set of linear and non-linear kernels. In our study, the kernel is actually a convex combination of basis kernels:

$$K\,(X_i, X_j) = \sum_{m=1}^{M} \beta_m k_m(X_i, X_j)\ ;$$

With $\beta_m \ge 0, \sum_{m=1}^{M} \beta_m = 1$

Here linear kernels could be categorized as basic kernels, radial basis function kernel, polynomial kernel etc. For all samples, MKL computed weights of kernels that are same. Therefore, for each test image, each kernel has a fixed share in deciding the class of that test image. After combination of MKL + SVM using a dual formulation of optimization problem. The learning of the hyperplane in linear SVM is done by transforming the dual optimization problem using polynomial kernel. This is where the kernel plays role.

$$\max_{\alpha}\Big[\sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i\alpha_j y_i y_j k(X_i, X_j)\Big]$$

Subject to $\sum_{i=1}^{N} \alpha_i y_i = 0, \quad 0 \le \alpha_i \le C$

At last step, a novel approach applying the multi-labeling classification for flower image transform the training task into one or a more single-label classification tasks. For that, Label Powerset (LP) is one of the multi-labeling approach that is used. Here, from a multi-label [49] training set, each different set of labels converts to a different class in a new single-label classification task. Whenever a new data introduced, the single-label category of label powerset produces the foremost probable class. This class contains a set of labels.

IV. Experiment Result Analysis and Discussion

This section covers all the details of an experiment carried out on the flower image dataset. Moreover, it contains the result analysis and comparisons between the classifications algorithms considered in the experiment.

A. Flower Image Dataset

Accumulation of related dataset is a crucial for carrying out experiment and producing results. In this experiment, we have considered the benchmark dataset of Oxford flower dataset (total 8189 number of image with 102 flower of spices) [50]. Along with this, many flower images were procured from internet resources. These were true-color photos with varied resolution. Since we have not managed the image acquisition and camera activity, the images in the dataset having completely different distinction and illumination. Therefore, it is very much required to apply a correct pre-processing technique(s). The final dataset considered for experiment consists with 25000 flower images with 102 different species.

B. Experimental Results

The experiment was carried out using MATLAB R2018a, on an Intel (R), Core i3 - 3110M Central Processing Unit @ 2.40GHz contains 4GB Memory and x64-based processor, Windows 10. The images of the dataset is resized with dimension of 250 × 250 pixels. After applying the appropriate pre-processing and segmentation techniques, the features are extracted and the dataset is prepared to apply the proposed MKL + SVM algorithm with the integration of multi-label power dataset. Individually, for 102 classes (flower species), the classification of flower images with ANN, KNN, SVM and Random forest achieves some accuracy, whereas the combination of MKL and SVM achieves better classification accuracy compare to other classifiers with multi-labeling. As per represented in the figure 13, the prediction model is predicted the flower botanical name, its family and whether the flower is herb flower or not.

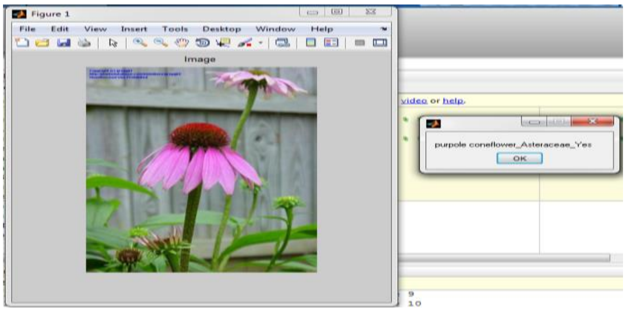

Fig. 12 the output of Prediction Model with flower botanical name, family of flower and whether the flower is herb flower or not

C. Performance Analysis

To measure the performance of the classification algorithms can be obtained through accuracy, precision, recall, and F-measure.

TABLE 3
Description of performance analysis

Sr. No.	Performance Analysis	Formula
1.	Accuracy	TP+TN / TP+FP+FN+TN
2.	Precision	TP / TP+FP
3.	Recall	TP / TP+FN
4.	F-measure	2*(Recall*Precision)/(Recall + precision)

The following table 4 depicted the performance of KNN, Random Forest, ANN, SVM and MKL+SVM with its accuracy and other parameters. 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.

TABLE 4
Performance Analysis with various classifier

Algorithm	Accuracy	Precision	Recall	F-Measure
KNN	0.21	0.195	0.203	0.201
Random Forest	0.69	0.687	0.690	0.688
ANN	0.71	0.705	0.708	0.706
SVM	0.72	0.720	0.721	0.720
MKL-SVM	0.78	0.775	0.770	0.774

V. Comparison with Previous Work

A comparative study with numerous other research has given in Table 5. It shows the details of the number of feature extracted, the size of the dataset, and the classifier used with its accuracy.

TABLE 5
A comparative study

Title	Author	Species	Total Images	Feature Extraction	ML Tech.	Acc. (%)
[51]	Elena et al. (2008)	102	8189	•Histogram of gradients • SIFT on the foreground boundary • SIFT on the foreground region • Color	Multiple kernel classifier	72.8
[52]	Elena et al. (2009)	102	8189	• Color • Texture • Shape	SVM	76.3
[53]	Kanan et al. (2010)	102	8189	• Color	SVM	71.4
[54]	Kubota et al. (2010)	102	8189	• Color • Texture • Shape	NIMBLE	53.9
[55]	Chai et al. (2011)	102	8189	• Color • Shape	SVM	80.0
[56]	Khan et al. (2011)	102	8189	• Color	LIBLINEAR	73.3
[57]	Almogdady et al. (2018)	102	8189	• Color • Texture • Shape	ANN	81.6
Our Method	The Proposed Approach	102	25000	• Color • Texture • Size • Disc flower • Petal type • No of petals •Aestivation of petal • Corona • Flower class	MKL-SVM	78.3

In order to study the efficiency of the proposed method with previously carried out work, we conducted a qualitative analysis of the results. The comparative analysis is depicted in Table 4. We have compared the performance of the proposed method on the 102-class flower dataset with 25000 images.

The flower features extracted in proposed method are Color, Texture, Size, Disc flower, Petal type, No of petals, Aestivation of petal, Corona, Flower class. The MKL-SVM classifier has given a recognition rate of 78.3%. It has been observed that, most of the researchers only considered basic flower features and the dataset taken into consideration is less. It should be noted that the combination of all 14 features even offers worthy accuracy than the discrete feature themselves.

VI. CONCLUSION

In this paper, the identification and classification of flower images with its species is discussed. A dataset is accumulated that contains 25000 flower images of 102 classes. Basic and morphology features of flower images are extracted using computer vison techniques with image pre-processing and image segmentation methods. A prediction model using machine learning supervised algorithm i.e. SVM is built with the integration of MKL and multi-labeling which can be helpful for the non-botanist's to classify different flower species easily, and the process becomes very fast, helping them in further research and study. It is observed the proposed approach achieved relatively a good classification accuracy with optimum possible extracted features.

REFERENCES

1. Jana Wäldchen, Michael Rzanny, Marco Seeland, Patrick Mäder, "Automated plant species identification—Trends and future directions", Pols, Computational Biology, April 5, 2018.
2. Andre C P L F de Carvalho and Alex A. Freitas, "A Tutorial on Multi-Label Classification Techniques, 05 May 2014.
3. Serhat S. Bucak, Rong Jin, and Anil K. Jain, "Multiple Kernel Learning for Visual Object Recognition: A Review"
4. Nilsback Maria-Elena, A. Zisserman, in: CVPR'06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1447–1454.
5. Y. Yoshioka, H. Iwata, R. Ohsawa, S. Ninomiya, Euphytica 139 (2004) 179–186.
6. T. Saitoh, K. Aoki, T. Kaneko, in: 17th International Conference on Pattern Recognition, ICPR'04, pp. 27–30.
7. T. Saitoh, T. Kaneko, in: 15th International Conference on Pattern Recognition, ICPR'00, pp. 507–510.
8. Fadzilah Siraj, Muhammad Ashraq Salahuddin and Shahrul Azmi Mohd Yusof," Digital Image Classification for Malaysan Blooming Flower" IEEE-2010.
9. Pavan Kumar Mishra, Sanjay Kumar Maurya, Ravindra Kumar Singh, Arun Kumar Misral "A semi-automatic plant identification based on digital leaf and flower Images" IEEE-2012.
10. Tanakorn Tiay, Pipimphorn Benyaphaichit, and Panomkhawn Riyamongkol" Flower Recognition System Based on Image Processing" ICT-ISPC-2014.
11. 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.
12. Warisara Pardee, Prawaran Yusungnern, Peeraya Sripian, "Flower Identification System by Image Processing", Conference Paper · August 2015.
13. Dr.K.Thangadurai, K.Padmavathi, "Computer Vision image Enhancement for Plant Leaves Disease Detection", World Congress on Computing and Communication Technologies, IEEE, 2014, pp. 173-175.
14. S.S. Agaian, K.P. Lentz, A.M. Grigoryan," A New Measure of Image Enhancement", International Conference on Signal Processing & Communication, 2000, pp. 19–22.
15. 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.
16. Sheikh Tania and Raghad Rowaida, "A Comparative Study of Various Image Filtering Techniques for Removing Various Noisy Pixels in Aerial Image", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.3 (2016), pp.113-124.
17. 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.
18. Monika Maru, M. C. Parikh, "Image Restoration Techniques: A Survey", International Journal of Computer Applications, Volume 160 – No 6, February 2017, pp. 15-19.
19. Anelia Angelova, Shenghuo Zhu, Yuanqing Lin, "Image segmentation for large-scale subcategory flower recognition", IEEE, 2013.
20. Asma Najjar, Ezzeddine Zagrouba, "Flower image segmentation based on color analysis and a supervised evaluation", ResearchGate, IEEE, 22 May 2014, pp. 402-406.
21. Nida M. Zaitoun, Musbah J. Aqel," Survey on Image Segmentation Techniques", International Conference on Communication, Management and Information Technology (ICCMIT 2015), Science Direct, 2015.
22. Isha Patel, Sanskruti Patel, "Analysis of Various Image Segmentation Techniques for Flower Images", JASC: Journal of Applied Science and Computations, Volume VI, Issue II, February-2019, pp. 1936-1943.
23. Samina Khalid, Tehmina Khalil, Shamila Nasreen, "A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning", Science and Information Conference, London, UK, August 27-29, 2014, pp. 1-8.
24. Mary Walowe Mwadulo, "A Review on Feature Selection Methods for Classification Tasks", International Journal of Computer Applications Technology and Research, Volume 5– Issue 6, 2016, pp. 395 - 402.
25. Syed Inthiyaz, B.T.P. Madhav and P.V.V. Madhav, "Flower segmentation with level sets evolution controlled by colour, texture and shape features", Cogent engineering, 28 April 2017, pp. 1-15.
26. Dibya Jyoti Bora, Anil Kumar Gupta, Fayaz Ahmad Khan, "Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation", International Journal of Emerging Technology and Advanced Engineering, Volume 5, Issue 2, February 2015, pp. 192-203.
27. Amira Ben Mabrouk, Asma Najjar and Ezzeddine Zagrouba, "Image Flower Recognition based on a New Method for Color Feature Extraction", pp. 201-206.
28. Takeshi Saitoh, Kimiya Aoki and Toyohisa Kaneko, "Automatic Recognition of Blooming Flowers", Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE, pp. 1051-4651.
29. Tzu-Hsiang Hsu & Chang-Hsing Lee & Ling-Hwei Chen, "An interactive flower image recognition system", Springer, Multimed Tools, Appl (2011), pp. 53–73.
30. Er. Gurleen Kaur Marwa, Er. Salam Din, "Development of Software for Flower Recognition Using Image Processing Technique", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 5, May 2015, pp. 4804-4811.
31. D S Guru, Y. H. Sharath, S. Manjunath, "Texture Features and KNN in Classification of Flower Images", IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR, 2010, pp. 21-29.
32. 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.
33. 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.
34. Anil Z Chitade, "Colour Based Image Segmentation Using K-Means Clustering", International Journal of Engineering Science and Technology, Vol. 2(10), 2010, 5319-5325.
35. Md. Rakib Hassan, Romana Rahman Ema & Tajul Islam, "Color Image Segmentation using Automated K-Means Clustering with RGB and HSV Color Spaces", Global Journal of Computer Science and Technology: Graphics & vision, Volume 1 7 Issue 2 Version 1.0 Year 2017.
36. Chomtip Pornpanomchai, "Leaf and Flower Recognition System (e-Botanist)", IACSIT, International Journal of Engineering and Technology, Vol.3, No.4, August 2011.

37. GURU D.S., SHARATH KUMAR Y.H., "SKELETON BASED APPROACH FOR FLOWER CLASSIFICATION", International Journal of Machine Intelligence, Volume 3, Issue 4, 2011, pp-191-198.
38. Y H Sharath Kumar, N Vinay Kumar, D S Guru," Delaunay Triangulation on Skeleton of Flowers for Classification", International Conference on Advanced Computing Technologies and Applications (ICACTA-2015), Science Direct, pp. 226 – 235.
39. Seiichi Uchida, "Image processing and recognition for biological images", Develop. Growth Differ. (2013), pp. 523–549.
40. https://socratic.org/questions/how-are-monocot-and-dicot-flowers-different
41. Elisa Veronese, Umberto Castellani, Denis Peruzzo, Marcella Bellani, and Paolo Brambilla, "Machine Learning Approaches: From Theory to Application in Schizophrenia", Computational and Mathematical Methods in Medicine, Volume 2013.
42. https://towardsdatascience.com/supervised-machine-learning-classification-5e685fe18a6d
43. Nupur Thakur, Deepa Maheshwari, "A Review of Image Classification Techniques", International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 11 | Nov -2017, pp. 1588-1591.
44. A Qur'ania and I Sarinah, "Identification of jasmine flower (Jasminum sp.) based on the shape of the flower using sobel edge and k-nearest neighbour", IORA-ICOR 2017.
45. Ravi Verma, Prof. J C Achutha, "Flower Species Classification using Random Forest Classifier", International Journal of Combined Research & Development (IJCRD), Volume: 5; Issue: 8; August -2016, pp. 1056-1060.
46. Thai Hoang Le, Hai Son Tran, Thuy Thanh Nguyen, "Applying Multi Support Vector Machine for Flower Image Classification", ICCASA 2012, Springer, pp 268-281.
47. Tomohiro SAKATA and Koichi KISE, "Flower Classification by Using Multiple Kernel Learning"
48. Jianqing Liang, Qinghua Hu, "Semi-Supervised Online Multi-Kernel Similarity Learning for Image Retrieval", IEEE Transactions On Multimedia, pp. 1-13.
49. Andre C P L F de Carvalho and Alex A. Freitas, "A Tutorial on Multi-Label Classification Techniques", Research Gate.
50. http://www.robots.ox.ac.uk/~vgg/data/flowers/102/
51. Nilsback M. Elena, Z. Andrew, in: ICVGIP'08: Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics and Image Processing, pp. 722–729.
52. Nilsback M E and Zisserman A (2009), an automatic visual flora: segmentation and classification of flower images (Doctoral dissertation, Oxford University).
53. Kanan C and Cottrell G (2010), robust classification of objects, faces, and flowers using natural image statistics. In Computer Vision and Pattern Recognition (CVPR), IEEE Conference on (pp. 2472-2479).
54. I to S and Kubota S, "Object classification using heterogeneous co-occurrence features", In Computer Vision–ECCV 2010 (pp. 701-714). Springer Berlin Heidelberg.
55. Chai Y, Lempitsky V and Zisserman A (2011), Bicos: A bi-level co-segmentation method for image classification. Master thesis
56. Khan F S, Weijer J, Bagdanov A D and Vanrell M (2011), Portmanteau vocabularies for multi-cue image representation. In Advances in neural information processing systems (pp. 1323-1331).
57. Huthaifa Almogdady, Dr. Saher Manaseer, Dr.Hazem Hiary, "A Flower Recognition System Based On Image Processing and Neural Networks", International Journal of Scientific & Technology Research Volume 7, Issue 11, November 2018.

AUTHORS PROFILE

Miss Isha Patel completed Bachelor of Computer Science from University of VNSGU, Surat, Gujrat in 2013 and Master of Computer Application (MCA) from NIRMA University, Ahmedabad in year 2016. She is currently pursuing Ph.D. and a Research Scholar in Faculty of Computer Science and Applications, Charotar University of Science and Technology, Changa since 2017. Her area of interest is Image Processing, Machine Learning, and Artificial Intelligence.

Dr. Sanskruti Patel is currently working as an Assistant Professor in Faculty of Computer Science and Applications, Charotar University of Science and Technology, Changa. She is having 13 years of experience including academics and research. She has published more than 25 research papers in various national/international journals. Her area of interest is Machine Learning, Deep Learning and Data Analytics.