Computer Vision Based Flower Recognition

BY

Asif Faysal
161-15-7340

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Md. Jueal Mia
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Anup Majumder
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2019

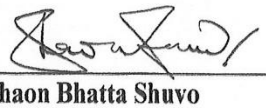
APPROVAL

This Project titled **"Computer-Vision Based Flower Recognition"**, submitted by Asif Faysal, ID No: 161-15-7340 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 5/12/2019.

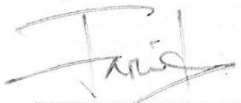
BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Saiful Islam — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Shaon Bhatta Shuvo — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dr. Dewan Md. Farid — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We thus announce that, this project has been done by us under the supervision of **Md. Jueal Mia, Lecturer, Department of CSE** Daffodil International University. We likewise proclaim that neither this undertaking nor any piece of this venture has been submitted somewhere else for honor of any degree.

Supervised by:

Md. Jueal Mia
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Anup Majumdar
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Asif Faysal
ID: -161-15-7340
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Md Jueal Mia**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Prof. Dr. Syed Akhter Hossain**, **Head,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University. We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Flower is the essential part of the beautiful nature. Its colors make the nature more colorful and balanced for the nature lover. Every day in our life, we have to go outside and so many things came front of us from the nature. Buildings, tress and so many things. Besides that, we also see some flowers on our walk way. May be those are in the rail line or in our balcony. But when we see them on the outside, we don't know the name of this flower. Sometimes we like the flower and then we wish to bring it on our garden or balcony. That time, we went to the Nursery and the seller told us the name of that flower. But the others flower that we don't want to bring in our garden we don't even know the name at all. That's why, we decided to do a research and develop a project on this. By this project, people will know the name of that flower which actually don't know by him. Our project will be working on mobile based camera for recognize the flower in real time. In our project we will use LMT (Logistic Model tree) which is a popular realm of machine learning, mainly used for image Classification and Regression analysis of any data. This is a little work we have done on this project but we are looking forward on it to make it more effective by upgrading our system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

We can say, Flower is the key of a beautiful nature because its color makes the nature more beautiful. Here, most of the plant in our nature we identify by its flower. Now, who is experienced such as Botanist or a nursery seller can identify a flower but a normal person can't do that. They have to find it on any web page or have to work on it with any other way. For this reason, our system can recognize those flowers which we actually want to know through a mobile camera. Currently, our project based on 10 flowers. But, till now we are very positive to our research. That's why, we decided that Day by Day, we will work on it for add more flowers. Daily life, we see so many flowers in our parks, houses, roadsides. But we don't know the name of this flower. We can find it on website or any other guidebooks for flowers. But when we go there with a picture its quite difficult to find out. Because, if we see, then we can realize that, there are so many flowers with same color and shape but those name and species are different. So, this is so much difficult to for a human to memorize those names of the flowers. This project will help for the recognition the flower by the real time.

1.2 Motivation

Whenever we start a project or work, there must be a motive behind this. I am a traveler. Love to travel for find out something new from the nature. For this reason, I have gone too many cities of BANGLADESH. Last year I went to JESSORE and I saw, so many flowers market out there. But I don't know the name most of them. So, I asked to the seller and then he told me the names. That time there are more traveler like me and foreigner also travelling there and they also facing the same problem. Then, I have decided to work on it. When I've returned, I have started work on it and then my title defense came and I choose to do my final year project on it. Now, we are

doing a research project that can identify the flowers automatically with their Scientific name. Hope this research project help those curious minds like me who want to learn new thing day to day.

1.3 Rational of The Study

When, we finally determined, that we are going to do this project, we have started our work for collecting information about this project. We start searching that, if there is any work done before like this which can identify the flower automatically. But our bad luck, there is so much short resources we found. There are lot of work should be done with food detection. But there are nothing much about flower. Then we took this project as a challenge for us and start our work which is identify the flower in real time. We took the challenge and now our challenge came true. Now, by this project we can identify about 10 of Bangladeshi common flowers with some foreign flowers also. For some kind of flower, this project can evaluate 100% accuracy rate in real time object. When we run the project, this will show the Scientific name. For doing this project we have used LMT (Logistic Model Tree) method. For the coding we have use the MATLAB for it.

1.4 Research Question

When we start and end up with this project there are few questions on our mind. Those are:

1. Can a system identify the different colors of same flower with the machine learning? Ex: Red Rangan/Purple Rangan.

2. Is LMT (Logistic Model Tree) method is efficient for doing this image processing project which can identify the flower accurately?

1.5 Expected Outcome

1. We can identify Bangladeshi along with some foreign flowers with this project.
2. LMT (Logistic Model Tree) which is the method we have used on this project. This method is used on Extreme cases of identifying any object. The average confidence level is 90% and sometimes it gives about 100% of accuracy rate for some kind of flowers.
3.There are almost 1000 images of data set which are taken in real life shot.

CHAPTER 2
BACKGROUND STUDY

2.1 Introduction

From the beginning of the era of science, Computer is the best inventions of it. Since of its invention it becomes our daily life hacker. From morning to night, it appears all the sectors of our daily life. It makes our life easier day by day by doing his operations. Machine Learning is the process, that develop the computer itself for think like a human being and do a job like a human. There are so many Applications done by the LMT method. Mostly it's used on an extreme cases image classification and its specially used for Flower Recognition System. For training on this project, we have used about 1000 data set which have taken in real life shot. That's why our data set is more reliable to work with it.

2.2 Related Work

We have used LMT image processing method for our Flower recognitions system project. There are some other projects like this which is developed till now. Recent history of modern software technology, this image processing technique is using in so many sectors. Police used this face detection technique for identifying the criminal. Traffic police find out the vehicles with this Image processing technique. Facebook use AI to make sure the given picture info of the users is correct or wrong.

There are some works also done on flower recognition sector also:

1. Flower recognition using the Shape and Size by the Morphological Technique.
2. Convolutional Neural Network (CNN) system is used for flower recognition system using the Pixels.
3. Gender Recognition System.
4. Color and Texture binary patterns.

5. Color and GIST features.
6. Flower Species recognition.
7. Deep learning.
8. Leaf for medicinal purpose
9. Classification system using LBP and SURF
10. Multiclass support vector machine

2.3 Research Summery

To implement this project, we have to use MATLAB, an open-source dataflow and machine learning library, to build an image classifying Logistic Model Tree (LMT) for classifying the flower image. For the clustering, we have used K-means clustering. After examining, inquiring, researching and comparing all the related works on Flower Identification, we come to have different and many types of conclusion. First of all, none of them are using real data for the detection and identification. All of them are just collecting the data from the google images. That's why sometimes its didn't work properly on the data base. Secondly is accuracy rate, most of the big project hit the accuracy rate up to 90% so they say. But the community feedback shows the different color. Users complain about the miss detection of plants and less information about it. It also takes long time for analyzing and searching their database.

2.4 Scope of The Project

❖ Developers

This project is finally done with LMT method with the real data set. So, if someone wishes to work on this field which is flower recognition system based on image processing then this is fully ready to go. Also, who want try it on CNN then must be good knowledge about it. It could be an Android based project also.

❖ Scientist or Botanist

There are almost 1000 data set of a flower. So, this will be very helpful project for the botanist or biologist for the building a unique data set. This project helps the scientists for classifying a large amount of data and generating a statistic for the data set.

❖ Traveler or Nature Enthusiasts

If there anyone like me, love to travel and find out something new than this project help him a lot. By this project, he can able to know the name of the flower. This will be really helpful for the Traveler and Nature Enthusiasts.

2.5 Challenges

1. We have used LMT method for this project. When we are performing our training and testing session there are some issues came out about identifying the some of flowers which are look like same by Size and Shape. We are doing this research project further day for solving this issue.

2. Another challenge is, identifying same flower of different colors. EX: Red Rangan, Purple Rangan. This was a tough work on LMT method.

3. Another challenge was collecting the data. We decided that, we will use the real data took on real life shot. That's why we are searched so many Nursery but season didn't help us. Because in Bangladesh the flower bloom in Winter. But it was Summer. So, this is the toughest challenge for us to collect the data.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Basic Introduction to Machine Learning

If we look deeply in our living circle, then we can realize that, we are living in an era of Artificial Intelligence (AI). Where, In Artificial Intelligence (AI), Learning is a very important feature. After invention of science, so many scientists make a definition of learning. One of them is (Simon 1987) who gave a definition that-Learning which is an adaptive skill, that can use as the same process better on next time. Later on, Mr. Feigenbaum (1977) says that, Learning is the only process that can use for the collecting Knowledge. Although, still there is no definite definition of learning skills but we have to define what is actually mean of Machine Learning. In 1997, Mr. Mitchell give a general thought about Machine Learning which is-Machine learning is used for improve the computer algorithm which can be perform automatically through the experience.

Based on the uses of Machine Learning, we can say that this is the most important field on Artificial Intelligence. When the developers start work on AI, there are some functions didn't work properly. One of them is thorough learning ability. That's why whenever the computer faces any kind of problems, the program stops working because of that time the AI is not self-adjusting. On the other hand, this time computer is not capable for collecting data and discover new knowledge. For this only one reason, computer can perform on that specific command which is given to it. It does not have any kind of ability to discover a new Logical Theory, rules and something like that.

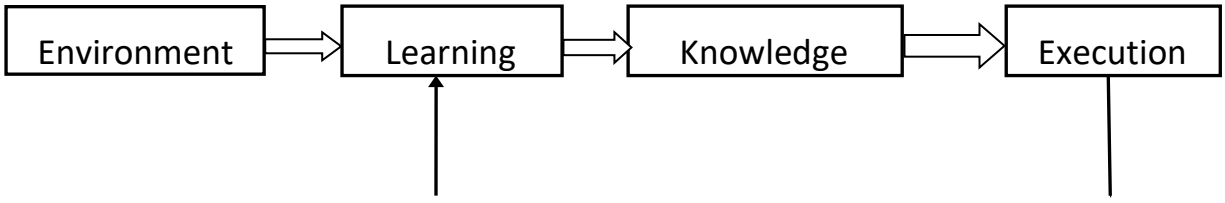


Figure 3.1: Learning System Structure

3.2 Logistic Model Tree

In the sector of computer science, a Logistic Model Tree((LMT) is a classification method with the associated supervised training algorithm which combines Logistic Regression((LR) and Decision Tree. This model trees are depending on the previous idea of model tree. In the logistic variant, the algorithm named LogitBoost, which is used to make an LR model at each node on the tree. Each of the LogitBoost supplication is warm started from its own results in the parent node. Last of all, the tree is pruned. The basic LMT method uses Cross-validation to find an amount of LogitBoost iterations which didn't overfit the training data.

For the clustering we have used K-means Clustering. This is an algorithm which is unsupervised and used for segmentation the interest area from the background. K-Means Clustering works for the unlabeled data.

The following figure showing the typical design of Logistic Model Tree:

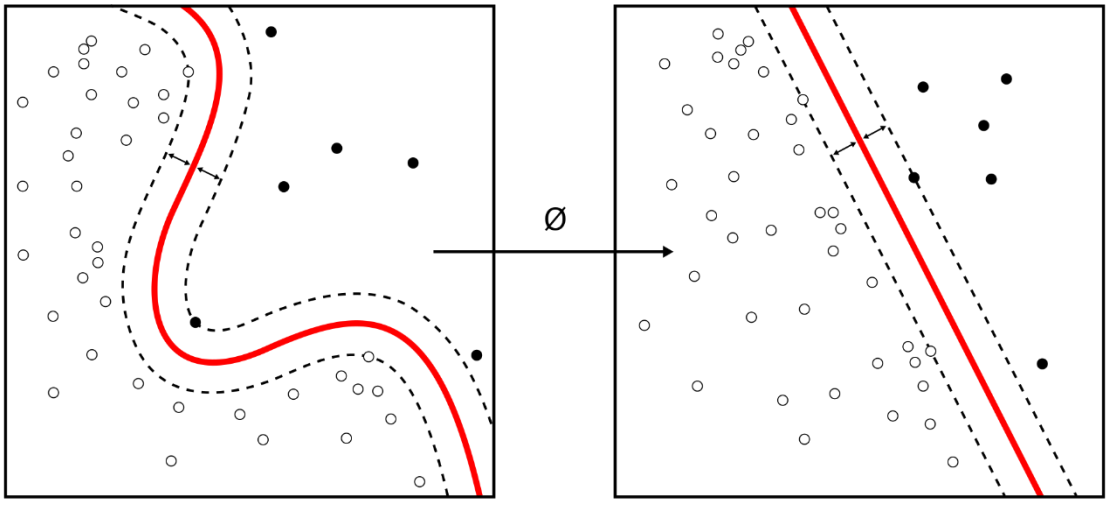


Figure 3.2: Typical LMT diagram

3.3 Data Collection Procedure

There are so many flower images data on Google. When we make a search on Google, we see that most of the researcher who has done a research on Flower Identification they are used the data from Google images. That time, they were faced a problem about image quality. Then we decided to collect the real data. We determined that, will be work on our own data set. That's why, we started searching the flower nursery in Dhaka city and finally found one at Savar area. But the season of that time, was Summer. So, there are small collection of flowers. Then we took the picture of the flowers averagely 100 each of them and use them as our data set.

The following figure shows the sample dataset image of Rangan flower.

Figure 3.3: Dataset of Rangan

3.4 Statistics of Data

The following table contains the statistics of raw data

Table 1: Statistics of Raw Data

Flower Name	Amount of Training Image
Allamanda Cathartica	150
Billbergia Pyramidalis	120
Catharanathus roseus	130
Joba	100
lantana Camara	150
Plumeria Rubra	120
Rangan	75
Rhododendron Ferrugienum	80
White Musanda	100

3.5 Implementation Requirements (Minimum)

❖ Software

- Windows / Linux / Mac OS
- Matlab
- Weka
- Microsoft Excel

❖ Hardware

- Intel i5 Processor
- 4 GB RAM
- 240 GB SSD

3.6 System Architecture

The procedure of system architecture of a machine learning-based expert system for identification the rare flower is shown in the figure 3.4. By Using this system, anyone can know the name of that specific unknown flower properly but before that they have to installed the application in the mobile first. Then, when the image will be taken, through the mobile application the user will send the picture to our expert system where the system is already placed on our back-end server. From the front-end software, picture will be received through the internet. Depending on the input picture, the system will send an confirmation to the user. Finally, user will get a response from the front-end mobile application.

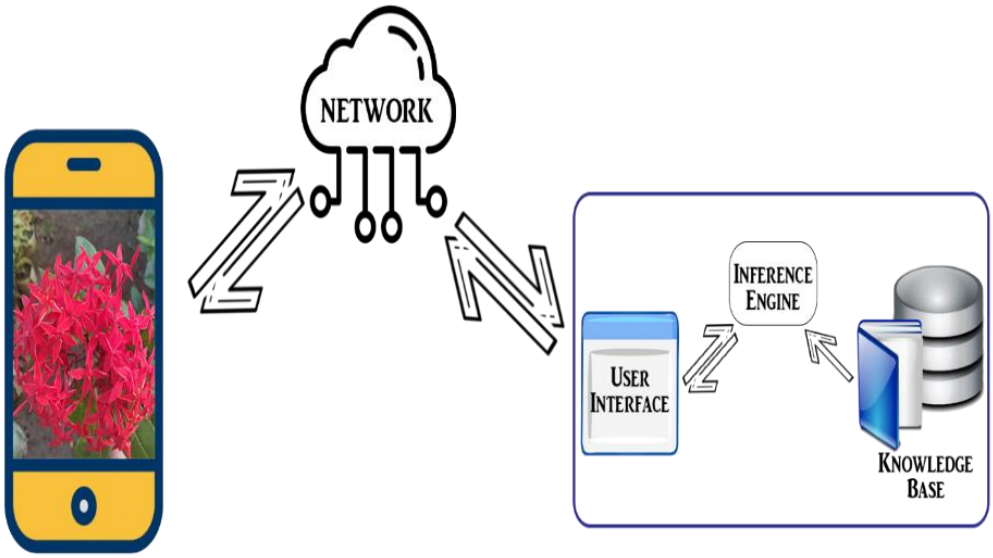


Fig 3.4: The architecture of machine vision based expert system for flower recognition.

3.7 Design Development

There are so many steps in image processing which are used for classification of flowers. Fig. 3.5 shows a machine vision-based solution framework for the recognition of Flowers.

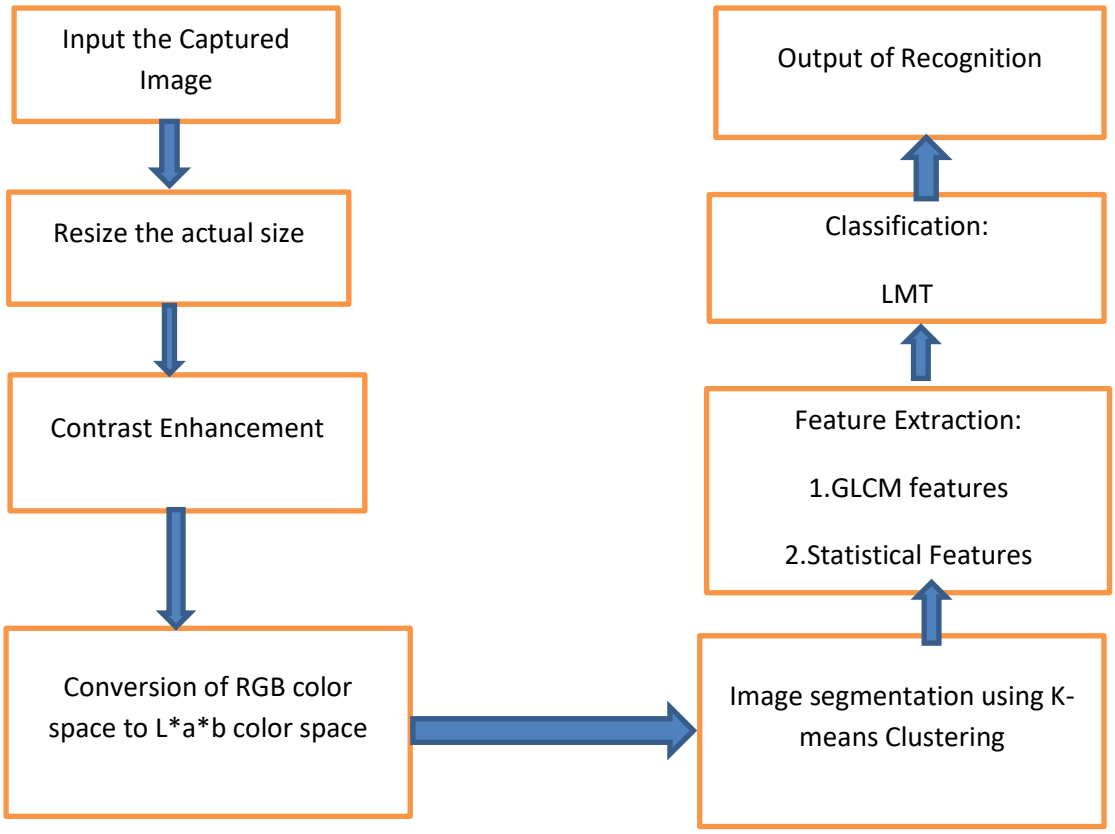


Fig 3.5: Methodology for local flower recognition system.

The system we are proposing starts with the image of flower. Flower data set has been created locally. At first, we have to use bicubic interpolation. Bicubic interpolation works with the color image of flower. The color image is transformed into a fixed-sized image. Let us assume intensity values as I, and p_x, p_y and p_{xy} are used as derivatives. (1, 1), (1, 0), (0, 1), and (0, 0) are four corners which can be recognized these derivatives of the unit square. Then the interpolated surface intensity is calculated with given equations:

$$p(x,y) = \sum_{i=0}^{3} \quad \sum_{j=0}^{3} \quad a_{ij} x^i y^j \ldots\ldots\ldots\ldots (1)$$

An image with good contrast can help as lot. So that, histogram equalization technique helped to develop the contrast of the image. Let,

R = the number of rows (height) in pixels,

C = the number of columns (width) in pixels

The number of pixels, which have intensity of color r_k, is n_k, and the total size of achievable intensity levels of color in the image is L. Then the color-mapped or prepared image is achieved by mapping each pixel with the intensity of color r_k into a similar pixel with color intensity s_k using the following equation.

$$s_k = \mathrm{T}(r_k) = \frac{L-1}{CR} \sum_{j=0}^{k} \quad n_j \ldots\ldots\ldots (2)$$

After completing contrast enhancement, RGB color space need to be changed. To change RGB color space it has to perform to acquire $L^*a^*b^*$ color space. et al. [9] said that k-means clustering segments image quite better in $L^*a^*b^*$ space than in RGB space. Transformation of RGB color space to CIE (International Commission on Illumination) XYZ color space as can be written as the following equation)

$$\begin{bmatrix} P \\ Q \\ Z \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \ldots\ldots\ldots\ldots (3)$$

From XYZ color space to transform into $L^*a^*b^*$ color space, P_n, Q_n, and Z_n are presumed as the tri-stimulus values of the reference white. The assumption can be written in the given way:

$$\int \quad f(t) = \begin{cases} t^{\frac{1}{3}} & \begin{cases} if\ t > 0.008856 \ldots\ldots\ldots\ldots\ldots (4) \\ if\ t \leq 0.008856 \end{cases} \\ 7.787\ t + \frac{16}{116} & \end{cases}$$

Then, according to the proposed technique of [10], the computation of L*, a*, and b* can be written as:

$$a^* = 500\left(f\left(\frac{p}{p_n}\right) - f\left(\frac{Q}{Q_n}\right)\right) \text{..... (5)}$$

$$b^* = 200\left(f\left(\frac{Q}{Q_n}\right) - f\left(\frac{z}{z_n}\right)\right) \text{..... (6)}$$

The mostly used *k*-means clustering method can help now to image segmentation of flower. After clustering, all the parts of the flower will be individually differentiated. Two types of feature vectors are obtained from the segmentation. Those two types are,

1. Statistical and
2. GLCM feature.

In the next section, two types of features are described in detail.

To apply LMT the expert feature vectors are used. In machine learning (ML) data classification is must. LMT is a machine learning (ML) algorithm. The aim of LMT is to provide huge accuracy. The capability of acquiring huge accuracy easily made it better than other classifiers. Reason behind using this method is, it can handle multiclass problem too. In our case the multiclass problem is our main problem. when the classifiers produce any output, two basic strategies are used.

- As the first strategy, classifier among 1 and the $K - 1$ other classes are trained (in entire classifiers K). This is called the 1-r (one-against-rest) approach.
- The one-against-rest (1-*r*) approaches including the number of LMT can be used to overcome this type of problem [12].

LMT are trained by using training data set. Finally, test data actually used to amount the performance of the classifier. Efficiency cannot be a true metric to analyze the performance of a classifier, and the reason is that may not be appropriately fitted for estimating classification types acquired from imbalanced data sets, i.e. the numbers of observations in different classes differ extensively. So, it can be state that to evaluate

the performance precisely of a classifier, not only calculating efficiency but also some other metrics [12] and [13] are important. For a binary decision the confusion matrix, i.e. 2-class problem summarizes the number of true positives (TPs), false positives (FPs), true negatives (TNs), and false negatives (FNs). In the multiclass purpose, i.e. longer than 2-class problems, the resulting confusion matrix dimension is $n \times n$ ($n > 2$). That is known that n is the matrix rows and n is the matrix columns and total entries are $n \times n$. Using that matrix, the number of TPs, FPs, TNs, FNs does not calculate shortly. According to the aforementioned procedure, the values of TPs, FPs, TNs, FNs for class i can be calculated: …

$$TP_i = a_{ii} \quad \text{(7)}$$

$$FP_i = \sum_{j=1, j\neq i}^{n} a_{ji} \quad \text{(8)}$$

$$FN_i = \sum_{j=1, j\neq i}^{n} a_{ij} \quad \text{(9)}$$

$$TN_i = \sum_{j=1, j\neq i}^{n} \sum_{k=1, k\neq 1}^{n} a_{jk} \quad \text{(10)}$$

By using confusion matrix

i. Accuracy
ii. Precision
iii. Specificity
iv. Sensitivity
v. FPR (false positive rate) and
vi. FNR (false negative rate)

of the expert system are calculated [11]. We divided our data into two separate sets. Those are,

a. Training data set
b. Testing data set.

We should keep the highest number of data in our training data set.

As per our expert system, data can be trained by the LMT. Test data set can help the performance of the system. This performance is measured in terms of these metrics. In

this way, LMT can classify the name of the flower from the image. From the confusion matrix accuracy, specificity, precision, sensitivity, FPR, and FNR are calculated in percentage can be evaluate as follows:

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \times 100\% \quad \text{..... (11)}$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad \text{..... (12)}$$

$$Specificity = \frac{TN}{FP+TN} \times 100\% \quad \text{..... ..... (13)}$$

$$\text{Sensitivity} y = \frac{TP}{TP+fN} \times 100\% \quad \text{..... (14)}$$

$$FNR = \frac{FN}{FN+TP} \times 100\% \quad \text{..... (15)}$$

$$FPR = \frac{FP}{FP+TN} \times 100\% \quad \text{..... (16)}$$

In order to estimate the overall performance of our classifier with the evaluation metrics equations are shown in (11), (12), (13),(14),(15) and (16).

3.8 Description of Local Flowers and Features

3.8.1 Description of Local Flowers

A large variety of tropical and sub-tropical flowers abound in Bangladesh. There are lots of flowers in our country. But now most of the flowers are becoming rare now. So, it is high time for retrieving the cultural heritage of Bangladesh. We use 10 flowers 'Allamanda Cathartica', 'Billbergia Pyramidalis', 'Catharanathus roseus', 'Joba', 'lantana Camara', 'Plumeria Rubra', 'Radha Chura', 'Rangan', 'Rhododendron Ferrugienum', 'White Mussaenda' as shown in Fig. 3.6 in our experiment.

- **Allamanda Cathartica:** Allamanda species mainly used in medicine purposes. This used for the treatment of liver tumors, jaundice, splenomegaly, and malaria.
- **Billbergia Pyramidalis:** This bromeliad is grownup for its flowers. Basically, this used

as an undergrowth plant or as a ground cover. This also used as a bottle plant and in one season it establishes from division. It can stand on a very little cooler weathers and is very easy to produce from division.

- **Catharanathus Roseus**: For herbal medicine and ornamental plant sector this species cultivated for so long.

- **Hawaiian Hibiscus (Joba)**: Hibiscus species represents so many countries national symbols, national flower. This is a national symbol of Haiti and the both Solomon Islands and Niue's national flower is this. Hibiscus which is the national flower of South Korea, and the national flower of Malaysia is Hibiscus Rosasinesis.

- **Lantana Camara**: L. camara has been used for so many herbal medicines treating like cancer, skin itches, leprosy, rabies, chicken pox, measles, asthma and ulcers.

Figure 3.6: Dataset of local flowers

- **Plumeria Rubra**: Plumeria Rubra testified to have anti-inflammatory, antioxidant, hepatoprotective, antimicrobial activities, anti-fertility and antioxidant.

- **Radha Chura**: It is an outstanding ornamental plant, usually grown in domestic and

public gardens and has a beautiful flowering in yellow, red and orange. This little size of flower is used for attracting the hummingbirds.

- **White Mussaenda**: They are well known to the African and Asian countries. Quite a lot of species are cultivated as an ornamental plants.

3.8.2 Description of Features

In this paper, we have selected two feature set, one is the gray-level co-occurrence matrix (GLCM) and the other one is statistical features to recognize the flowers. Statistical features value in the framework of textile fabric defect recognition is shown in [14]. From that paper, we have selected some statistical feature to identify the flowers. All the structures are existing in [14]. All features are defined exactly here.

- Mean (μ): If we signify object regions by *N* quantity of pixels, Background or object free regions by *M* quantity of pixels, and *GS* is the gray-scale color strength of a pixel in object regions. There will be an equation like:

$$\mu = \sqrt{\frac{\sum_{i=1}^{N} GS_i}{N}} \quad \text{..... (17)}$$

- Standard deviation (σ): If there is an object regions by *N* quantity of pixels, where μ , which is the mean gray-scale color intensity and on the other hand *GS* represent the gray-scale color intensity of a pixel separately, then the specific equation will be:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (GS_i - \mu)^2}{N}} \quad \text{..... (18)}$$

- Variance (σ^2): If we denote the object regions by *N* number of pixels, where μ is the mean gray-scale color intensity and *GS* represents the gray-scale color intensity of a pixel individually, then the variance equation is presented as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(GS_i - \mu)^2}{N} \text{ ..... (19)}$$

- Skewness (γ): If we represent deviation, mean, and mode of color intensity of all pixels of the gray-scale image in object regions are using *Mo*, μ, and σ individually, then

$$\gamma = \frac{\mu - Mo}{\sigma} \text{ ..... (20)}$$

In this work, we used 5 GLCM features, namely correlation (ρ), energy (E), contrast (C), homogeneity (H), and entropy (S). All of their corresponding equations are shown below from (19-23).

$$\text{Contrast: } C = \sum_{i=0}^{L=1} \sum_{j=0}^{L=1} (i-j)^2 p(i,j) \text{ ..... (21)}$$

$$\text{Correlation: } \rho = \frac{\sum_{i=0}^{L=1} i=j=p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \text{ ..... (22)}$$

$$\text{Energy: } E = \sum_{i=0}^{L=1} \sum_{j=0}^{L=1} P(i,j)^2 \text{ ..... (23)}$$

$$\text{Entropy: } S = -\sum_{i=0}^{L=1} \sum_{j=0}^{L=1} P(i,j) \log p(i,j) \text{ ..... (24)}$$

$$\text{Homogeneity: } H = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{p(i,j)}{1+(i-j)^2} \text{ ..... (25)}$$

Where μ_x, μ_y, σ_x and σ_y are the sum of required and variance values for the row and column matrix entries, respectively.

CHAPTER 4
EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction to Image Classification

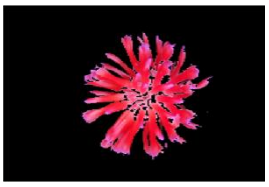
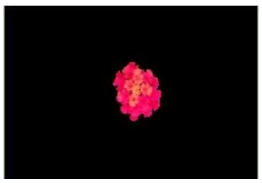
We have used LMT method for doing this project. This method is well known for the Image classification. For training and testing we have used MATLAB software. This is a commonly used software for these types of project. After that, For the clustering, K-means clustering used on this project. Lastly, when we completed our testing and training then we have used WEKA software for the accuracy test. Finally, with use of those methods and software we have completed our project

4.2 Experiment Result

Machine vision based expert system for flower recognition is shown in Fig. 3.4. At first, we have taken 1000 color images of 10 rare flowers that are captured by a different person with a different angle as well as a different size. Different size image is used for our experiment considering the different people. Then all the images are sent to our proposed system for further procedure as shown in Fig. 3.4.

Using color concentration mapping, contrast augmentation of the image is completed. Then using *k*-means clustering, color image is segmented into the number of clusters. Here, the reason of using k-means clustering because it works better than other existing segmentation algorithms. Thus, the percentage of the flowers are extracted from the background of the image. The stepwise changes of images for recognition of flowers is shown in Fig. 4.1. After all those things done, we can say that our expert systems delivered the better results. For getting the good quality of segmentation is required since it affects the performance of subsequent steps. Good quality of segmentation leads to good quality of feature extraction. This is why, good quality of segmentation is needed for improving the performance of the expert system. We have taken 1000 color images of 10 different local flowers for our experiment. Experimental outcomes from feature

extraction is shown in

	Train Image	Contrast Image	Segment Image
Allamanda Cathartica:			
Billbergia Pyramidalis:			
Catharanath s roseus:			
Joba:			
lantana Camara:			
Plumeria Rubra:			
Radha Chura:			

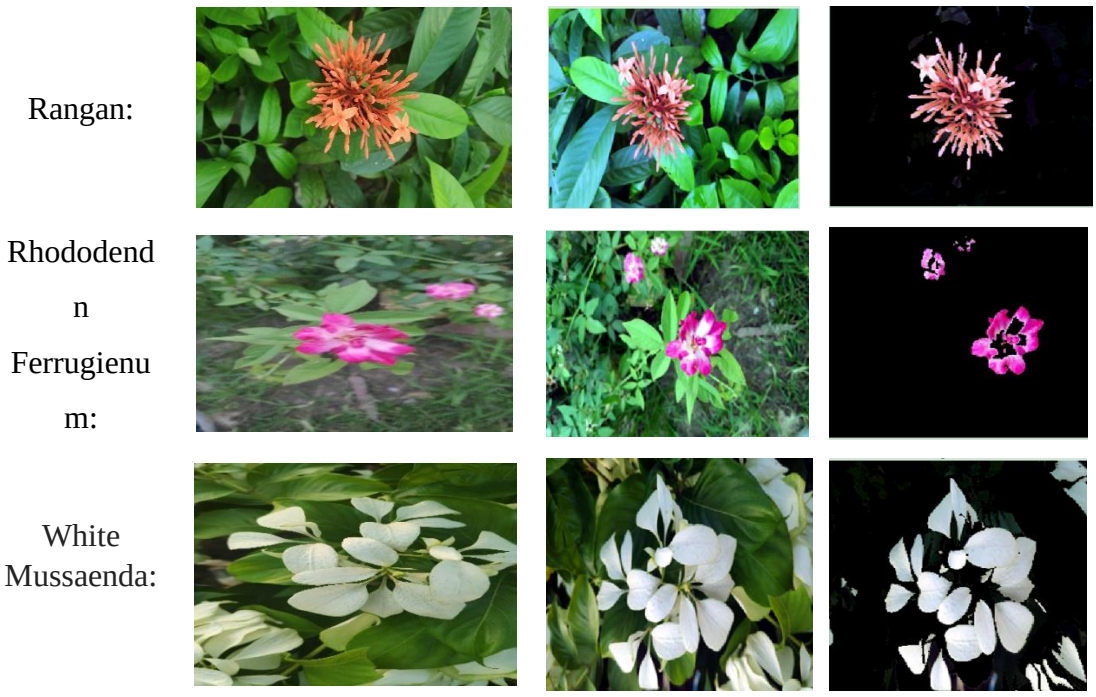


Figure 4.1: Stepwise changes of the image of each flower

Entire sample data are split into two parts, namely testing part and training part. For selecting the training and testing data, we used the holdout method. Based on the holdout method we choose about two-thirds data as training data set and the rest of them are used for testing sample data.

Figure 4.2: Extracted feature from some flowers, where most them are classified correctly and rest of them are misclassified

When we implementing the data set, there are many problems rises during experiments using those data. We have chosen the data justification set to avoid low training and simplification error. It is also known as overfitting problem. For solving this issue, we have divided the original training data set into two subsets. The first one is used for training data set, and the other one is used for data validation. From all the training data we have taken two-thirds as training set for building classifier and the rest of the image are used for error correction. To approach the highest performance classifier, we apply holdout methods for the five times repeatedly. After completing the training of the data set, we portion the performance of the system using the test data set. From all the results, we can calculate the average value for build multiclass confusion matrices.

In this viewpoint, we have to use the classifier. So, we used a linear SVM classifier. If the data set for training is {($\mathbf{x}1$, $d1$), ($\mathbf{x}2$, $d2$)….. ($\mathbf{x}84$, $d84$)}, where $\mathbf{x}i$ = (μ, σ, S, σ^2, κ, γ, C, ρ, E, H) is the input vector; di = ±1. For maximizing the objective function, Lagrange of multipliers {$\alpha 1$, $\alpha 2$,……, $\alpha 84$} is calculated as follows:

$$Q(a) = \sum_{j=1}^{28} a_i - \frac{1}{2}\sum_{i=1}^{28} \sum_{j=1}^{28} a_i a_j d_i d_j X_i^T X_j \text{……..(26)}$$

For measuring the performance of all the classifier, we have calculated average specificity, accuracy, precision, FNR, FPR. Also, we have calculated the accuracy level for all the flowers individually.

Table 2. Confusion matrix in binary structure for every local flower

Class	Matrix				Class	Matrix			
Allamanda Cathartica			Predicted Class		Billbergia Pyramidalis			Predicted Class	
			+	-				+	–
	Actual Class	+	26	3		Actual Class	+	30	0
		-	0	240			-	0	239
Class	Matrix				Class	Matrix			
Catharanathus roseus			Predicted Class		Hawaiian Hibiscus			Predicted Class	
			+	-				+	-
	Actual Class	+	29	2		Actual Class	+	27	2
		-	0	238			-	2	238
Class	Matrix				Class	Matrix			
Lantana Camara			Predicted Class		Plumeria Rubra	Actual Class		Predicted Class	
			+	-				+	-
	Actual Class	+	29	1		Actual Class	+	30	0
		-	1	238			-	0	239

Class	Matrix				Class	Matrix			
Rangan			Predicted Class		Rhododendron Ferrugienum			Predicted Class	
			+	-				+	-
	Actual Class	+	29	1		Actual Class	+	29	1
		-	1	238			-	2	237

We perceive from Table 2 that the accuracy level is 96.17%; which specifies the detection rate of our expert system is high enough. The average Sensitivity rate is 92.25%; that means the identifying rate of flowers name from the input is so good; and the FN rate is (7.75%), which is good. Here, Also the average specificity rate is 96.26%; that means flower identification rate is better enough; for this reason, FP rate is too low (3.74%); which is better enough for our experimentation. From our experiment, we prove that our system is working better than other present systems for identifying flowers.

Table 3. Results of metric-wise performance of LMT classifier used.

Metric	Value
Accuracy	96.17%
Sensitivity	92.25%
Specificity	96.26%
Precision	93.46%
False positive rate	3.74%
False negative rate	7.75%

In our experiment, we only described for only LMT method but for comparison the performance, we can implement other two classifiers which are SVM and Random Forest. Here, we found Comparative performance between those three classifiers is shown in Table 4. From all other classifier LMT method performs better in terms of confusion matrices, where SVM in the mid-level and Random Forest stands in last.

Table 4. Comparison of the three experimentally evaluated classifiers

Classifier	AUC
LMT	94.89%
SVM	90.31%
Random Forest	87.69%

As we say, LMT is showing better result than the SVM and Random Forest.

4.3 Experiment Summary

Our developed application can identify local Bangladeshi flowers with some foreign flowers also. We deeply focused on the accuracy level. The accuracy rate depends on the amount of data. So that we use more images with different angel to improve the confidence level. Currently some flower identifies with 100% confidence level. It is one of the success of our research and project. We used about 100 images for per flowers for training step. The dataset contains around 1000 flower images.

CHAPTER 5
CONCLUSIONS

5.1 Summery of The Study

Flower is a very important part of nature. Mostly we identify a plant through its flower. Experienced botanists do this identification of flower but a naive person will have to consult flower guidebooks or browse any relevant web pages on the Internet through keywords searching. This is a system that recognizes the flower. Presently this system can identify around 10 flowers. We are continuously working to add more flowers to identify. Every day we see a huge number of flower species in our house, parks, roadsides, in farms, on our rooftop but we have no knowledge of that flower species or their origin. Even we have no idea about its name. There are several guidebooks for flowers knowledge but it becomes quite difficult to find the name when have the picture. Even the Internet sometimes is not useful. But it is quite difficult for human brain to memorize all the species they see. Even some flower is similar to look at. This software recognizes the flower in real time by using mobile camera.

5.2 Future Work

This project has plenty of room for future work, by myself or a future interested student.

1. Improved the LMT design. There are much more research and practice are needed to optimize the design.
2. There are very short data set, if we think the whole sector of flower. So, we have to collect more data set and work on it.
3. We have to find those flowers which is actually look like same along with Color, Shape and Size. After that, we have to work on it separately.
4. Have to try on other Technique like CNN. Then, there will be something new will come.
5. Plannig to make an Android based app and put it on the Google Playstore very soon.

5.3 Conclusions

In present era, if we see that, the computer doing those things which is commanded by the human. Human make a command and then the computer work on this specific command. By the Machine Learning, Computers doing so many things, that one time those are un-imaginable. When the Artificial intelligence start to develop, that time if there any problem faced by the Computer it cannot be self-adjusting. Because of, for solving the issue there is no command given by the human. That's why, AI did not work on that time. Computer only can conduct with the existing truths. It does not have any ability to discover something new. So, the ultimate goal of this project is to find out the flowers name along with Scientific name and Uses of those flowers. For make this more easy and helpful, we will make an android based mobile app on it.

REFERENCES

[1] Xuanxin Liu, Fu Xu, Yu Sun, Haiyan Zhang, and Zhibo Chen, "Convolutional Recurrent Neural Networks for Observation-Centered Plant Identification," Journal of Electrical and Computer Engineering, vol. 2018, pp. 1–7, 2018.

[2] Yu Yang and Patric Granholm, "A study of pattern recognition of Iris flower based on Machine Learning, ". Turku University of Applied Sciences, Turku Finland , Bachelor's Thesis, 2013.

[3] Er. Gurleen Kaur Marwa and Er. Salam Din, "Development of Software for Flower Recognition Using Image Processing Technique, ". Dept. of SEEIT, COAET, PAU, Ludhiana, India,
vol. 3, Issue 5, May 2015.

[4] Vijaya Bylaiah, "Leaf Recognition and Matching with MATLAB, ". A Thesis Presented tothe Faculty of San Diego State University, Spring 2014.

[5] Y. Li, Q. Zhu, Y. Cao, and C. Wang, "A leaf vein extraction method based on snakes technique," in Proceedings of the International Conference on Neural Networks and Brain (ICNN&B '05), pp. 885–888, 2005.

[6] Bache, K.& Lichman, M. 2013. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[7] Bishop, C. 2006. Pattern Recognition and Machine Learning. New York:Springer, pp.424-428.

[8] Robin. 2010. Machine-Learning - Artificial Intelligence. Available at: http://intelligence.worldofcomputing.net/category/machine-learning. Consulted 22 Nov 2018.

[9] Burney, SM Aqil, and Humera Tariq. "K-means cluster analysis for image segmentation." International Journal of Computer Applications 96, no. 4 (2014).

[10] Color Conversion Algorithms. Available online: http://cs.haifa.ac.il/hagit/courses/ist/Lectures/Demos/ColorApplet2/t_convert.html#RGB%20to%20XYZ%20&%

[11] Confusion Matrix, Available online: https://en.wikipedia.org/wiki/Confusion_matrix

[12] Tax, David MJ, and Robert PW Duin. "Using two-class classifiers for multiclass classification." In Object recognition supported by user interaction for service robots, vol. 2, pp. 124-127. IEEE, 2002.

[13] J. Han, M. Kamber, and J. Pei, Data Mining Concepts and Techniques. Waltham, USA: Elsevier, 2012

[14] Md. Tarek Habib, M. Rokonuzzaman, "Distinguishing feature selection for fabric defect

classification using neural network J. Multimedia," 6 (5) (2011), pp. 416-424

[15] Haralick, R. M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Trans. Syst., Man, Cybern. SMC-3 (6), 610–621.

[16] L. J. Rozario, T. Rahman, M. S. Uddin "Segmentation of the Region of Defects in Fruits and Vegetables," International Journal of Computer Science and Information Security, vol. 14, no. 5, pp. 399-406, May 2016

[17] P.-N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. New York: Addison-Wesley, 2006.

Flower

ORIGINALITY REPORT

19%	10%	8%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Md. Tarek Habib, Anup Majumder, A.Z.M. Jakaria, Morium Akter, Mohammad Shorif Uddin, Farruk Ahmed. "Machine vision based papaya disease recognition", Journal of King Saud University - Computer and Information Sciences, 2018 Publication	5%
2	Submitted to Daffodil International University Student Paper	4%
3	ijircce.com Internet Source	1%
4	plantfileonline.net Internet Source	1%
5	www.vijayafarms.com Internet Source	1%
6	en.wikipedia.org Internet Source	1%
7	Submitted to Higher Education Commission Pakistan	1%

	Student Paper	
8	Submitted to Our Lady of Fatima University Student Paper	1%
9	tuetego.net Internet Source	<1%
10	www.plantsguru.com Internet Source	<1%
11	YANMIN SUN, ANDREW K. C. WONG, MOHAMED S. KAMEL. "CLASSIFICATION OF IMBALANCED DATA: A REVIEW", International Journal of Pattern Recognition and Artificial Intelligence, 2011 Publication	<1%
12	Habib, Md. Tarek, and M. Rokonuzzaman. "Selection of distinguishing features for fabric defect classification using neural network", 2010 13th International Conference on Computer and Information Technology (ICCIT), 2010. Publication	<1%
13	floralfridayfoto.blogspot.com Internet Source	<1%
14	etheses.whiterose.ac.uk Internet Source	<1%
15	ijarcsse.com Internet Source	<1%

16	www.floweradvisor.co.kr Internet Source	<1%
17	Submitted to University of Nottingham Student Paper	<1%
18	www.joetsite.com Internet Source	<1%
19	Submitted to VIT University Student Paper	<1%
20	Submitted to Sheffield Hallam University Student Paper	<1%
21	Li, Guanlin, Zhanhong Ma, and Haiguang Wang. "Development of a Single-leaf Disease Severity Automatic Grading System Based on Image Processing", Energy Procedia, 2011. Publication	<1%
22	Lokender Kumar, John Brice, Linda Toberer, Judith Klein-Seetharaman, Daniel Knauss, Susanta K. Sarkar. "Antimicrobial biopolymer formation from sodium alginate and algae extract using aminoglycosides", PLOS ONE, 2019 Publication	<1%
23	tel.archives-ouvertes.fr Internet Source	<1%
24	X He. "Identification of combustible material with	