MACHINE VISION BASED POTATO SPECIES RECOGNITION

BY

Md Nuruzzaman
ID: 162-15-8016

MD. Shahadat Hossain
ID: 162-15-7783
AND

Md Mostafijur Rahman
ID: 162-15-7863

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

Supervised By

MD. TAREK HABIB
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

MS. NAZMUN NESSA MOON
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

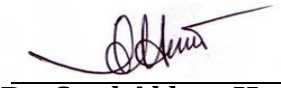
DHAKA, BANGLADESH

JULY 2020

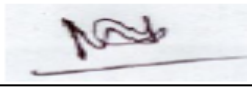
APPROVAL

This Project titled "**Machine Vision Based Potato Species Recognition**", submitted by **Md Nuruzzaman 162-15-8016**, **MD. Shahadat Hossian 162-15-7783** and **Md. Mostafijur Rahman 162-15-7863** 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 July 2020.

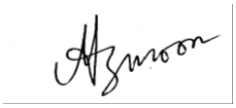
BOARD OF EXAMINERS



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



Dr. Md. Ismail Jabiullah **Internal Examiner**
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Nazmun Nessa Moon **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

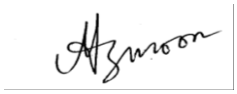
DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Md. Tarek Habib, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Md. Tarek Habib
Assistance Professor
Department of CSE
Daffodil International University

Co-Supervised by:



Ms. Nazmun Nessa Moon
Assistance Professor
Department of CSE
Daffodil International University

Submitted by:

Md Nuruzzaman
ID: 162-15-8016
Department of CSE
Daffodil International University

Md Mostafijur Rahman
ID: 162-15-7863
Department of CSE
Daffodil International University

MD. Shahadat Hossian
ID: 162-15-7783
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we must confess the grace of Almighty **Allah**, it was difficult for us to finish the work without His grace. Then I would like to say that with the care of the supervisor and all the others who have helped in different ways.

I would like to express sincere gratitude to my **supervisor, Md. Tarek Habib, Assistant Professor**, **co-supervisor, Ms. Nazmun Nessa Moon, Assistant Professor**, Department of CSE, Daffodil International University, Dhaka, who is truly a talented man and adapts himself to new technology. He encouraged me to do the project with his talent and sincerity and gave me the courage to finish the right job successfully even after the road had been hard. The goal of this project could not have been realized without his unwavering support. His contribution to my success today is unimaginable.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain**, **Head**, Department of CSE, for his kind help to finish our project and **Dr. Md. Ismail Jabiullah**, **Professor** and **Ms. Nazmun Nessa Moon**, **Assistant Professor,** Department of CSE, Daffodil International University, Dhaka and also to other faculty members and the staff of the 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, I must acknowledge with due respect to my parents for their constant support and our supervisor, for their kind co-operation and encouragement which helped us in the completion of this Research project.

ABSTRACT

One of the bound vegetables element of our daily life is potato. There are 4200 kinds of potato species and among them 82 kinds are found in Bangladesh. It is usually grown for numerous reasons apart from having as a food. Different potatoes are used in different cases. If there's a curry making procedure there will a different potato be used and if there's a French fried making recipe there also a different potato be used. But the problem is many people of our country fail to understand the right kind of potato for the right use. Thus, a big confusion is created in every genre of these potato consuming sector especially for the people who are unaware about potato species. This is where our project idea born, using a machine vision based recognition of these potato species which can help people to recognize them. In this paper, we perform an in-depth exploration of a machine vision approach for recognizing different species of potatoes of Bangladesh. A number of potatoes are classified based on the figures analyzed by their images. For our experiment, we collected the data of 4 verities total 1200 potato images. In this process, we wanted to identify the real image of potatoes. we have applied machine learning algorithms as like Random Forest Classifier (RF), Linear Discriminant Analysis (LDA), Logistic Regression, Support Vector Machine (SVM), CART, NB and KNN on our datasets. Here we apply our developed algorithms to (color, texture and shape) features. The data obtained from image processing were classified using extraction, resize and gray scale convert. After applying all algorithms each of them produced different results. Random Forest Classifier (RF) show best result and its accuracy rate 100% and SVM gives lowest rate. Its accuracy rate 74.074%, which is not only good but also promising for future research.

TABLE OF CONTENTS

CHAPTER 3: RESEARCH METHODOLOGY

CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSIONS

CHAPTER 5: CONCLUSION AND IMPLICATION FOR FUTURE RESEARCH

LIST OF TABLES

LIST OF FIGURES

CHAPTER 01
INTRODUCTION

1.1 Overview:

One of the plants of the nightshade family is the Potato (Solanaceae), It is usually grown for numerous reasons apart from having as a food, like its tenacious esculent tubers. Native potatoes has enormous species (more than 4000) and mostly of them are found in different countries like Andes, Specially in Bolivia, Peru, Ecuador, Colombia and Chile. These come in different shapes and sizes. This also includes the wild potato species which is more than 180. Even though they taste bitter, their vital biodiversity is tremendous which includes natural counteraction pests, malady other climatic states. Potato cultivars comes in different diversities of color, shape and size.

Table 1.1: Potato scientific classification chart

Scientific classification	
Kingdom:	Plantae
Clade:	Tracheophytes
Clade:	Angiosperms
Clade:	Eudicots
Clade:	Asterids
Order:	Solanales
Family:	Solanaceae
Genus:	Solanum
Species:	S. tuberosum
Binomial name	
Solanum tuberosum	

Potato plant shrubs are perennial kinds which is estimation heights are about 60 cm or 24 inches. Of course it depends on the variety, the flowers that dies after the formation of flowers, fruits and tubers. They usually contain pink, white, blue, purple or red flowers with yellow pollens. In general, different types of tubers come with white flowers mostly have skins like white, but colorful flowers have different types of pink skin.

1.2 Motivation:

The potato is a bound vegetable element in our daily life, the plant itself a perennial kind in the potato family of Solanaceae with a starchy tuber of the plant Solanum tuberosum. The rapid growing characteristics of potato allow poor families to cultivate it on small plots and break the cycle of poverty. Millions of people around the world depends on potatoes for their livelihood. Potatoes are grown almost every countries in some environmental circumstances like under temperate, tropical and some subtropical conditions as it is world's third most important food element after rice grain and wheat. China is now the largest producer of potatoes, followed by India, the Russian Federation and the United States.

Potato is a nightshade of the Solanum variety like tomatoes. The vegetable and fruit parts of the potato contains a toxic ingredient solanine which is harmful for human consumption. Ordinary potato tubers are which are properly grown in agricultural lands are preserved and produced sufficient amounts of glycoalkaloids, which are insignificant for human hygiene. The green parts of the plant (such as leaves, sprouts and skins) are directly exposed to sunlight, but a high concentration of glycoalkaloids of the tuber’s can accumulate, will throw.

Potatoes are mostly infected like any other plants by the insects such as grasshoppers, flees, bumbles which carries pollen from other plants, although a sufficient amount of self-fertilizer is given. Tubers are formed in the response to reduced the day length, although this bias has declined among in commercial species. Potatoes are a great food

that provides sugars, proteins, vitamins (mostly vitamin C, vitamin B6) and minerals (mainly iron, potassium). So now it's used in national dishes and most popular dishes.

- Different varieties of potatoes are now available in Bangladesh.
- Many people are ignorant of varieties of potatoes.
- Many people are deceived in shopping due to the ignorance of varieties of potatoes.
- Many people are unaware of Bangladeshi heritage of potatoes.

1.3 Objectives:

About 28 local varieties of potatoes are cultivated in different parts of the country. They have different well-known local names. It's estimated that between the year of 2004-2005, local varieties has cultivated potatoes on about 1,833,466 acres of land, of which 5, 99, 51 metric tons of tubers were produced. In Bangladesh's aspect, potato is commonly used as a vegetable in daily life. But is not same with other countries. Many countries of the world potato is the foremost food. In fact, it contributes more than 92% of the carbohydrate food source. Countless numbers of potatoes are processed annually in countries like Europe in dextrose, potato meal, alcohol, starch, and flour. Sometimes huge number of potatoes are processed into potato chips and french fries, processed into dehydrated mashed potatoes, and some into canned. Almost every Asian country uses more rice than potatoes to meet the shortage of sugar in the national diet. In this case, the countries of Europe are the opposite. In Bangladesh, however, potatoes, along with fish, meat and eggs, are mainly used to make curry. There is a variety of uses in different products made of potatoes. Among the potato dishes, many varieties have now become essential for breakfast e.g. boiled potato, fried potato, spicy potato and baked potato are notable. Nowadays, Bakery as well as the fast-food shops now have a variety number of delicious dishes made of potatoes to entice anyone such as french fries, potato chips, potato singara etc. They have started preparing various food delicacies based on potatoes. It has created a lot of employment as well as human needs. In Bangladesh, urban or rural

areas have a different impact. Different types of potatoes are used to make different types of food. And now the biggest problem is that we don't have much idea about potatoes. And we are using technology to solve this problem.

- To solve the problems of potatoes that identify them.
- To give a clear idea about them.
- To help Bangladeshi people in shopping.
- To help retain Bangladeshi heritage.

1.4 Expected outcome:

- Successful deployment of existing or new machine learning algorithm for recognizing potato varieties.
- Publication of one or more articles in international conference proceedings or journals.
- Reduction of deception in shopping.
- Building of a large data set for potato varieties in the context of Bangladesh.

CHAPTER 02

BACKGROUND STUDY

2.1 Introduction:

We have done research on this kind of work for others for the convenience of our work. In this we wanted to get an idea about the work of others in this regard. This kind of work is basically a lot of challenges. Related tasks, research summaries, problem areas and decision challenges. We compare our work with the work of others, we summarize a few research papers, related works, the underlying methods and the accuracy related to our work. We show summaries of some related work in the middle of the brief section in the study. The Problems section discusses the various problems of working time and the different solutions to the problems. Finally, in the Challenge section we talk about the various issues of our working time that made it difficult to do the work.

2.2 Related Works:

Habib et al. [1] presented a agro-medical expert system based on machine vision. It narrates a two-feature set for the papaya disease recognition solution consisting of a total of ten feature. Image processing techniques are implemented to extract some features. SVMs, C4.5, Naïve Bayes, Logistic Regression, K-NN, Random Forest, BPN, CPN, RIPPER using those techniques. The classification of papaya disease was done with SVM and then they analyzing relative quality of their work was determined by results of similar work. Achieved 90.15% accuracy, which is good as well as committed.

Zawbaa et al. [2] paper presents a technique which is used for recognizing fruit images automatically. The proposed model's formation is in three stages. The first stage is pre-processing, second is feature extraction and the third stage is classification. Feature extraction consists of two algorithms; The first algorithm is for shaping. Color algorithms uses color moments (color variance, color mean, color kurtosis, and color skewness) and shape (elevator number, centroid, and concentricity) for creating feature vectors. The second algorithm is known as the Scale Invariant Feature Transform (SIFT). The classification can work, once feature vectors are created for each image. Data set use

Orange-46 and Strawberry-55 and SIFT, K-NN ,SVM algorithm for methodology and best result on using the SVM apple - 90.91% orange -78.89% using SIFT apple - 96.97% strawberry - 85.71% as feature extraction.

Lak et al. [3] develop a more general algorithm in various natural lighting conditions of the main discussion. Through this, long-end light sources will be used for luminance controllers. The image was created by two algorithm and the images could be viewed with an apple. An Algorithm based on Edge detection, Algorithm based on Color-Shape and Gaussian low-pass filter, Canny (1986) method are used for classifier. To standardize the luminance of the obtain images, no light control was used. Here the edge detection based procedure did not seem triumphant but the Algorithm based on color-shape was capable of detecting 83.33% of those images.

Arivazhagan et al. [4] Try to develop a propose method that can be analyzed, identified and processed fruits based on features like color and texture. To improve the functionality and flexibility of the recognition system, features like size and shape can be combined with texture and color features. In fact, by increasing the number of images in the database the recognition rate can be increased. This algorithm can be used for service scales like smart selfie. Here they use 15 categories total 2633 image and algorithms use Minimum distance criterion (MDC) and statistical analysis. Using Color and Texture Features show best result on Cashew fruits 99.047%.

Hussain et al. [5] Work with Deep Convolution Neural Network (DCNN). Try to propose a novel fruit recognition classification system and the proposed method could recognized fruits images easily with many challenges. The proposed method has too efficiently growth the recognition rate to meet the requirements of actual word application. For data set them use 15 different categories 44406 images and use technology K-means clustering algorithm, CNN and DCNN for classification. The experiments were performed on their datasets and the results showed that taxonomy system had achieved a fair accuracy rate of 99%.

Rochaa et al. [6] this paper has reached the multi-class classification of many classes in such a way that a person can combine different features simultaneously and interact with the classifiers according to the part of the problem and It represents a unified solution which can be combined of many classifications and features. This type of strategy is required less training and better performance compared to a naïve method. The outcome reveals that the proposed solution can reduce the errors occurred by the classification by 15 percentage points(respect of the baseline). For data set they use 15 different categories 2633 images and use technology Linear Discriminant Analysis (LDA), K-means clustering algorithm, Classification Trees, Ensembles of Classifiers and Vector Machines (SVMs) show better result between naïve method.

Astuti et al. [7] another Automatic fruit classification system has been put and amplified. It is found that the existing method based on artificial neural network (Ann)'s accuracy is not as good as The accuracy of the SVM-based fruit classification method. Use 2 different categories images and technology KNN, Multilayer perceptron (MLP), SVM, artificial neural network (Ann) and result show SVM-based classification 100% accuracy. This was also distinct from simulation of our computer results that the proposed method required much less training than the existing one.

Naik et al. [8] here supervisions the initial process of fruit grade and classification. Feature extraction strait for shape, color, size and texture are properly expound with properties like HOG, LBP and SRF. Eventually some machine learning methods like Decision Tree, KNN, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), Principal Component Analysis (PCA), SVM, Fuzzy Logic and CNN are briefly discussed. 20 categories fruits used for experiment and best result show SVM (Tomato 92%), Fuzzy (Mango 90%) ANN (Orange 88%). Even though there are some appeals still need to be overcomed, Future's non-destructive fruit classification and grading will be Machine Vision .

Sekar L et al. [9] Machine vision fruit grading system has the capability to replace labor work to inspect fruit and morphological has given a accuracy rate which is highest among all of them. For grading, HIS (Hue, Saturation, Intensity) color model is frequently used

because it relatesd to human Psychology. In here use 5000 images data set and machine learning techniques showed the best result out of these techniques. On the other hand, Fuzzy is easy to implement but gave lowest accuracy rate result.

Zawbaa et al. [10] this paper presents automatic fruit classification different types of techniques on machine learning. The proposed classification system has pre-processing, feature extraction and collision-splitting order. At the pre-processing stage, the the fruit images is reshaped. Also it features Two algorithms which are used to extract the properties of the fruit images at the extraction stage. Random Forest (RF) algorithm is applied for classification. The outcomes of the Random Forest (RF) algorithm are compared k-nearest neighbors (KNN) and Support Vector Machine (SVM) algorithms. 178 type fruit images are used for data set and best result for KNN (71.42% orange and 72.72% strawberry) For RF (87.50% orange and 90.91% strawberry.

2.3 Research summary:

Table 2.1: Summary of previous researches

SL	Problem Domain	Data Set	Algorithm	Number of Feature	Accuracy
1	Machine vision based papaya disease recognition	129 color images of papaya	K-means clustering, LogisticRe-gression, Random Forest, KNN, SVM, Decision Tree, Naive Bayes	Contrast, correlation, energy, entropy and homogeneity	Accuracy has been 90.15% on using SVM
2	Automatic Fruit Image Recognition System Based on Shape and Color Features	Orange - 46 Strawberry - 55	SIFT, K-NN, SVM algorithm	Pre-processing, feature extraction,cla-ssification	Using the SVM-apple - 90.91% orange - 78.89% Using SIFT

					apple - 96.97% strawberry - 85.71% as feature extraction
3	Apple Fruits Recognition Under Natural Luminance Using Machine Vision	Apple –(25-80)	Edge detection based Algorithm, Color-Shape based Algorithm, Gaussian low-Pass filter, Canny (1986)method	Color, Shape, Edge detection	Color-shape based algorithm detect 83.33% of images
4	Fruit Recognition using Color and Texture Features	15 categories total 2633 image	Statistical Analysis, Minimum distance criterion (MDC)	Contrast, energy, local homogeneity, cluster shade and cluster prominence	Cashew fruits 99.047 Using Color and Texture Features
5	Automatic Fruit Recognition Based on DCNN for Commercial source Trace System.	15 different categories 44406 image	K-means clustering algorithm , CNN, Deep Convolution Neural Network (DCNN)	Color, shape	Accuracy rate of 99%

6	Automatic fruit and vegetable classification from images	15 different categories 2633 images	K-Means, Linear Discriminant Analysis (LDA), Support, Vector Machines (SVMs), Classification Trees, Neural Networks(NNs), and Ensembles of Classifiers	Shape, texture and color	Better than naïve method
7	Automatic fruit classification using support vector machines: a comparison with artificial neural network	2 categories fruits	KNN, Multilayer perceptron (MLP), SVM, artificial neural network (Ann)	Image preprocessing, Feature Extraction, Pattern Matching	SVM-based classificati on 100% accuracy
8	Machine Vision based Fruit Classificatio n and Grading - A Review	20 categories fruits	KNN, Decision-Tree, Principal Component Analysis (PCA), SVM, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), Fuzzy Logic	Maturity, Size, Volume Shape, Weight, Texture	SVM (Tomato 92%), Fuzzy(Mango 90%) ANN(Oran ge 88%)
9	Fruit Classificatio n System Using Computer	5000 images	Gray LevelCo-occurrence Matrices (GLCM), K means clustering method, Artificial	Colour, Morpholog-ical, Texture	GLCM(Chi ckoo 94%, Apple 93%, Sweet

	Vision: A Review		Neural Networks (ANN), Adaptive Neural Fuzzy Interference System(ANFIS).		lemon 93%, Orange 92% and Mango 92%)
10	Automatic fruit classification using random forest algorithm	178 type fruit images	KNN and RF algorithms, Support Vector Machine (SVM)	Shape, Colour characteris-tics and Scale Invariant Feature Transform (SIFT)	For KNN (71.42% orange and 72.72% strawberry) For RF (87.50% orange and 90.91% strawberry
11	Automated Fruit Classificatio n System	Different type	Canny method, Automated and embedded based due to system required several hardware, Support Vector Machine (SVM)	Size and the colour	Edge detection converted in to gray image and grey image detecting the edges of the fruit
12	Potato Classificati-on Using Deep Learning	1200 images	CNN, Computer model learns to perform classification tasks directly from images, text, or sound	Size re-scaling, rotations of 40, horizontal shift, image zooming, and horizontal flipping	Accuracy of 99.5%

2.4 Scope of the problem:

Previous works in other domains have predicted many Automated fruit classification and Grading system by using varies machine learning techniques and different types of algorithm. With the help of previous research and analysis of works, we have decided to work with Potato varieties detection and classification with machine vision.

2.5 Challenges:

Different varieties of potatoes were our biggest challenge. There are about 4,000 varieties of potatoes in the world. There are about 72 varieties of potatoes in our country. There is very little difference between such huge potato varieties .Can't catch people's eyes very easily. The differences between them are mainly the color, size, structure, small holes in the middle of the body, the upper covering, including a number of things. Over time, as the potatoes mature, they also change color. Potatoes, on the other hand, are mainly seasonal fruits and not all types of potatoes are available in all season's .And because some potatoes are more expensive to produce and not profitable in business, production is stopping. In general, for the structural properties of potatoes, it is difficult to diagnose with the human eye, but it is possible to diagnose through technology. A lot of this kind of work is happening now where machines are giving much more successful results than humans.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter, we will talk about data collection method, data preprocessing, proposed methodology and implementation requirements. First of all, we will discuss how we collected our data in the data collection method. Next, in the data preprocessing part, we have discussed how we pre-processed it for our model. Then in the proposed methodology, here we have discussed in detail those algorithms that we have used for our classification. Finally, the chapter is closed by a clear concept about what we used for the project.

3.2 Data Collection Procedure

There are about 81 varieties of potatoes in our country. We have basically collected the data of the potatoes that are commonly available in local market. We have collected about 1200 data set of four varieties of potatoes.

3.2.1 Image Samples

Show an image of the potato samples. Here we use four varieties of potatoes.

(a) (b) (c) (d)

Figure 3.1: Images of potato samples (a) BARI Alu-14 (Cleoptra), (b) BARI Alu-18 (Baraka), (c) BARI Alu-28 (Lady Rosetta), (d) BARI Alu-55 (Red Fantasy)

3.3 Data Preprocessing

After data collection there may be some inconsistent, missing and uncategorized data. We preprocessed our data by the sklarn library in two ways one is LabelEncoder and another is MinMaxScaler. Here, For normalizing lebels in a way that they only contain values between 0 and n, there is an utility class know as LebelEncoder. MinMaxScaler converts properties by scaling each feature to a specific range. The scaling and translating operation is done by the estimator for each feature such that it falls within the range which was provided by the training organization.

Image pre-processing steps:

- Read image
- Resize image
- Features descriptor
    - Gray scaling of image
    - Hu Moments: COLOR_BGR2GRAY
    - Haralick Texture: COLOR_BGR2GRAY
    - Color Histogram: COLOR_BGR2HSV

3.4 Proposed Methodology

We could do our research on different platforms like image processing, deep learning, machine learning, but we used a machine learning algorithms. Because of Machine learning is a branch of artificial intelligence connect with algorithm making that can metamorphosis itself except human intervention to acquire the desired results. We want to classify machine learning algorithms to help categorize such images. It usually requires structured data and these algorithms are designed to understand labeled data and

then work effortlessly on new data to give better results. That's why we applied machine learning algorithm.

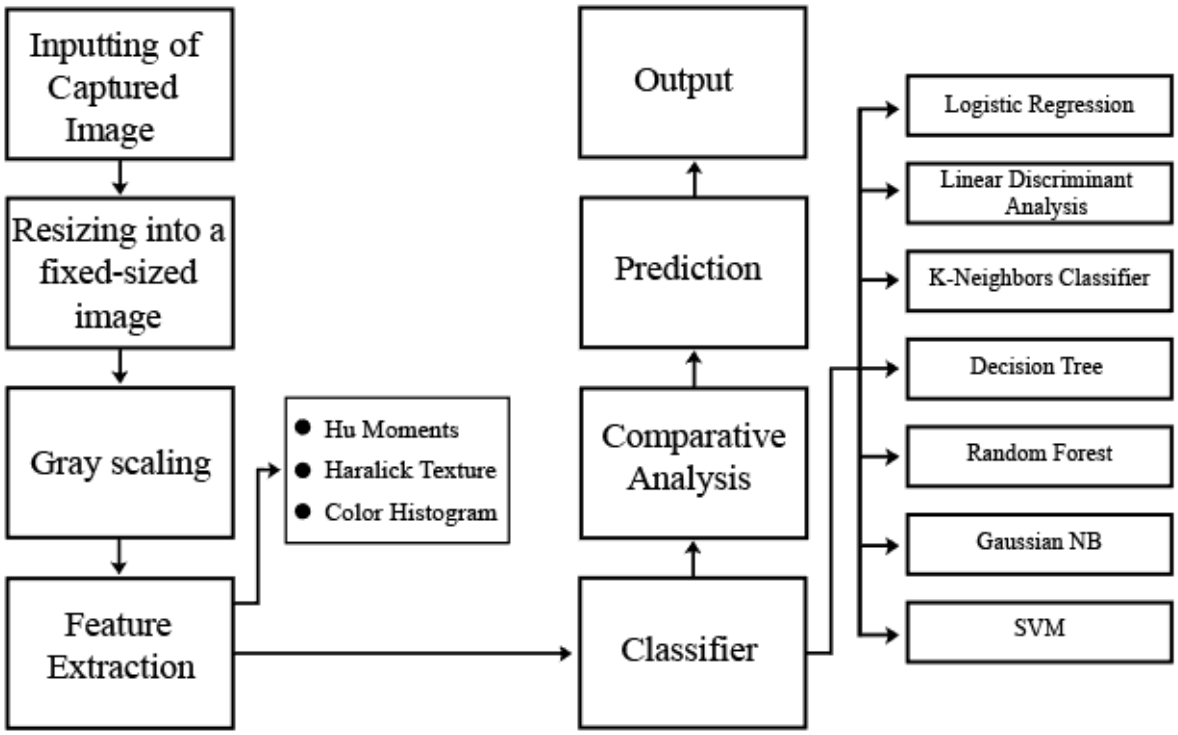


Figure 3.2: Non-interactive proposed systems and Interactive proposed system.

3.4.1 Feature Description

Hu Moments:

Hu Moment is that a set of seven numbers using indomitable central moments in image transformations. For translation, scale, and rotation, and reflection 6 moments have been proved to be invariant and when the 7th moment's appear sign changes for image reflection.

This 7 moments are indomitable using that formulae:

$$h_0 = \eta_{20} + \eta_{02}$$
$$h_1 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$h_2 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$h_3 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$h_4 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$h_5 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})]$$
$$h_6 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

Figure 3.3: Figure of Hu Moments formulae.

Shape Matching by Hu Moments:

Before we tell, 7 Hu Moments are invariant under translations, scale and rotation.

This is an example for Hu Moments.

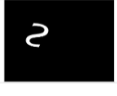
id	Image	H[0]	H[1]	H[2]	H[3]	H[4]	H[5]	H[6]
K0		2.78871	6.50638	9.44249	9.84018	-19.593	-13.1205	19.6797
S0		2.67431	5.77446	9.90311	11.0016	-21.4722	-14.1102	22.0012
S1		2.67431	5.77446	9.90311	11.0016	-21.4722	-14.1102	22.0012
S2		2.65884	5.7358	9.66822	10.7427	-20.9914	-13.8694	21.3202
S3		2.66083	5.745	9.80616	10.8859	-21.2468	-13.9653	21.8214
S4		2.66083	5.745	9.80616	10.8859	-21.2468	-13.9653	-21.8214

Figure 3.4: Shape Matching using Hu Moments

Here we can see, At first, we have moved the letter Ko in K, and S0 in the letter S. Similarly, we moved the letter S in S1, and scaled it in S2. Here we used rotation to make S3 and for make S4 we flipped more.

Here we see that Hu Moments for S0, S1, S2, S3, and S4 are close to each other, and the sign of last Hu moment of S4 is flipped and K0 is different from other.

Haralick's texture features:

Haralick's texture features were calculated using the haralick () function of the cytometry tool box for Khoros. The principle of these properties is the gray-level co-presence matrix. Here, the matrix is square with dimension Ng, where Ng is the number of gray levels. Matrix elements i and j are generated by counting the number of pixel bars with a value attached to the pixels and then dividing the entire matrix by comparing. Each entry is calculated as a probability that a pixel value with the value of i is found adjacent to the j pixel.

$$\mathbf{G} = \begin{bmatrix} p(1,1) & p(1,2) & \cdots & p(1,N_g) \\ p(2,1) & p(2,2) & \cdots & p(2,N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,2) & \cdots & p(N_g,N_g) \end{bmatrix} \quad (1)$$

Color Histogram:

At present, the color histogram is employed to represent the distribution of colors in an image. The color histogram represents the number of pixels that have colors in a fixed list of color range that span the image's color space. For monochromatic images, the set of possible color values is sufficiently small that each of those colors may be placed on a single range; then the histogram is merely the count of pixels that have each possible color. For color images using RGB space, the space is divided into an appropriate number

of ranges, often arranged as a regular grid, each containing many similar color values. Figure 2(a) is an original rainbow image with RGB channels from 0 to 255, so there are totally 256 × 256 × 256 = 224 colors. Figure 2(b) uses four bins to represent each color component, Bins 0, 1, 2, 3 denote intensities 0-63, 64-127, 128-191, 192-255, respectively, so there are in total 4 × 4 × 4 = 64 colors. Figure 2(c) is the histogram of Figure 2(b), where the x-axis denotes the index of the 64 colors, and the y-axis denotes the number of pixels.

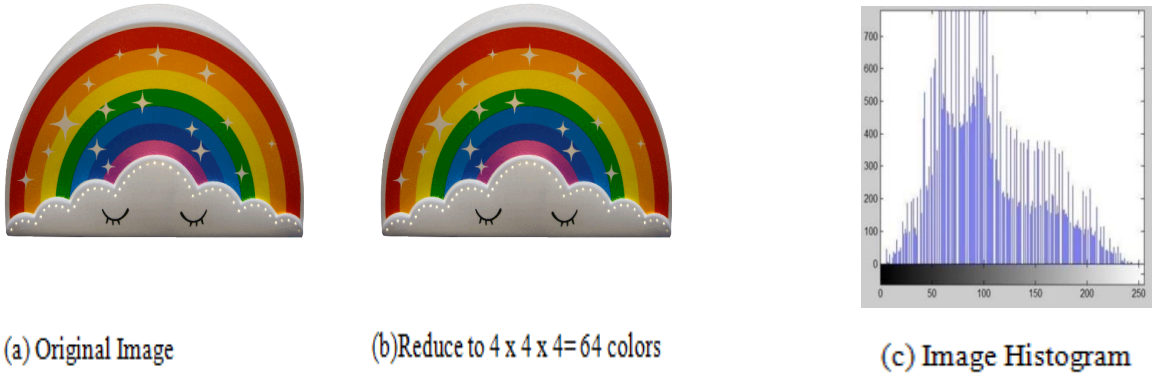


Figure 3.5: Image of Color Histogram

The histogram provides a compact summarization of the distribution of data in an image. The color histogram of an image is relatively invariant with translation and rotation about the viewing axis. By comparing histograms signatures of two images and matching the color content of one image with the other, the color histogram is well suited for the problem of recognizing an object of unknown position and rotation within a scene.

3.4.2 Classifier

K-Nearest Neighborhood:

K-Nearest Neighborhood (K-NN) classifier execution time is very fast and it is very dominant. It works on the basis of the minimum distance between the curry example and the training data set to determine the k-nearest neighbors. Using distance function (x, y)

computing the distance between two scenarios, where x, y are query example and the training data set created of N features as showing in equation 2. Using one of the two functions as in the equations 3, and 4 distance measure is computed.

$$x = \{x_1, x_2, \ldots . x_n\},\ y = \{y_1, y_2, \ldots . y_n\} \quad (2)$$

$$d_i(x, y) = \sum_{i=1}^{n} x_i - y_i \quad (3)$$

$$d_i(x, y) = \sum_{i=1}^{n} \sqrt{x_i^2 - y_i^2} \quad (4)$$

After gathering k-nearest neighbors, simple majority of these k-nearest neighbors is taken to be the prediction of the query instance. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors.

Support Vector Machine:

Support Vector Machine Classifier is a higher-level machine learning algorithm that is used to regression and classifies of high-dimensional datasets with good results. There is an optimal individual hyperplane for solving classification problems which looks into effectively to split between classes known as SVM. SVM aims to maximize the margin around a hyperplane that separates a positive class from a negative class. Let us a training dataset with n samples (x_1,y_1), (x_2,y_2),...., (x_n ,y_n),where a property vector x_i in n-dimensional property spaces and y_i∈ { −1, 1} with the labels corresponding to two separate linear classes C_1 and C_2. Share equations 5 and 6 with decent funds in optimal happiness plan to solve optimistic problems with maximum margin between two classes.

$$\text{Maximize } \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i\, a_j y_i y_i\, .K(x_i x_j) \quad (5)$$

$$\text{Subject} - \text{to:} \sum_{i=1}^{n} a_i\, y_i\, , 0 \leq a_i \leq C \quad (6)$$

Where, a_i the weight determined in the training sample is determined x_i.when $a_i > 0$, the x_i is a support vector. Parameter C of a regulation permit used to use trade-off training assurances and the model complexity has a higher general capacity .Here, K is a kernel function, which can measure the equality between two samples.

Decision Tree Algorithm:

The decision tree is a flowchart-like structure of the tree where the internal node represents the feature, the branch represents the decision rule, and each leaf node represents the result. The node at the top of the decision tree is called the root node. It learns to divide based on attribute values. This is called repetitive partitioning of the tree in a repetitive manner. Structures like this flowchart help you make decisions. It's a scene like a flowchart diagram that easily copies human-level thoughts. That's why it is easy to understand and explain.

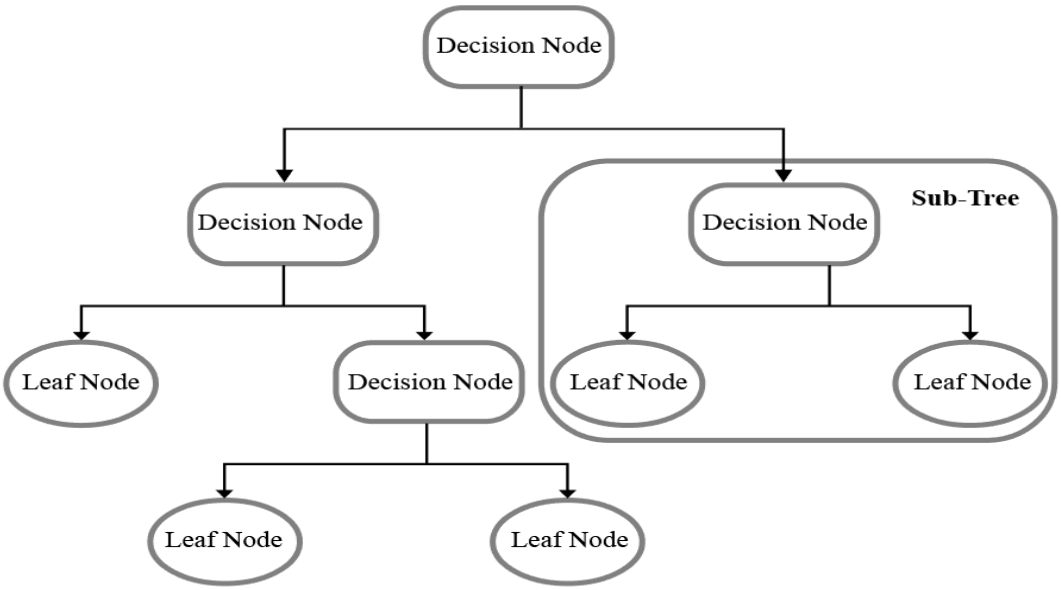


Figure 3.6: Flowchart diagram of Decision Tree

We know about many types Machine Learning Algorithm. Decision Tree is one of them. Basically, it is a type of Machine Learning algorithm known as white box. It shares the logic of internal decision making, which is not found in black box type algorithms like

neural networks. Its training time is faster than with neural network algorithms. Time complexity is depends on several numbers of thing for any function. For the decision tree, it goes to the same. It is a function of records number and attributes number in the given dataset. The decision tree is a method which is distribution free or non-parametric. It simply means it does not depend on any probability assumptions. Decision tree has the capability to handle a good accuracy rate of high dimensional data.

The main concept of the decision tree algorithm is as follows:

- Best attributes should get selected , Here Attribute Selection Measures (ASM) can be used to split the records.
- Transform the attribute to a decision node which break the datasets into smaller sets(subset).
- Tree building should be started building by recursively repeating this process for each child.
- Same attribute value should own all the tuples
- No attributes should remain.
- No instances should remain.

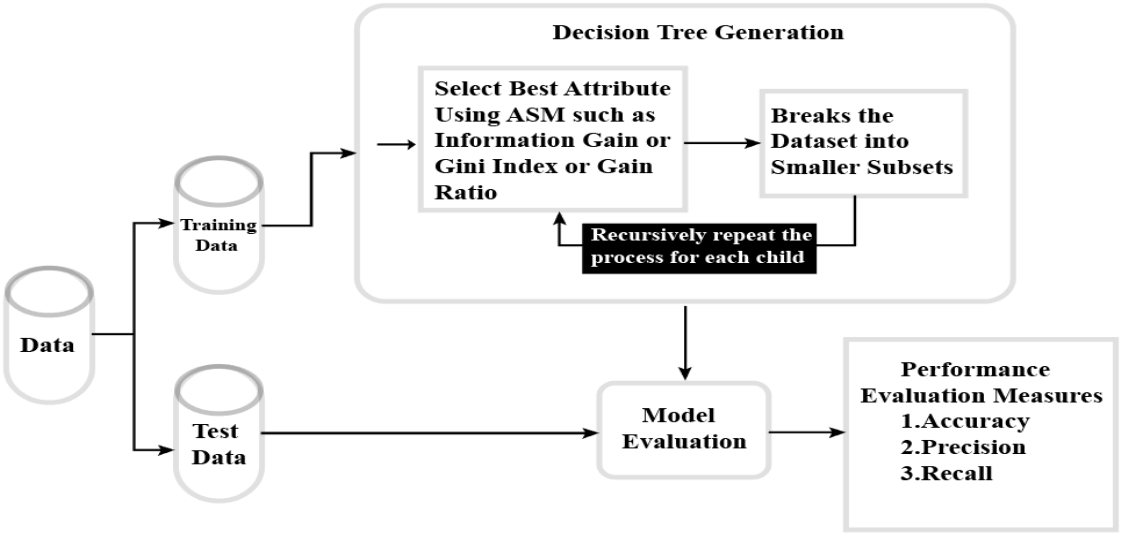


Figure 3.7: Decision Tree Generation

Linear Discriminant Analysis:

Linear Differential Analysis (LDA) is a dimensional reduction strategy, which means that dimension reduction techniques reduce the number of dimensions in a datasheet while retaining as much information as possible. For an example, suppose we plot the relationship between two variables where each color represents a different class.

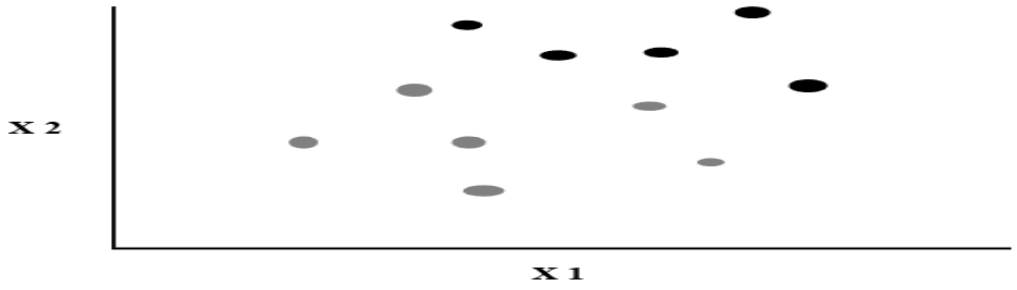


Figure 3.8: Plot diagram of Linear Discriminant Analysis

Logistic regression:

A regression analysis is required when a dichotomous(binary) dependent variable is found. Logistic regression is the right analysis to perform. The logistic regression is predictable same as all other regression analyses. Mostly this function is called when a data needed to be described and explained relationship between one dependent binary variable and other independent variables(nominal, ordinal, interval or ratio-level).

Random Forest algorithm:

The random forest is an algorithm that is a supervised classification algorithm. The name itself tells its purpose which is to propagate a forest of data in some way randomly. Outcomes are directly dependent on the number of trees in the forest. The result would be more accurate if there is a larger number of trees. But there is one thing which should

be noted that forming a forest is not as same as comprising decision with gain index approach or information gain.

We used Random Forest algorithm because of four advantages. One of the notable advantages are this algorithm can not only be applied for classification but also for regression tasks. Secondly, Major problem is overfitting which can make the results worse for a random forest algorithm, if enough trees in the forest, it will not be suitable compared to the classified model. Third is random forests classifiers are used to model the classified values . Forth is handle missing values. The random forest algorithm has two levels, one is creation and the other is the prediction from random forest classification created in the first stage. The total process is given below, and for easy to understand here we used a figure .

Pseudocode of Random Forest creation:

- From total 'm' features, it randomly selects 'K' where (K <<m).
- Calculating 'd' node using the best split point from the 'K' features.
- Using the best split divide the nodes into leaf nodes.
- Repeating to the step 'c' until 'I' number of nodes has been reached out.
- Repeating the steps for 'n' number of times for creating 'n' number of trees.

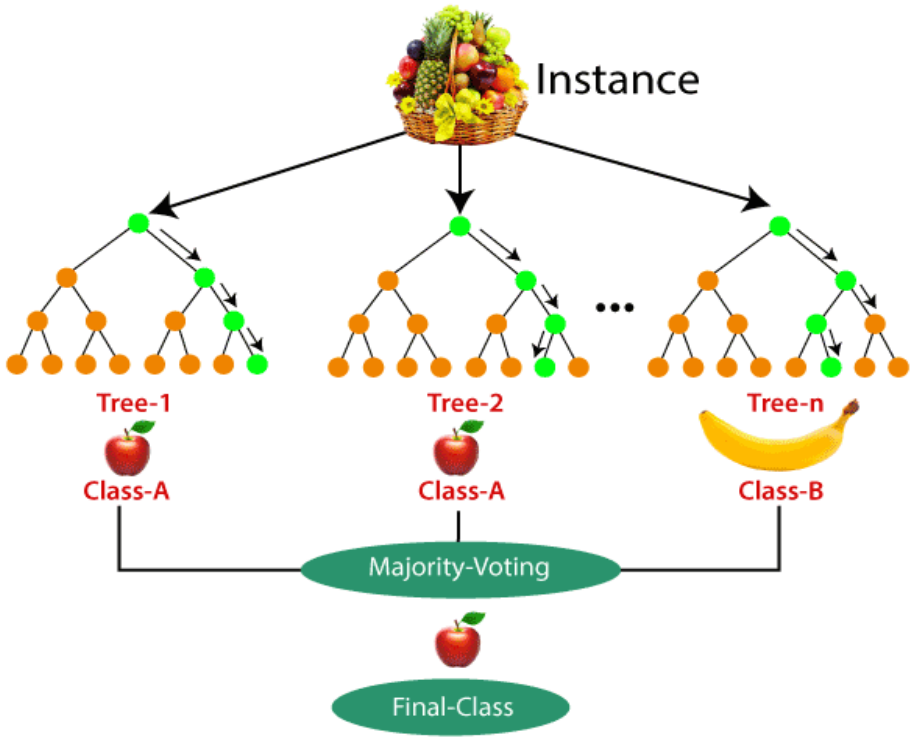


Figure 3.9: Random Forest Simplified

3.5 Implementation Requirements

After Proposed Methodology, we created a list of developing tools we need for classification and identify of potatoes. The probable necessary things are:

Developing Tools

- python 3.7
- Anaconda
- Jupyter notebook
- Pandas
- Sklearn

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSIONS

4.1 Introduction:

In the previous section we explain about the data collection method, data processing, proposed methodology, statistical analysis, and different implementation requirements for solving this type problem. In this section we will finally explain what we have found so far throughout our entire research period. We will try to describe our whole process here.

4.2 Experimental Evaluation:

We collected the data of 4 verities total 1200 potatoes image. We collected all of potato from local Market. First, we divided our data set with 4 class variables with different labeling. In feature extraction for preprocessing use sklearn library for Python and import LabelEncoder, MinMaxScaler. In this process, we wanted to identify the real image of potatoes. In this process, we wanted to identify each potato individually. Finally we applied machine learning algorithms as like Random Forest Classifier (RF), Linear Discriminant Analysis (LDA), Logistic Regression, Support Vector Machine (SVM), CART, NB and KNN on our datasets. After applying all algorithms each of them produced different results. Random Forest Classifier (RF) show best result and its accuracy rate 100% and SVM gives lowest rate. Its accuracy rate 74.074%.

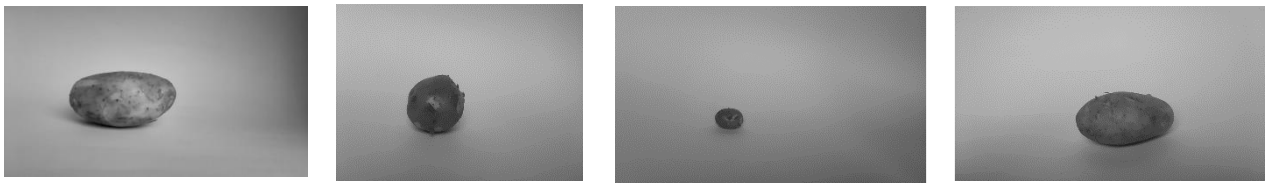
Table 4.1: Summary of Accuracy from all using algorithm

Using Algorithm	Accuracy
LR	99.35%
LDA	99.35%
KNN	98.33%
CART	99.53%
RF	100%
NB	76.94%
SVM	73.42%

4.3 Descriptive Analysis:

Basically we divided the whole process in 3 main parts: feature extraction, training our model and testing our model. The problem we mentioned can be solved in different ways. Notable among them are traditional machine learning, image processing, and deep learning. They all solve the problem separately. Individual algorithms are used in each case. We have chosen traditional machine learning for our work. We use Anaconda Python for our background support. Anaconda is a free and open-source distribution of Python. And it's an R programming language for scientific computing that is: data science, machine learning applications, large-scale data processing, predictive analytics, etc. whose goal is to facilitate package management and deployment. First step is feature extraction which is the part of data preprocessing. Here we use 1200 images which actually in different shape. We used 1,200 images in our work, each with different shapes and sizes. We took the pictures in different light conditions. For ease of work (350 * 450) resized the pixels and make sure all image are same size. We have fixed the path of our

data set. For preprocessing use most popular library (sklearn) for Python and import LabelEncoder, MinMaxScaler which work for labeling the different class and resize all of images. And import numpy, mahotas, cv2, os, h5py, re library. In processing there are: Hu Moments, Haralick Texture and Color Histogram. In Hu Moments with help of CV2 library image converted BGR2GRAY scale. By using mahotas library so that compute the haralick texture feature vector. By using Color Histogram feature, its converted HSV color-space from original image. Hu Moments feature-descriptor BGR – GRAY image color converted. Haralick Texture feature-descriptor BGR – GRAY image color converted,



(a) (b) (c) (d)

Figure 4.1: Gray scaling image (a) BARI Alu-14 (Cleoptra), (b) BARI Alu-18 (Baraka), (c) BARI Alu-28 (Lady Rosetta), (d) BARI Alu-55 (Red Fantasy)

After that we train labels with os library and sorting label with sort() function. We developed a functional structure with all sub-folders where loop over shape, color and others feature with training data. Then get the current training label using os.path.join(train_path, training_name) this function. In each sub-folder loop over the images with the help of CV2 library get the image file name and read the image first then with a fixed-size it resize.

Get the overall feature vector size using this function format(np.array().shape and then encode the target labels after that save the feature vector using h5py library finally end of training data. Here show this result,

```
Feature vector size (1200, 532)
Training Labels (1200,)
Target labels: [0 0 0 ... 3 3 3]
Target labels shape: (1200,)
End of training data.
```

Figure 4.2: Status for training data

Here can see that total train 1200 images, vector size (1200, 532) target labels are train with [0, 1, 2, 3] numbering.

Second step is training our model. In time developed a model for cassifier with help of those algorithm: Logistic Regression, Decision Tree Classifier, Random Forest Classifier, K-Neighbors Classifier, Linear Discriminant Analysis, Gaussian NB, SVM find a result which one is best and how much our data accuracy rate. After complete verifying then split the training and testing data in individual part,

```
[STATUS] splitted train and test data...
Train data  : (1080, 532)
Test data   : (120, 532)
Train labels: (1080,)
Test labels : (120,)
```

Figure 4.3: Status for Split the training & testing data

After splitting train and testing data we get all Algorithm performing result. Here we can see mean and standard deviation. And this means actually accuracy rate. This accuracy rate close to 1 gives best result. Here RF gives best result. Next we testing our data with RF model because RF gives best result.

```
Different Type of Algorithom performing result

LR: 0.995370 (0.006211)
LDA: 0.992593 (0.006929)
KNN: 0.978704 (0.012457)
CART: 0.993519 (0.008333)
RF: 1.000000 (0.000000)
NB: 0.766667 (0.027716)
SVM: 0.740741 (0.048290)
```

Figure 4.4: Output result from all using algorithm

After find Algorithm result now comparison experimental result for displaying with boxplot, violinplot and plot.

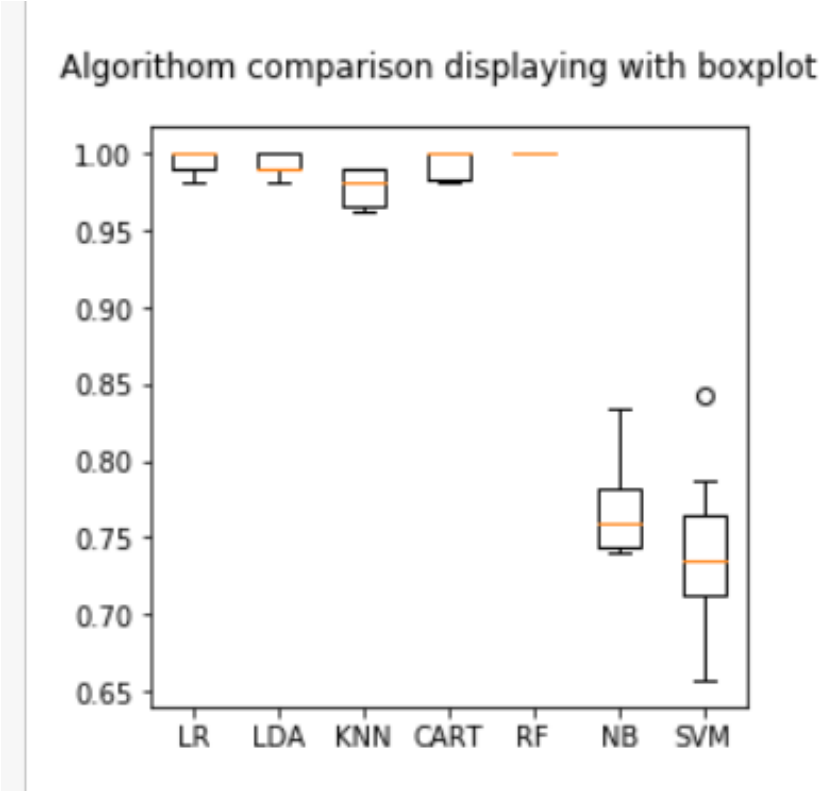


Figure 4.5: Boxplot algorithm comparison

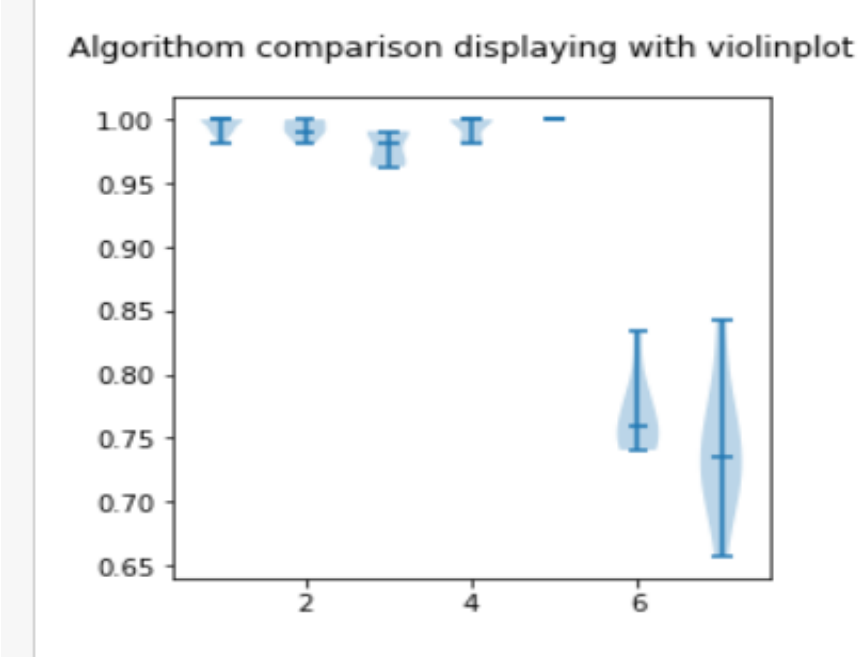


Figure 4.6: Violin plot algorithm comparison

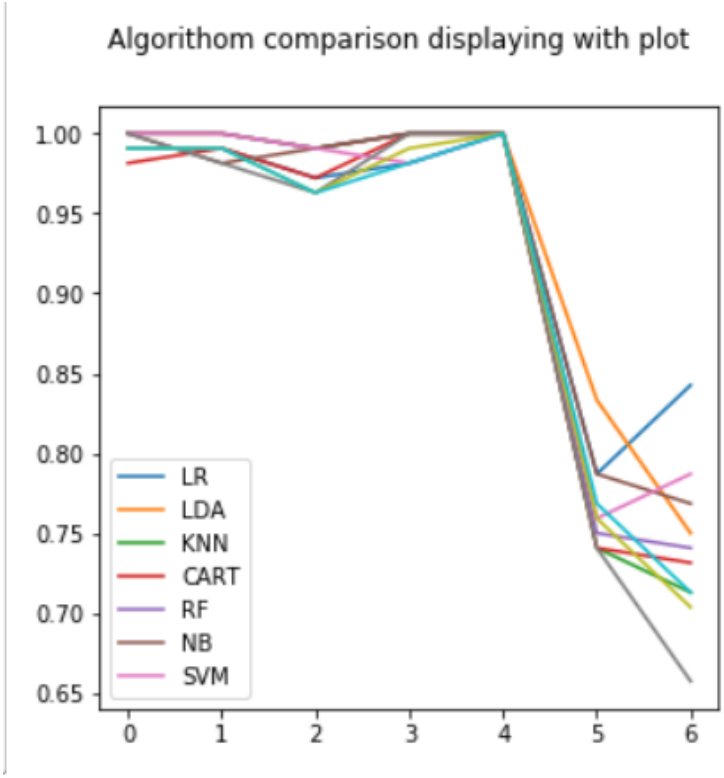


Figure 4.7: Comparison compare with plot displaying

Now third and final step is testing model. For test this model use 4 category different images and it can detect all of them properly. One of them is given bellow –

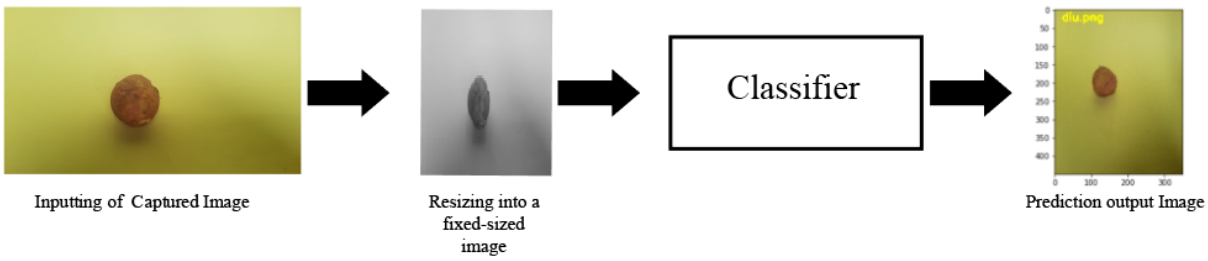


Figure 4.8: Output prediction of input image

4.4 Summary:

Must have a proper idea of all meanings and individual performance on every model to improve the accuracy of the model. These parameters are to find the best value for each individual part. Here we used Different types of parameters we used here. They are such that number of trees, maximum feature, oob score, random state of the object and others. That's why In RandomForest inherent better and more accurate models because here optimization of all used parameter values and finally give a best result. By comparison with others RF shows best results and its accuracy rate is 100% and LR, LDA, KNN, CART those are very close to each other. SVM and NB are also closer to each other but they show the lowest accuracy rate. They show 77% and 74% respectfully.

CHAPTER 5

CONCLUSION AND IMPLICATION FOR FUTURE RESEARCH

5.1 Conclusion:

In order to make proper potato classification and identification, external properties of potatoes are very important attributes like color, shape, size, texture and defects. Now-a-days machine vision has become so advanced and its very much available with low cost of hardware and software which is why in many cases potato classification and identification is replaced with automated vision systems. Another reasons behind this initiative is it has the ability to produce more accurate, rapid, fast, objective and efficient results over manual process.

The entire paper schemes the basic process genre of potato classification and identification. With Hu Moments, Color Histogram and Haralick's texture features, feature extraction methods for color, shape, size and texture are shortly explained. Finally, some machine learning approaches are briefly discussed like Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Logistic Regression, CART, Random Forest Classifier (RF), NB and K-NN. Although still there are some challenges that needed to be overcome, but someday machine vision will certainly become a technique for non-destructive potato classification and identification in the future.

5.2 Implication for Further Study:

After achieving experiences from this project, we want to work on image classification for different fruits and vegetables in the future. We will figure fruits and vegetables grading and then prepare algorithm and machine vision. It is a kind of a process that can also be used for classifying and identifying plants, leaves or flowers. A system can be developed for identifying plants, leaves or flowers and give all the information regarding

them. Also, we can work on some other features for identification and classification like a system which can identify different types of defects or disease or texture structure of fruits and vegetable. Each of them can be used for future opportunities in different directions. For example, an individual can work on these ideas and prepare prototype model based on his vision which can be used in industries. Or he might develop mobile applications for the same cause based on above techniques. It will be easier for the agriculturist or common people for easily operate it for different purposes like horticultural products identification, classification and grading.

APPENDIX

Appendix: Research Reflections

We didn't have much knowledge about application of machine learning in classification and identification before we started. Our very kind and helpful supervisor helped us choose our research topic and gave the complete guideline needed to get our research done. We took a long time to learn, practice and get ready for work. In the first case we worked on a small number of dataset to check if the model works properly. When we saw that the model was working properly, then we applied it to the whole dataset to see our expected outcomes. Throughout this whole journey we have learned a lot and developed some new ideas to do something in our future works.

One thing from our first research we are learning that once you have the whole picture of your work in your head and you already know how to do it, it won't take long to finish.

Reference:

[1] M. T. Habib, A. Majumder, R. N. Nandi, F. Ahmed and M. S. Uddin, "Machine Vision Based Papaya Disease Detection", Journal of King Saud University – Computer and Information Sciences, vol. 32, pp. 300-309, March 2020

[2] H. M. Zawbaa, M. Abbass, M. Hazman and A. E. Hassenian , "Automatic Fruit Image Recognition System Based on Shape and Color Features", A.E. Hassanien et al. (Eds.): AMLTA 2014, CCIS 488, pp. 278–290, 2014

[3] M. B. Lak, S. Minaei, J. Amiriparian and B. Beheshti, "Apple Fruits Recognition Under Natural Luminance Using Machine Vision", Advance Journal of Food Science and Technology, vol. 2, pp. 325-327, November 2010

[4] S. Arivazhagan, R. N. Shebiah, S. S. Nidhyanandhan and L.Ganesan, "Fruit Recognition using Color and Texture Features", Journal of Emerging Trends in Computing and Information Sciences, vol. 1, NO. 2, Oct 2010

[5] I. Hussain, Q. He and Z. Chen "AUTOMATIC FRUIT RECOGNITION BASED ON DCNN FOR COMMERCIAL SOURCE TRACE SYSTEM", International Journal on Computational Science & Applications (IJCSA) vol.8, No.2/3, June 2018

[6] S. Jana, S. Bask and R. Parekh, "Automatic Fruit Recognition from Natural Images using Color and Texture Features", 2017 Devices for Integrated Circuit (DevIC), 23-24 March 2017

[7] W. Astuti, S. Dewanto, K. E. N. Soebandrija and S. Tan, "Automatic fruit classification using support vector machines: a comparison with artificial neural network", IOP Conference Series: Earth and Environmental Science, v. 195, 1 December 2018

[8] S. Naik and B. Patel, "Machine Vision based Fruit Classification and Grading - A Review", International Journal of Computer Applications, vol. 170, No.9, pp. 0975 – 8887, July 2017

[9] R. S. L, A. N, D. V and K. T, "Fruit Classification System Using Computer Vision: A Review", International Journal of Trend in Research and Development, vol. 5, ISSN: 2394-9333, Jan - Feb 2018

[10] H. M. Zawbaa, M. Hazman, M. Abbass and A. E. Hassanien, "Automatic fruit classification using random forest algorithm", 2014 14th International Conference on Hybrid Intelligent Systems, April 2015

[11] U. S. R., N. Dawande and A. Mulajkar, "Automated Fruit Classification System", International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, July 2016

[12] A. A. Elsharif, I. M. Dheir, A. S. A. Mettleq and S. S. Abu-Naser, "Potato Classification Using Deep Learning", International Journal of Academic Pedagogical Research (IJAPR), vol. 3, pp. 1-8, December – 2019

[13] Y. Song, C. A. Glasbey, G. W. Horgan, G. Polder, J. A. Dieleman and G.W.A.M. van-der-Heijden, "Automatic fruit recognition and counting from multiple images", Biosystems Engineering, vol. 118, pp. 203-215, February 2014

[14] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, "Automatic Fruit Classification Using Deep Learning for Industrial Applications", IEEE Transactions on Industrial Informatics, vol. 15, pp. 1027-1034, October 2018

[15] A. Azizi, Y. Abbaspour-Gilandeh, M. Nooshyar and A. Afkari-Sayah, "Identifying Potato Varieties Using Machine Vision and Artificial Neural Networks", International Journal of Food Properties, val. 19, pp. Pages 618-635, Dec 2015

[16] G. Athanikar and P. Badar," Potato Leaf Diseases Detection and Classification System", International Journal of Computer Science and Mobile Computing, vol. 5, pp. 76-88, February 2016

[17] A. Yadollahinia, A. Latifi and R. Mahdavib, "New method for determination of potato slice shrinkage during drying", vol. 65, pp. 268-274, March 2009

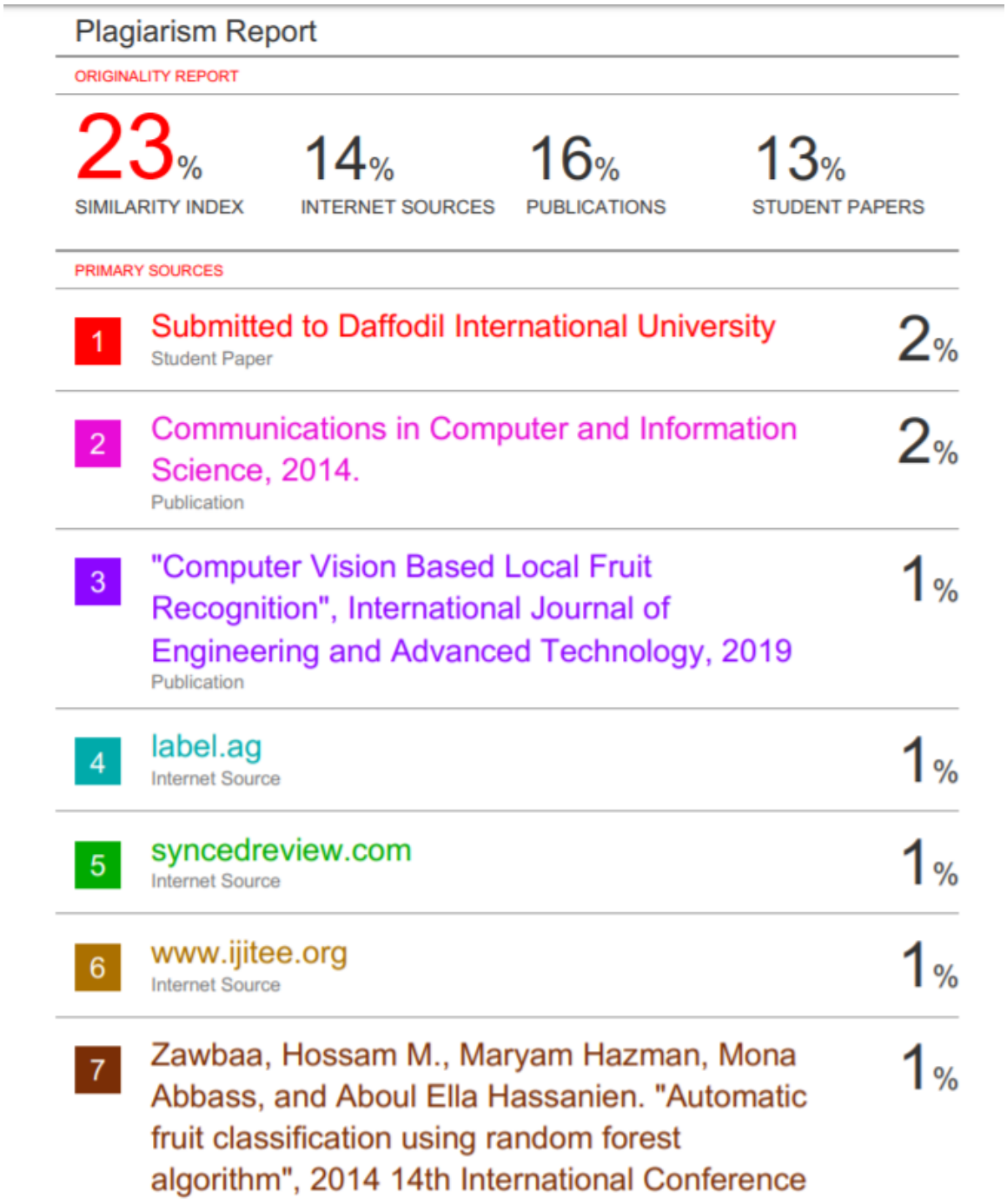
[18] N. Pandey, S. Kumar and R. Pandey, "Communications in Computer and Information Science", vol. 839, CNC, pp. 329-339

[19] Learn about Proven Ways for improving Accuracy of a Machine Learning Model, available at << https://www.analyticsvidhya.com/blog/2015/12/improve-machine-learning-results/>>, last accessed on 02-05-2020 at 08:06 PM.

[20] Learn about Potato and potato plant, available at << http://en.banglapedia.org/index.php?title=Potato >>, last accessed on 02-05-2020 at 08:11 PM.

[21] Learn about Potato, available at << https://en.wikipedia.org/wiki/Potato >>, last accessed on 02-05-2020 at 08:26 PM.

Plagiarism Report



Plagiarism Report

ORIGINALITY REPORT

23%	14%	16%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	Communications in Computer and Information Science, 2014. Publication	2%
3	"Computer Vision Based Local Fruit Recognition", International Journal of Engineering and Advanced Technology, 2019 Publication	1%
4	label.ag Internet Source	1%
5	syncedreview.com Internet Source	1%
6	www.ijitee.org Internet Source	1%
7	Zawbaa, Hossam M., Maryam Hazman, Mona Abbass, and Aboul Ella Hassanien. "Automatic fruit classification using random forest algorithm", 2014 14th International Conference	1%

	on Hybrid Intelligent Systems, 2014. Publication	
8	docplayer.net Internet Source	1%
9	www.ijsr.net Internet Source	1%
10	Donald J. Norris. "Machine Learning with the Raspberry Pi", Springer Science and Business Media LLC, 2020 Publication	1%
11	Suchet Bargoti, James Underwood. "Image classification with orchard metadata", 2016 IEEE International Conference on Robotics and Automation (ICRA), 2016 Publication	1%
12	medium.com Internet Source	<1%
13	ie.binus.ac.id Internet Source	<1%
14	Winda Astuti, Satrio Dewanto, Khristian Edi Nugroho Soebandrija, Sofyan Tan. "Automatic fruit classification using support vector machines: a comparison with artificial neural network", IOP Conference Series: Earth and Environmental Science, 2018 Publication	<1%

15	Submitted to The Robert Gordon University Student Paper	<1%
16	Submitted to University of Southampton Student Paper	<1%
17	arxiv.org Internet Source	<1%
18	Ismael-Antonio Davila-Rodriguez, Marco-Aurelio Nuno-Maganda, Yahir Hernandez-Mier, Said Polanco-Martagon. "Decision-Tree Based Pixel Classification for Real-time Citrus Segmentation on FPGA", 2019 International Conference on ReConFigurable Computing and FPGAs (ReConFig), 2019 Publication	<1%
19	Jasman Pardede, Milda Gustiana Husada, Asep Nana Hermana, Sri Agustina Rumapea. "Fruit Ripeness Based on RGB, HSV, HSL, L*a*b* Color Feature Using SVM", 2019 International Conference of Computer Science and Information Technology (ICoSNIKOM), 2019 Publication	<1%
20	Submitted to University of Bedfordshire Student Paper	<1%
21	Submitted to University College London Student Paper	<1%

22	murphylab.web.cmu.edu Internet Source	<1%
23	philpapers.org Internet Source	<1%
24	M. Muthu Rama Krishnan, Chandan Chakraborty, Ranjan Rashmi Paul, Ajoy K. Ray. "Hybrid segmentation, characterization and classification of basal cell nuclei from histopathological images of normal oral mucosa and oral submucous fibrosis", Expert Systems with Applications, 2012 Publication	<1%
25	www.irjet.net Internet Source	<1%
26	worldsciencepublisher.org Internet Source	<1%
27	Submitted to Napier University Student Paper	<1%
28	Submitted to De Montfort University Student Paper	<1%
29	download.atlantis-press.com Internet Source	<1%
30	ijarcet.org Internet Source	<1%

31	www.science.gov Internet Source	<1%
32	wireilla.com Internet Source	<1%
33	Submitted to University of Nottingham Student Paper	<1%
34	"Advanced Machine Learning Technologies and Applications", Springer Science and Business Media LLC, 2021 Publication	<1%
35	www.ijetcse.com Internet Source	<1%
36	Anuja Bhargava, Atul Bansal. "Fruits and Vegetables Quality Evaluation Using Computer Vision: A Review", Journal of King Saud University - Computer and Information Sciences, 2018 Publication	<1%
37	Rocha, A.. "Automatic fruit and vegetable classification from images", Computers and Electronics in Agriculture, 201001 Publication	<1%
38	dstore.alazhar.edu.ps Internet Source	<1%
39	Lecture Notes in Electrical Engineering, 2014.	