



Faculty of Computer Science and Information Technology

FRUIT RECOGNITION APPLICATION USING MACHINE LEARNING

LOH WENG KEONG

Bachelor of Computer Science with Honours (Software Engineering)

2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE **Fruit Recognition Application Using Machine Learning**

ACADEMIC SESSION: 2022/2023

LOH WENG KEONG
(CAPITAL LETTERS)

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (✓)

<input type="checkbox"/>	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
<input type="checkbox"/>	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
<input checked="" type="checkbox"/>	UNRESTRICTED	

Validated by



(AUTHOR'S SIGNATURE)

Permanent Address

902, Lorong Bercham 4,
Kampung Bercham,
31400 Ipoh

Date: 30/6/2023

(SUPERVISOR'S SIGNATURE)

Date: _____

Note * Thesis refers to PhD, Master, and Bachelor Degree

** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

FRUIT RECOGNITION APPLICATION USING MACHINE LEARNING

LOH WENG KEONG

This project is submitted in partial fulfilment of the requirements for the degree of Bachelor of
Computer Science with Honours (Software Engineering)

Faculty of Computer Science and Information Technology

UNIVERSITI MALAYSIA SARAWAK

2023

APLIKASI PENGECAMAN BUAH MENGGUNAKAN PEMBELAJARAN MESIN

LOH WENG KEONG

Projek ini merupakan salah satu keperluan untuk Ijazah Sarjana Muda Sains Komputer dengan
Kepujian (Kejuruteraan Perisian)

Fakulti Sains Komputer dan Teknologi Maklumat

UNIVERSITI MALAYSIA SARAWAK

2023

DECLARATION

I hereby declare that this project is my original work. I have not copied from any other student's work or from any other sources except where due reference or acknowledgement is not made explicitly in the text, nor has any part had been written for me by another person.



.....
(LOH WENG KEONG)

1st July 2023

Matric No: 70236

ACKNOWLEDGEMENT

Firstly, I am deeply grateful to my supervisor, Prof. Dr Wang Yin Chai, for his invaluable guidance, assistance, and advice during the development of this project. His expertise and support were crucial in helping me complete my work. I would also like to thank my examiner, Dr. Tan Ping Ping, for their valuable comments and suggestions that helped to improve the quality of my project.

Next, I am thankful to Prof. Dr Wang Yin Chai, the final year project coordinator again, for providing me with the necessary information and guidance to complete my final year project. I would also like to thank my course mates for their support and for providing me with the information I needed to finish this project.

Lastly, I am extremely grateful to my parents and friends for their constant support and encouragement throughout this project. Their support has given me the motivation and strength to see it through to completion. Without their help, this project would not have been possible.

Table of Contents

ACKNOWLEDGEMENT	i
LIST OF FIGURES	v
LIST OF TABLES	vii
LIST OF EQUATIONS	viii
ABSTRACT	ix
ABSTRAK	x
CHAPTER 1: INTRODUCTION	1
1.1 Introduction.....	1
1.2 Problem Statement	3
1.3 Objectives	3
1.4 Brief Methodology	4
1.5 Scope.....	5
1.6 Significance of Project.....	6
1.7 Expected Outcomes	6
1.8 Project Schedule.....	7
1.9 Outline of Project Report.....	7
CHAPTER 2: LITERATURE REVIEW	9
2.1 Introduction.....	9
2.2 Existing Techniques in Fruit Recognition	10
2.2.1 CBIR using K-Nearest Neighbours (KNN).....	10
2.2.1.1 Color Extraction using RGB & Shape Feature Extraction	11
2.2.1.2 Classification using KNN	12
2.2.1.3 Results from KNN	13
2.2.2 CBIR using Support Vector Machine (SVM)	14
2.2.2.1 Images Acquisition & Images Pre-Processing (SVM)	14
2.2.2.2 Color Features Extraction Using LAB Color Model, Texture & Shape Features Extraction.....	14
2.2.2.3 Classification using SVM	17
2.2.2.4 Results from SVM	18
2.2.3 CBIR using Back Propagation Neural Networks (BPNN).....	19

2.2.3.1	Image Acquisition & Image Pre-processing (BPNN).....	19
2.2.3.2	Color Feature Extraction using HSV Color Model & Shape Feature Extraction ...	20
2.2.3.3	Classification using BPNN	23
2.2.3.4	Results from BPNN	25
2.3	Comparative Analysis of Reviewed Techniques	26
2.3.1	Review on Color Model	26
2.3.2	Review on Features Extraction.....	27
2.3.3	Analysis of Review.....	29
CHAPTER 3: METHODOLOGY		31
3.1	Introduction.....	31
3.2	Methodology	31
3.2.1	Process flow of image processing technique.....	33
3.2.1	Image Acquisition	33
3.2.2	Image Pre-Processing.....	35
3.2.3	Feature Extraction.....	36
3.2.3.1	Color Extraction.....	36
3.2.3.2	Shape Extraction	39
3.2.4	Classification	40
3.2.4.1	Training the neural network.....	40
3.2.4.2	Testing the neural network.....	42
3.5.5	Results and Evaluation	43
3.6	Conclusion	43
CHAPTER 4: IMPLEMENTATION AND TESTING		44
4.1	Introduction.....	44
4.2	Implementation Environment	44
4.3	Model Training Process	45
4.3.1	Step 1: Image Acquisition	45
4.3.2	Step 2: Image Pre-processing	46
4.3.3	Step 3: Features Extraction	48
4.3.3.1	Color Features Extraction.....	48
4.3.3.2	Shape Features Extraction	50

4.3.3.3	Combine Extracted Features.....	51
4.3.4	Step 4: Classification Using Back-propagation Neural Network.....	51
4.3.5	Step 5: Testing Result & Evaluation	53
4.4	Key Contributions.....	54
4.4.1	Parameters Selection and Threshold Determination	54
4.4.2	Neural Network Properties	55
4.5	Graphical User Interface Prototype	56
4.6	Conclusion	59
CHAPTER 5: Discussion.....		60
5.1	Introduction.....	60
5.2	Results and Evaluation.....	60
5.3	Conclusion	67
CHAPTER 6: CONCLUSION AND FUTURE WORK		68
6.1	Introduction.....	68
6.2	Project Achievements	68
6.3	Project Limitations.....	69
5.4	Project Future Work.....	69
References.....		70
APPENDICES		73
APPENDIX A: GANTT CHART OF PROJECT SCHEDULE (FYP1).....		73
APPENDIX B: GANTT CHART OF PROJECT SCHEDULE (FYP 2).....		74

LIST OF FIGURES

Figure 2. 1 Lab Color Space (Shukla & Desai, 2016)	15
Figure 2. 2 HP1 with smaller margin, HP2 with greater margin (Shukla & Desai, 2016)	17
Figure 2. 3 HSV color model space (Erdoğan, 2014).....	20
Figure 2. 4 Example of fruit contour extraction image.....	22
Figure 2. 5 Structure of BP neural network (Maru et al, 2021)	24
Figure 3. 1 Process flow of proposed technique	33
Figure 3. 2 Example of collected fruit images	34
Figure 3. 3 Flowchart of Image Pre-processing.....	35
Figure 3. 4 Flowchart color feature extraction.....	38
Figure 3. 5 Flowchart of training neural network.....	41
Figure 4. 1 Source code of dataset input directory path	46
Figure 4. 2 Source code of load train image and labeling	46
Figure 4. 3 Source code of image preprocessing	47
Figure 4. 4 Binary image after preprocessing.....	48
Figure 4. 5 Source code of color features extraction	49
Figure 4. 6 HSV color histogram.....	49
Figure 4. 7 Source code of shape feature extraction.....	50
Figure 4. 8 Source code of combine feature vector	51
Figure 4. 9 Source code of split datasets.....	51
Figure 4. 10 Source code of model training.....	52
Figure 4. 11 Source code to generate confusion matrix.....	53
Figure 4. 12 GUI of prototype	56
Figure 4. 13 GUI of prototype after select image	57

Figure 4. 14 Recognition result and details	58
Figure 4. 15 HSV color histogram from result	58
Figure 5. 1 Scatter graph of percentage of predicted mango	63
Figure 5. 2 Scatter graph of percentage of predicted durian.....	63
Figure 5. 3 Scatter graph of percentage of predicted rambutan.....	64
Figure 5. 4 Precision and Recall Chart	66

LIST OF TABLES

Table 1. 1 Stored details of shape roundness, area and perimeter for each type of fruit in the system during training (Seng & Mirisae, 2009).....	12
Table 2. 1 Recognition Rate Based on Different Characteristics (Tang et al, 2020).....	25
Table 2. 2 The comparison of different color model (Sánchez & García, 2013)	26
Table 2. 3 Parameter used in reviewed techniques	27
Table 2. 4 Comparison of reviewed techniques.....	29
Table 3. 1 Result prediction of fruit recognition using BPNN	42
Table 4. 1 Implementation Tools	45
Table 5. 1 Result of test images from trained model using Back-Propagation Neural Network..	60
Table 5. 2 Confusion Matrix of accuracy of recognition for each fruit types	65
Table 5. 3 Values of Precision and Recall	66

LIST OF EQUATIONS

Equation 1. Shape Roundness.....	11
Equation 2. Euclidean Distance	13
Equation 3. Aspect Ratio	17
Equation 4. Calculation of Margins Between Two Hyperplanes	18
Equation 5. Feature Vector of Global Color Histogram.....	21
Equation 6. Roundness.....	39
Equation 7. Sigmoid Function	40
Equation 8. Accuracy of Recognition.....	43

ABSTRACT

In this study, we present a fruit recognition system using machine learning techniques. The system consists of several distinct phases, including image acquisition, image pre-processing, feature extraction, classification, and results and evaluation. In the image acquisition phase, the system gathers a dataset of images that will be used for training and testing. These images may be acquired from a variety of sources, such as digital cameras or online databases. In the image pre-processing phase, the system prepares the images for further analysis by cleaning them up and standardizing them. This may include tasks such as removing noise, correcting for distortion, or resizing the images. In the feature extraction phase, the system extracts meaningful features from the images that can be used to distinguish one fruit from another. These features include color and shape features. In the classification phase, the system uses the extracted features to classify the fruits using a machine learning algorithm – a backpropagation neural network (BPNN). Finally, in the results and evaluation phase, the system assesses the accuracy and effectiveness of the classification process. Next, a fruit recognition prototype is developed to implement the trained model. The prototype is ready to use to recognize input image and display the result. Overall, our approach demonstrates a thorough and systematic approach to developing a fruit recognition system using machine learning techniques.

ABSTRAK

Dalam kajian ini, kami menyajikan sistem pengenalan buah menggunakan teknik pembelajaran mesin. Sistem ini terdiri daripada beberapa fasa yang berbeza, termasuk pengambilan gambar, pra-penyediaan gambar, pengambilan ciri, pengkelasan, dan keputusan dan penilaian. Dalam fasa pengambilan gambar, sistem mengumpul satu set data gambar yang akan digunakan untuk latihan dan ujian. Gambar-gambar ini boleh diperolehi dari pelbagai sumber, seperti kamera digital atau pangkalan data dalam talian. Dalam fasa pra-penyediaan gambar, sistem menyediakan gambar untuk analisis lanjut dengan membersihkan dan menyamaradarkan mereka. Ini mungkin termasuk tugas-tugas seperti mengeluarkan bising, membetulkan distorsi, atau mengubah saiz gambar. Dalam fasa pengambilan ciri, sistem mengambil ciri-ciri yang bermakna daripada gambar yang boleh digunakan untuk membezakan satu buah dari yang lain. Ciri-ciri ini termasuk warna dan bentuk. Dalam fasa pengkelasan, sistem menggunakan ciri-ciri yang diambil untuk mengkelaskan buah-buahan menggunakan algoritma pembelajaran mesin - satu rangkaian neural pembalikan (BPNN). Akhirnya, dalam fasa keputusan dan penilaian, sistem menilai ketepatan dan keberkesanan proses pengkelasan. Secara keseluruhannya, pendekatan kami menunjukkan pendekatan yang terperinci dan sistematik dalam mengembangkan sistem pengenalan buah menggunakan teknik pembelajaran mesin.

CHAPTER 1: INTRODUCTION

1.1 Introduction

Malaysia is a great melting pot of cultures and cuisines due to its different races of people. Malaysia also known as paradise for those who love the local tropical fruit. This is mainly made possible by the year-round sunny weather and regular rains in the land of country (Staff, 2015). While certain seasonal fruits are only available during certain seasons, others are easily available all year round. Seasonal fruits that are well-known are durian, rambutan, mangosteen, pomelo and so on. Besides, bananas, papayas, guavas, and pineapples are non-seasonal fruits that may be served all year round. These fruits become unique symbol that represent Malaysia with its position (Wong, 2021). These local fruits are generally affordable and can be easily found at local supermarkets, road stalls or even night markets during the peak season.

Thanks to Malaysia's fertile land and sufficient water resources, we have been able to harvest a huge variety of delightful and fresh local fruits. The fruit festival at Kota Padawan' was jointly organized by Padawan Municipal Council and Mambong BN Youth Wing. Local fruit festivals should be held more frequently in Sarawak as they have the potential to be a tourist attraction, said Deputy Tourism Minister Datuk Dr James Dawos Mamit (Lee, 2012). Those events accumulate some experiences for the organizer to become host. Therefore, we should use our advantages to promote the local fruit toward tourists and foreigner as attraction to stimulate the tourism economy in our country.

Fruit recognition is important to assist foreigners and tourists as many local fruits in Malaysia may looks strange to them. Foreigners may confuse to figure out what types of fruit in their first sight on it. A fruit recognition application can help them with the problems. The

system would improve the agricultural sector of the country as well. Fruit recognition from images is now possible thanks to classifications based on deep learning from existing projects in this field. Therefore, the technology of machine learning is applied to design and develop an application that fulfills such requirements to the users.

We are living in the era of data. The internet of things (IoT) and Artificial Intelligence (AI) have become the common technologies in our daily life. Therefore, it is necessary to come up with a system that recognizes the fruit types based on its appearance to the foreigners. A highly scalable system that uses machine learning for distributed data processing may be created using Artificial Intelligence of things (AIoT), the merging of AI and IoT (Boesch, 2022). A fruit recognition application will be able to improve the learning process experience on local fruits and even the culture since it helps people to comprehend fruit types efficiently.

A few research studies that focused on such topic are providing solutions for the detection and classification of local fruits. The reviewed techniques were using machine learning techniques using Neural Network for the classification (Sakib et al, 2019). However, classification of fruits faced several complex challenges during development such as the wide variety, irregularity, and inconsistency in shape, color, and texture (Africa et al., 2020). Using a fruit recognition application would definitely assist the users to differentiate the types of fruit by its features and labels. In this project, the technique of machine learning is chosen instead of deep learning. The results to the users are from a process of machine learning models. Object detection and recognition through image processing are processed by OpenCV. Features extraction helps to find out the most crucial features from the datasets to increase accuracy of the machine learning models (Zita, 2022).

1.2 Problem Statement

Most the foreigner could not recognize the types of fruit by its looks. Foreigner may feel confused due to similar appearance of local fruits. Moreover, the manual inspection of each individual is subjective, the accuracy of the result may affect. This may cause a lack of interest in being familiar with the local fruit among them. Thus, a good introduction to local fruits with guidance will be necessary. These attract foreigners' attention in their journey meanwhile promotes local fruit to tourists worldwide. There are existing applications available to help the users. However, those applications failed to process due to the different environmental factors. Different levels of brightness, size of the fruit and angle of images taken are the reasons for those applications not adapted to. Also, some of the applications may not update with the latest fruit sample causing the dataset to be incomplete. Hence, there are improvements that need to be made to increase the accuracy and usability of the application.

1.3 Objectives

The objective of this project is to develop a prototype of fruit recognition application using image processing and machine learning techniques. The technique of image processing is to identify each fruit features and parameters that correspond to each type of fruit. The technique of machine learning is to train a model that recognizes different types of fruits based on the features extracted.

- I. To identify fruit features and parameters through image processing.
- II. To identify and interpret an algorithm for fruit recognition using machine learning.
- III. To develop a prototype for fruit recognition application.

1.4 Brief Methodology

The tools and skills used in this project involved Python programming language and library, image processing and machine learning technique are used. To design and develop a prototype based on algorithms, several phases are required and brief in this section. The details will be explained in Chapter 3.

i. Image Acquisition

This step will be the foundation step in the project. This is a process in which the unprocessed target fruit images are captured and later undergoes analysis purpose. The step needs to ensure the collected images are valid and consistent for the next phase.

ii. Image Pre-processing

Image collected will be converted into a standardized format that eases image processing. The main purpose of this phase is to segment the collected images from the background image to obtain the target image that contains useful information of fruit which is the fruit features.

iii. Feature Extraction

The target fruit images will be analyzed to identify each fruit type by its unique features. Color and shape features will be extracted. The pattern of the unique features becomes the parameters to classification for each fruit. Then, the parameter of features will be extracted for model training and testing later.

iv. **Classification**

The images are categorized into training and testing image dataset. The process of classification involved model training and result testing. The model will first train before it is used to predict the result using testing image. Next, the result presented is being recorded and will be discussed to identify whether the fruit is recognized correctly based on the results.

v. **Results and evaluation**

Result from the trained model is recorded and the accuracy is calculated to get a complete fruit recognition result. Evaluation is made to discuss the prediction result and improvement on machine learning model.

1.5 Scope

The scope of this project focuses on the implementation and evaluation of a fruit recognition prototype using machine learning. In terms of fruit recognition, the scope of the prototype is focused on three specific fruits which are Durian, Rambutan, and Mango. These fruits have been selected as the targets based on the fruit's distinct visual characteristics and the availability of image datasets for training and testing purposes.

1.6 Significance of Project

This project application helps users to recognize the fruit by identifying the fruit features. Users especially foreigners and tourists come and visit our country for the vacation. In their journey, the variety of local fruit may look strange to them due to their appearance including their shape. Users would be able to identify the types of fruits even if they see the fruits for the first time using the application. Users will have to upload the images of fruit taken by the device camera. After that, the images will be processed and passed to analysis through the machine learning model. The results of the fruit types will be shown on the screen for display.

1.7 Expected Outcomes

At the end of the project, the expected outcome would be a fruit recognition prototype that predicts fruit images accurately to identify and classify fruit types. The system will utilize image processing techniques, feature extraction methods, and machine learning algorithms to achieve a high level of recognition accuracy for the selected fruits, Mango, Durian and Rambutan. The project outcome also includes conducting a performance evaluation of the fruit recognition prototype by using confusion matrix. Next, the project will produce a well-documented report that outlines the entire development process, including the methodology, algorithms used, implementation details, result evaluation and future works.

1.8 Project Schedule

The FYP1 started from 11st October 2022 until 25th January 2023. By using Gantt chart, the project plan for FYP1 is shown in appendices section. While FYP2 is started from 20th March 2023 until 24th June 2023, the Gantt chart for project plan for FYP2 will be shown as well in appendices section.

1.9 Outline of Project Report

There are a total of six chapters in this report.

Chapter 1: Introduction

In this chapter, a brief introduction of the project is given including the overview of fruit recognition with its background and importance. It also consists of problem statement, objectives of project, methodology and procedures adapted, scope, significance of project, project schedule and expected outcomes.

Chapter 2: Literature Review

This chapter will review the used techniques of existing fruit recognition system. Each of the existing techniques is discussed with its benefits and disadvantages. A comparison among the existing techniques will be made and come up with a proposed solution in the end.

Chapter 3: Methodology

In this chapter, methodologies employed in the project are presented and discussed in detail. The planning of procedures will be arranged in sequence. Next, some requirement analysis will be done to see whether the methods and techniques used fulfill the specified objectives and scope.

Chapter 4: Implementation and Testing

This chapter focuses on implementation of image processing and machine learning techniques approaches to the fruit recognition system. Detailed explanation of implementation of model training and prototype will be included in this chapter. After that, the prototype of the system is developed and tested to ensure that the prototype meets the requirements.

Chapter 5: Discussion

This chapter discusses the evaluation of the trained model within the fruit recognition prototype. The main purpose is to focus on assessing the model's performance in accurately recognizing different types of fruits.

Chapter 6: Conclusion and Future work

A summary of the project is made to conclude the chapters above. Achievement and limitations are included in this chapter. Future works to be done for improvement are discussed.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter illustrates the background study from existing fruit recognition techniques in the field. From background studies, computer vision is usually applied in most of the fruit recognition techniques to identify the fruit types. Machine learning uses algorithmic models that enable a computer to show itself about the settings of visual information. A well-fed model with enough data will teach itself from one image to another. Thus, the algorithms enable the machine to learn itself instead of programmed to recognize an image (IBM, n.d.). This technology is used widely for a variety of purposes which are surveillance, motion capture, machine inspection, biometrics, optical character recognition (OCR), medical imaging, automotive safety and so on. To make computer vision work, lots of data, tons of images, and various analyses are required to differentiate and recognize the objects (Sharma, 2021). During computer vision, Image Acquisition is a process that obtains images in digital form whether is from image sensor from camera or scanner (Singh et al., 2019). Since digital images exists in different format with different size, grayscale in color and more, standardization process named Image Pre-processing must be made before sending them into the machine learning model (Fredrick, 2021). Image analysis and classification is the method involved with extracting significant features from digital images such as shapes tracking, RGB identification or object properties recognition (Mathworks, Image Segmentation and Analysis, n.d.). This chapter will focus on the different types of approach in computer vision technique with machine learning model.