COMPUTER VISION ESTIMATION OF VOLUME AND WEIGHT OF APPLE BY USING 2D NONCONTACT MEASURING METHOD

AFIQ IKHWAN BIN MOHD FAUZI

B.ENG (HONS.) ELECTRICAL ENGINEERING (ELECTRONICS)

UNIVERSITI MALAYSIA PAHANG

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name   : AFIQ IKHWAN BIN MOHD FAUZI

Date of Birth   : 14 APRIL

Title   : 1E16 COMPUTER VISION ESTIMATION OF VOLUME AND
                            WEIGHT OF APPLE BY USING 2D NONCONTACT
                            MEASURING METHOD

Academic Session   : SEM 2 2021/2022

I declare that this thesis is classified as:

☐ CONFIDENTIAL   (Contains confidential information under the Official
                       Secret Act 1997)*
☐ RESTRICTED   (Contains restricted information as specified by the
                       organization where research was done)*
☑ OPEN ACCESS   I agree that my thesis to be published as online open access
                       (Full Text)

I acknowledge that Universiti Malaysia Pahang reserves the following rights:

1. The Thesis is the Property of Universiti Malaysia Pahang
2. The Library of Universiti Malaysia Pahang has the right to make copies of the thesis for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

_____   _____
(Student's Signature)   (Supervisor's Signature)

                                    Ir. Ts. Dr. Mohd Zamri bin Ibrahim

_____   _____
New IC/Passport Number   Name of supervisor
Date: Date: 27 June 2022

NOTE: * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.

THESIS DECLARATION LETTER (OPTIONAL)

Librarian,
Perpustakaan Universiti Malaysia Pahang,
Universiti Malaysia Pahang,
Lebuhraya Tun Razak,
26300, Gambang, Kuantan.

Dear Sir,

CLASSIFICATION OF THESIS AS RESTRICTED

Please be informed that the following thesis is classified as RESTRICTED for a period of three (3) years from the date of this letter. The reasons for this classification are as listed below.

Author's Name	AFIQ IKHWAN BIN MOHD FAUZI
Thesis Title	COMPUTER VISION ESTIMATION OF VOLUME AND WEIGHT OF APPLE BY USING 2D NONCONTACT MEASURING METHOD

Reasons	(i)
	(ii)
	(iii)

Thank you.

Yours faithfully,

_____
(Supervisor's Signature)

Date:

Stamp:

Note: This letter should be written by the supervisor, addressed to the Librarian, *Perpustakaan Universiti Malaysia Pahang* with its copy attached to the thesis.

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Electrical Engineering (Electronics).

_____

(Supervisor's Signature)

Full Name   : IR. TS. DR. MOHD ZAMRI BIN IBRAHIM

Position   : SENIOR LECTURER

Date   : 27 JUNE 2022

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

_____

(Student's Signature)

Full Name   : AFIQ IKHWAN BIN MOHD FAUZI

ID Number   : EA18090

Date   :

COMPUTER VISION ESTIMATION OF VOLUME AND WEIGHT OF APPLE
BY USING 2D NONCONTACT MEASURING METHOD

AFIQ IKHWAN BIN MOHD FAUZI

Thesis submitted in fulfillment of the requirements
for the award of the
B.Eng (Hons.) Electrical Engineering (Electronics)

College of Engineering

UNIVERSITI MALAYSIA PAHANG

JUNE 2022

ACKNOWLEDGEMENTS

In the Name of Allah, the Merciful, the Compassionate.

First and foremost, I want to express my sincere gratitude to several individuals and organizations for supporting me through this journey. First, I would like to express heartfelt appreciation to my supervisor Ir. Ts. Dr. Mohd Zamri bin Ibrahim for his passion, patience, insightful remarks, valuable information, practical guidance, and never-ending ideas, all of which have greatly aided me during my Final Year Project and writing of this thesis. It was a great honour and joy to work and study under his guidance. I appreciate his constant support in completion of this project. Without his assistance, completing the Final Year Project will be more challenging.

I offer my greatest motivators, my parents, my gratitude. I appreciate your words of support and all of the daily prayers you have said for me. Thank you so much for always being there for me. I appreciate all of my family members' encouragement and support. Thank you for assisting me in surviving all of the stress and for refusing to let me quit.

I also would like to offer my gratitude to all my friends who have been there all the time. They deserve gratitude for assisting me in finishing the thesis when I was overwhelmed and bewildered, as well as for their steadfast support. Thanks to my wonderful friends who refused to let me give up.

Finally, I would like to express my gratitude to everyone who was directly or indirectly involved in the completion of my Final Year Project, I much appreciate your efforts. Thank you very much.

ABSTRAK

Epal merupakan antara buah yang terkenal yang telah dipilih oleh ramai keluarga di Malaysia. Terdapat beberapa faktor yang menentukan kualiti epal iaitu isipadu dan berat. Kedua-dua parameter tersebut telah digunakan sebagai penanda aras untuk mengenal pasti kualiti epal. Kedua-dua parameter ini boleh diukur dengan mudah secara individu dengan menggunakan neraca penimbang untuk mengukur berat dan kaedah anjakan air untuk mengukur isipadu. Walau bagaimanapun, kedua-dua kaedah ini tidak sesuai digunakan dalam industri kerana kedua-dua kaedah memerlukan banyak masa untuk mendapatkan hasil akhir. Oleh itu, teknologi baru diperlukan untuk menganggarkan berat dan isipadu epal. Objektif utama projek ini adalah untuk membangunkan sistem berasaskan penglihatan yang mampu menganggar isipadu dan berat epal. Isipadu epal dikira dengan menggunakan lebar dan ketinggian epal melalui imej. Kamera perlu menentukan untuk mendapatkan nisbah piksel/cm yang digunakan untuk penukaran data. Detectron2 dengan set data tersuai akan digunakan untuk mengesan dan membahagikan imej epal. Output imej pengesanan dan pembahagian akan digunakan untuk mencipta kotak sempadan untuk memperoleh ketinggian dan lebar epal. Sistem yang dibangunkan diuji oleh empat tetapan berbeza yang mempunyai pembolehubah berbeza. Isipadu dan berat epal mampu memperoleh ketepatan 85.41% dan 85.26% dengan menggunakan formula. Secara keseluruhan, projek ini dapat menganggarkan semua parameter epal dengan ketepatan lebih daripada 80% yang boleh disimpulkan sebagai projek yang berjaya.

ABSTRACT

Apple is one of the most famous fruits that have been picked by a lot of families in Malaysia. There are several factors that determine apple quality which are volume and weight. The two parameters have been used as a benchmark to identify the quality of apples. These two parameters can be easily measured individually by using a weighing balance to measure weight and water displacement method (WDM) to measure volume. However, these two methods are not suitable to apply in industries since both methods required a lot of time to obtain the final output. Therefore, new technology is needed in order to estimate the weight and volume of apples. The main objective of this project is to develop a vision-based system that is able to estimate the volume and weight of apples. The volume of the apple is calculated by using the width and height of the apple via images. The camera needs to calibrate in order to get the ratio of pixel/cm which is used for conversion data. Detectron2 with a custom dataset will be used to detect and segment apple images. The output of detection and segmentation images will be used to create boundary boxes to obtain the height and width of apples. The developed systems are tested by four different settings which have different variables. The volume and weight of apples are able to obtain 85.41% and 85.26% of accuracy by using the formula. Overall, the project is able to estimate all the parameters of apple with an accuracy of more than 80% which can be concluded as a successful project.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	2-Dimension
3D	3-Dimension
AI	Artificial Intelligence
AIA	Automated Imaging Association
CCD	Charge-Coupled Device
COCO	Common Object in Context
CVS	Computer Vision System
GPS	Global Positioning System
IDE	Integrated Development Environment
LSVSM	Least Squares Support Vector Machine
MP	Megapixel
OID	Open Images Database
PLS	Partial Least Square
VGG	Visual Geometric Group
VIA	VGG Image Annotator
WDM	Water Displacement Method

CHAPTER 1

INTRODUCTION

1.1   Project Background

Computer Vision System (CVS) is one of the advanced technologies that are still not widely explored and implemented in daily life. Computer vision is a branch of artificial intelligence (AI) that allows computers and systems to extract useful information from digital photos, movies, and other visual inputs [1]. Both systems are depending on each other which AI enables computers to make decisions while computer vision enables the systems to see and observe. The systems work the same as human vision. However human sight has advantages from lifetimes of context to learn how to distinguish objects apart, how far away they are and many more. While computer vision trains machines to do these functions with much less time and higher accuracy.

Computer vision has a huge influence on many sectors from retail to security, healthcare, automotive, industrial automation, and agriculture [2]. The huge number of applications is due to the system's ability to automate several tasks much faster. Agriculture is one of the highest sectors that monopolize computer vision. This is due to the concern's observation on the product. Moreover, it is also used for quality monitoring and inspection in agriculture products and food processing. The applications on fruits and vegetables vary from defect detection on fruits, colour inspection, automatic classification, size estimation and many more.

In this project application on computer vision are used in order to estimate volume and weight of an apple. Apple has been chosen as the model due to vary reasons. One of the reasons is due to there are many varieties of apples that exist around the world. Based on the statistics in 2018, apples are places on number 4 of the most harvested fruits in the world with 86.14 million tons [3]. Furthermore, there are 7500 known cultivators of

apples which have various methods and tastes. Due to this, the quality of apples is very important. The volume and weight of apples are two essential characteristics in determining their exterior quality. Apple volume and weight metrics are not only used to influence customer preferences and hence marketing values, but they are also used to determine apple quality. In agricultural goods and food processing, computer vision systems are frequently employed for quality monitoring and inspection.

Moreover, in this project application of visual data that obtains from camera also will be used in this project. There are many types of cameras that have been applied in many sectors especially related to electronics such as Gaming, Mobile phones, drones and many more [4]. However, application of this technology in agriculture sectors is still not widely used. It is not an argument that this technology can boost the current method to another level.

1.2 Problem Statement

Computer Vision is one of the possible solutions of the most major errors in industry which are due to human error. This is due to the fact that human vision is unable to work all the time. In industry, it is crucial to have a constant workforce in order to maintain the production rate. This problem is undeniable and requires help from machines to overcome this problem. By applying computer vision in industry, the production rate is much higher due to machines able to work much longer with constant levels of concentration. As a result, high-quality products and services will be delivered more quickly.

Moreover, the parameters of the Apples which are able to determine the quality of the fruits such as weight and volume are very important. Various methods that have been used in order to estimate the volume and weight of an apple with the help of many instruments. This method is undeniably able to achieve the objective, however the speed of this method is time consuming since it needs to be measured one by one using multiple instruments for each parameter. This method is great to use for any applications that do not prioritize speed as the main concern such as night markets or small grocery shops. However, most factories required speed to process in order to obtain high profits. Therefore, this method is not suitable.

Lastly, lag in applying newer technology. Computer vision is the latest technology which uses visual data to do analysis on the object. This technology has been used in many other departments such as automotive, medical, automation, electronics and many more. All the departments show positive results and achievement by applying this technology. Therefore, it is very wasteful to not apply this technology in agriculture sectors.

This paper is searching for methods which provide much higher accuracy on estimation while reducing the complexity of the system by exploiting the new technology which is analysis with computer vision data.

1.3   Objective

The objective of this project can be listed as below:

- To detect and segment the apple models.

- To develop a vision-based system to estimate the volume and weight of apple.

- To analyze accuracy of the systems with the manual measurement

1.4   Project Scope

One of the most important things for this project is the data preparation. The current dataset is prepared by open source such as COCO Dataset and OID. This dataset contains millions of pictures and thousands of models which is very useful for this project. However, in this project there is only one model (apple) that is needed. Therefore, a custom dataset needs to be created in order to accomplish the first objective. VGG Image Annotators are used to prepare the custom apple dataset. Detectron2 is used to detect and segment the apple models from the picture in order to obtain the apple parameter.

There are multiple types of visual data in this world that are taken using different tools and technology. Each technology has its own capabilities and deficiency depending on the situation used. The current industries use many types of high end cameras such as CCD and CMOS. This camera contains multiple technologies which are able to take advantage in industries. However, in this project, only phone cameras with different resolutions are used for this project. The camera consists of 16MP and 106MP to collect test data of 20 apples.

Last but not least the results obtained from the systems are compared with the actual data that obtains from the instrument such as vernier calliper and weight scale. This has to be done to analyse the accuracy and the error of the developed systems. Moreover, from this result it also can show the capability of the developed systems and its qualification before applying in industries.

CHAPTER 2

LITERATURE REVIEW

2.1   Introduction

This chapter briefly covers the literature review and conclusion which are related to this research. There are three main parts in this chapter which are estimation volume and weight, computer vision systems and apple volume and weight calculation. All basic information is covered in this chapter in order to accomplish this research. In order to develop a system that exceeds the current technologies capabilities, other research which is related needs to take consideration.

This project will focus more on computer vision estimation of volume and weight of an apple by using a phone camera. Therefore, techniques that are used are taken into consideration in order to improve the capability of the systems. Following that, the approaches utilised in this work will be further examined based on the findings of other researchers in order to develop systems that can accurately estimate volume and weight of the models. Moreover, approach on detect and segment the model also be discussed in this chapter which also included other algorithms that related to this project. For this project an additional feature is added to improve the capabilities of this project and to conduct experiments on application of phone cameras.

2.2 Volume and Weight Estimation

2.2.1 Relationship between Mass, Weight and Volume

Objects are described using parameters such as mass, weight, and volume. The terms mass and weight, in particular, are frequently used interchangeably to represent the same thing, despite the fact that they have very different meanings. The mass of an object, whether liquid, gaseous, or solid, refers to the amount of substance it contains. It's usually expressed in kilograms or grams, and it's a constant amount that doesn't change no matter where an object is located. Weight refers to the pull of gravity on an object. Since the gravity differs through the solar system, an object's weight doesn't remain constant (5). The force for gravity greatly affects the mass of an object. This can be explained by the equation below.

$$F = m \times a \qquad\qquad 2.1$$

This is a general equation that relates force to mass and acceleration. Force (F) represents the object's weight in Newtons (N) while mass (m) in the equation represents the object's mass in kilograms, and the letter a represents the downward acceleration due to gravity. As this equation demonstrates, mass and weight are inextricably linked. The weight of an object grows in proportion to its mass. When mass doubles, for example, weight doubles as well.

Volume refers to the amount of space that an object takes up. Solid volumes are measured in meters cubed or centimeters cubed, while liquid volumes are measured in literals or milliliters.

Weight (W) is a product of mass and gravity (g) which lead to the following equation.

$$W = m \times g \qquad\qquad 2.2$$

In addition, volume is related to density (ρ) which measures the mass of a substance per unit volume as follows equation.

$$\rho = \frac{m}{V} \qquad\qquad 2.3$$

From the equation above that have been shown, volume and mass can be related by the following equation

$$V = (W/g)/ \rho \qquad\qquad 2.4$$

In conclusion, by applying this formula it can be saved that volume and weight are related to each other which are two parameters that are sought in this project. Therefore, it can be concluded that by applying mathematical techniques both parameters can be achieved which greatly reduces the complexity of this project (6).

2.2.2 Weight calculation by volume

As mentioned before, weight and volume are related to each other. However, there is a point of intersection that connects these two parameters which is density (ρ). The mass of a material substance per unit volume is called density. Density is calculated using the following formula.

$$\rho = \frac{m}{V} \qquad\qquad 2.3$$

The parameter above represent as follows, where ρ represents density, m is mass, and V is volume. The density of a substance is usually measured in grams per cubic centimetre. This parameter is able to correspond to another parameter. Moreover, most solid objects already have a constant value for density.

Therefore, in order to convert the calculated weight by volume, it is possible just by applying the formula above. As for the density value, the model that was used in this project is apple with the density value as follows (6).

$$\rho = 0.96 g/cm3$$

2.2.3 Weight and volume application

Weight and volume are parameters that are considered important in many sectors. Various applications have been applied in order to achieve certain objectives. Moreover, this parameter provides important information which enables us to quantify many things in life such as characterizes, represent, and predict certain items, concepts or outcomes.

In terms of automotive, volume and weight greatly impact this sector. It is due to determine the best design which is able to monopolize every condition. For example, it is able to determine the fuel consumed. The heavier the vehicle, the more energy required in order to get moving. Other than that, it also acts as a safety feature since vehicles that have higher weight value are able to increase vehicle rolling resistance and provide better crash protection. This shows that the weight of each component in a car is a very important parameter.

In terms of the food sector, volume and weight also impact this sector. This is due to exact values which can be seen in packaging. All the food products are precisely calculating the exact value of each ingredient in the product. This information is important which needs to be measured precisely in order to create great food. Moreover, weight and volume also determine the quality of agricultural products. This information greatly impacts on many factors which determine such as the value price and the category. This shows that it is a very important weight and volume parameter in industry.

2.3   Computer Vision System

If artificial intelligence enables computers to think, computer vision enables them to see, observe and understand. Computer vision works much the same as human vision, except humans have a head start. Human sight has the advantage of lifetimes of context to train how to tell objects apart, how far away they are and whether they are moving.

Computer vision is used in industries ranging from energy and utilities to manufacturing and automotive and the market is expected to reach $48.6 billion by 2022. Computer vision can analyze thousands of products or processes a minute, noticing imperceptible defects or issues, and it can quickly surpass human capabilities.

2.3.1   Computer Vision and Machine Vision

Computer Vision is a vast system that covers a lot of important things which enable it to function and work for different purposes. First and foremost, it is important to recognize the difference between computer vision and machine vision. Machine vision, according to the Automated Imaging Association (AIA), covers all industrial and non-industrial applications in which a combination of hardware and software provides operational assistance to devices in the execution of their activities based on image

capture and processing (2). Computer vision is a sort of artificial intelligence that allows computers to interpret and analyse visual data, emulating how people see and comprehend their surroundings.

The similarities of both systems are that they acquire image capturing and analysis to perform tasks that are able to surpass the capabilities of human vision in terms of speed and accuracy. The term "computer vision" refers to the automation of image capture and processing, with a focus on visual analysis. In other words, Computer Vision's purpose is to process and offer relevant conclusions based on the observation, not only to see. Machine vision is a subsection of computer vision that pertains to the use of computer vision in industrial settings.

There are four main categories in machine vision which are guidance, identifications, gauging and inspection (7). Each category has a much more specific and detailed process. Image segmentation, object detection, facial recognition, edge detection, pattern detection, image classification, and feature matching are the categories of computer vision.

2.3.2 Computer Vision Benefits and Applications

This technology is pursued due to the several benefits that are provided such as faster and simpler processes. This is due to capabilities in carrying out monotonous and repetitive tasks in a single time at a faster rate. Moreover, it also has better products and services which due to the possibility of receiving errors are very low. Lastly it is also a method to reduce cost reduction by reducing fixing on flawed products due to errors.

Computer vision is a very competent system that is able to assist human workload in many major categories. There are many sectors that have applied computer vision such as manufacturing, healthcare, agriculture, transportation and many more. Many applications of computer vision that applied in these sectors that able to help human surpass common capabilities in order to achieve better results.

Throughout past years, image processing and computer vision technology have gained significance and potential in a variety of fields, including agriculture. Several researches have been done which apply computer vision such as defect detection on apples [8,9], common defect detection on citrus [10], size assessment of berries [11],

defect detection on bananas [12], automatic classification of fruits [13], size estimation of sweet onions [14] and color inspection of various fruits and vegetables [15,16].

2.3.3   Computer Vision Volume and Weight Estimation

In the past, the volume and weight of an apple were determined by using eyes and hands to categorize the type of apple. Even with the creation of weighing balance this method is still used in determining the fruits types in many major fruit orchards. This method is also still applied in fruit processing factories. It is clear that this method is quite subjective which has lower accuracy and speed processing.

Using noncontact measurement methods, various computer vision technologies have been studied to estimate the volume and weight of agricultural items. Because the 2D digital images captured by computer vision systems are made up of pixels, image processing techniques can be used to measure the projected area, perimeter, or length and breadth properties in the images. The projected area is the most often utilised picture characteristic and basic simple measurement for volume and weight evaluation.

Ellipsoid approximation and image processing have been used to determine volume of watermelon [17]. In this system, the key points in determining the volume lies on the length and major and minor diameters of watermelon. The data are collected with eight-bit surface images to be used as an image processing method. Using the paired t-test and the Bland–Altman methodology, the volume obtained through ellipsoid approximation and image processing was compared to the volume calculated by the water displacement method. The volume calculated using the ellipsoid approximation differed considerably from the volume calculated using water displacement ($P < 0.05$), whereas the volume calculated using image processing did not differ significantly from the volume calculated using water displacement ($P > 0.05$).
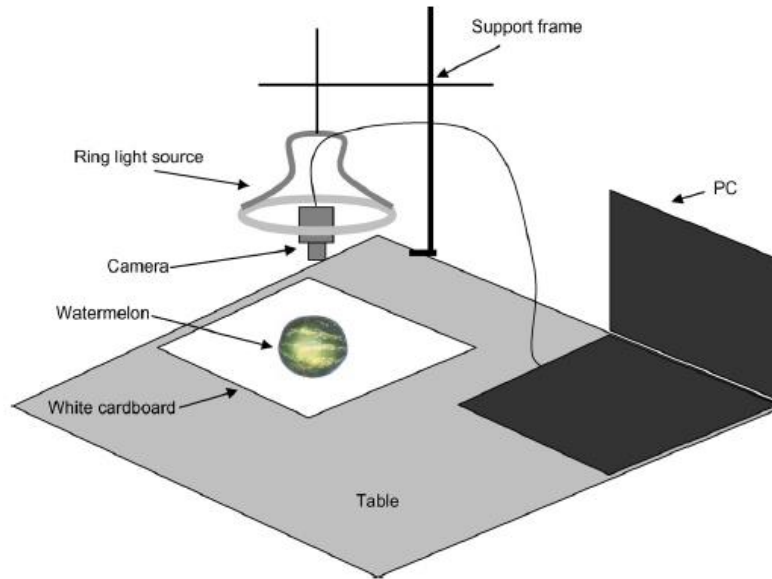
Figure 2.1: Ellipsoid Approximation and Image Processing

Images and analysis techniques for estimating weight of Chokanan mangoes [18]. This method also uses the same concept as before which utilised captured images. This method counted the number of pixels of mango reign in the captured image to determine the weight by analyse using statistical method of regression. The process showed the result of correlation coefficient value of 0.9769 with mean absolute percentage error 3.76%. Both methods which utilised 2D-imaging techniques able to estimate accurately due have strong correlation with some dimensional parameters of 2D projected area. Moreover, Volume estimation of strawberries, mushrooms, and tomatoes with a machine vision system (19) using the same method to determine the volume of models. The only difference between this methods is the technique to analyze 2D-data imaging.
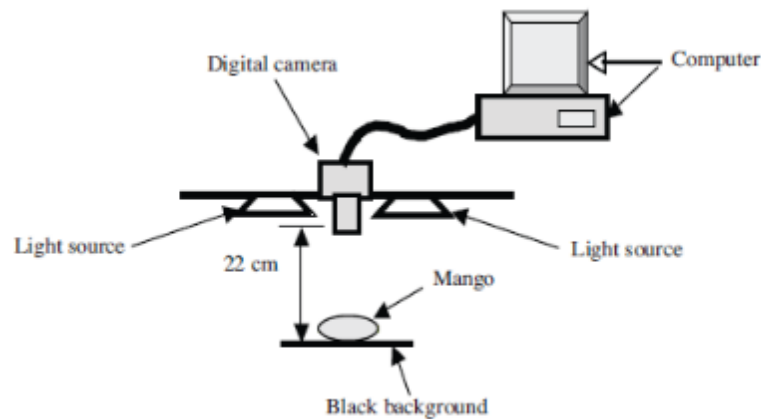
Figure 2.2: 2D-Imaging Technique

Computer vision systems consist of 3 types of systems which are 1D Vision System, 2D Vision Systems and 3D Vision Systems [1]. Each system has its capability and flaws. In estimating the volume and weight of objects the most suitable systems that can be used are 2D Vision Systems and 3D Vision Systems. This is due to the axis that is able to determine the coordinate of the picture. The vision systems work by applying a mapping system in which points are coordinated. This point contains various information which can be used to determine a lot of things. The major difference between this system is the axis existence in the model's data. By applying 3D- Vision System, a new axis is added which is the Z-axis. This axis is able to collect height data of objects which is a major influence that is able to manipulate the output data [8].

Various sensors and techniques could be employed to acquire 3D images of the items. The most frequent ways to generate 3D images in 3D detection are binocular stereo vision techniques, which are based on binocular CCD cameras. This method obtains Z-axis information by using a pair of cameras [15]. It is undoubtedly that this method is able to estimate volume and weight of models accurately but this method is not suitable to apply due to the complexity of the systems.

Other than that, a technique that obtains 3D images by extracting and sorting the features [20]. This system worked with multiple image's view of models. Therefore, the data are needed in multiple angles to achieve the objective of the data. The obtained data are segments in order to extract silhouette regions of models. The combinations of these silhouettes will produce a course of 3D shapes of the models. The volume of the models can be estimated by using the coarse model.

22
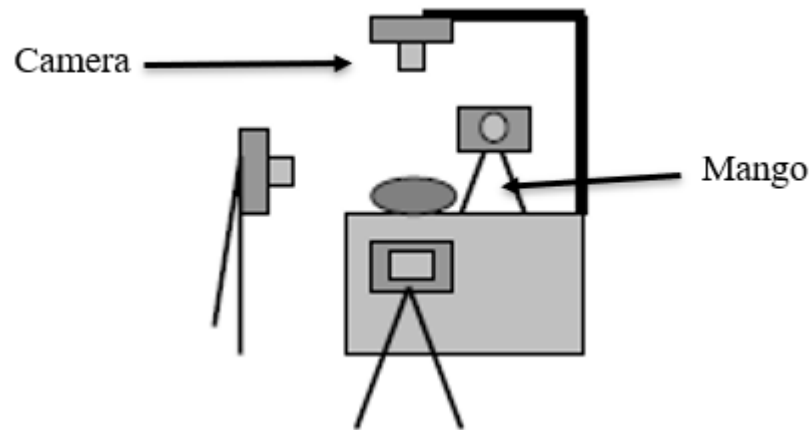
Figure 2.3: 3D Images by Extract and Sorting the Features

The results by applying in this method are compared with the data obtained by using laboratory instruments. The accuracy of the data is slightly lower than other methods which is 87.65 %. This result shows that the accuracy is lacking in these systems and further improvement needs to be done. Moreover, in order to obtain the data, multiple image views are required in order to satisfy data requirements. Therefore, this method is not suitable to apply in industries due to time consuming and the cost of the setup.

Other than that, a method that has been developed to estimate mass and volume of citrus fruits by applying image processing [21]. Consider a three-dimensional model of an axis-symmetric item. The image of the item can be divided into a number of right cone frustums. The lateral surfaces are considered to be elliptical to improve system accuracy. These basic cylindrical objects are assumed to be of equal pixel height, h, without sacrificing generality.
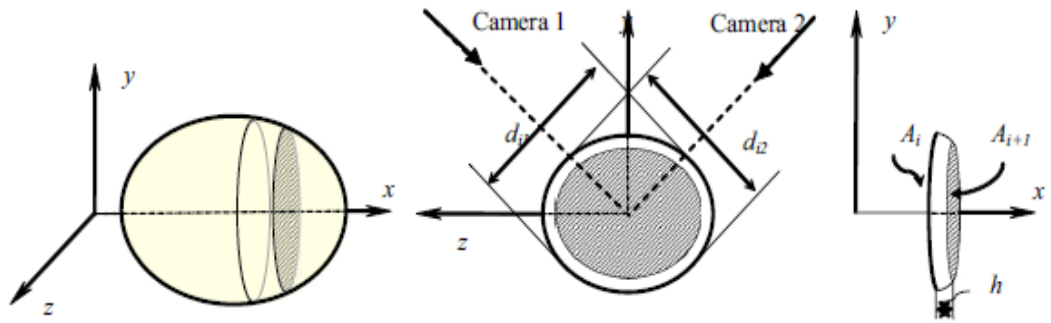
Figure 2.4: Cone Frustums

By aggregating these constituent volumes, the entire volume of the object under inquiry can be calculated. The top and bottom diameters, as well as the height of the frustum, are needed for dimensional properties. Using the two perpendicular diameters measured by the two cameras, the cross-sectional areas through the elliptical frustum may be determined precisely. Two cameras are used in this technique to provide perpendicular images of the fruit. To simplify the volume are calculated by dividing the fruits image into a number of elementary elliptical frustums. Each elliptical frustum is determined by the volume and the total volume are represents the volume of citrus fruits.

Figure 2.5: Estimate Mass and Volume of Citrus Fruits by Applying Image Processing

The above method showed a positive result which is that the computed coefficients of determination for lemon, lime, orange, and tangerine were 0.962, 0.970, 0.985, and 0.959, respectively. Therefore, this method can be concluded to be able to accurately determine the volume of the citrus fruits. However, this method is not suitable to apply in industries due to several reasons. First and foremost, this method is assumed to measure each model individually which is impossible to be done in the industries. Moreover, this method is only able to apply to any fruit which has a high correlation shape to 2D- dimensions. Therefore, this method is limited to a certain number of fruits.

3D Reconstructions are also one of the methods that are developed in order to estimate weight and volume of an apple [22]. This system proposed to reconstruct the 3D surface of the apples by using a single multispectral camera and near infrared linear array light. In order to estimate weight and volume of apples both the traditional images features and height data are extracted. In this method two types of height are extracted in order to combine the data with a projection area to form combination features.
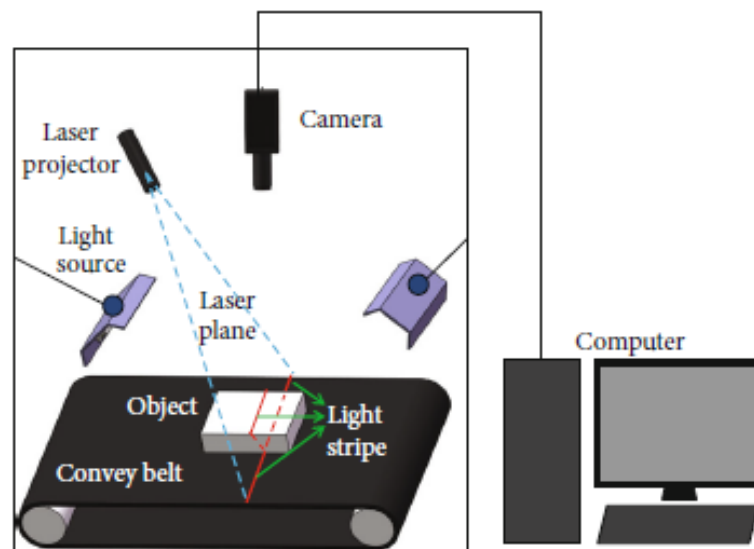
Figure 2.6: 3D Reconstructions Using a Single Multispectral Camera

The 3D surface of the upper half of the inspected apples would be recreated after the scanning. The height map would also be featured in pseudocolor height map and grey level height map graphics to make it more intuitive and visual.
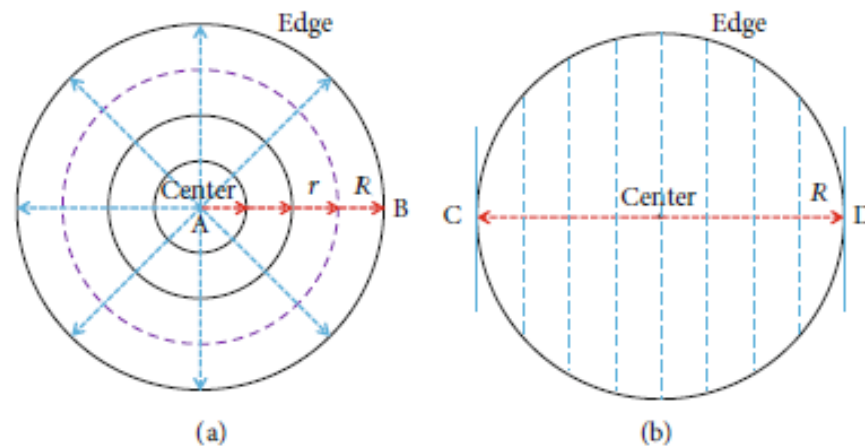
Figure 2.7: Pseudocolor Height Map and Grey Level Height Map

The two types of features are used to determine the height of the model. Two different features are used to extract the model height. The first type of height features (marked as Type I feature) was extracted from the 50 concentric annuli equally distributed in the heightmap with an adaptive distance to the size of the inspected apples by averaging the height values of all pixels as shown in Figure. The second type of height features (marked as Type II feature) was extracted from the 50 vertical lines equally distributed in the height maps with an adaptive distance to the size of the inspected apples by averaging the height values of all pixels as shown in Figure.

After obtaining the feature, the process is proceeded with applying multiple algorithms to determine the volume of the apples. One of the algorithms is Partial Least Square (PLS) which are used to determine original independent information. This method consists of two steps which are a decompose matrix and process by linear regression. The ideal amount of PLS components to optimise the model's prediction power should be established in PLS analysis. Cross validation is commonly used to make this decision.

Other than that, Least Squares Support Vector Machine (LV-SVM). LSSVM, a cutting-edge statistical learning method, can handle both linear and nonlinear multivariate analysis and solve them quite quickly. Furthermore, with minimal training data, the support vector machine (SVM) may learn in a high-dimensional feature space. To derive the support vectors, it uses a series of linear equations rather than quadratic programming problems.

The result of the method shows the possibility to reconstruct the 3D-Surface of an apple and estimate the volume and weight. Applying 3D-height data greatly improves the prediction of the systems. In this system LS-SVM models with combination Feature II are suitable to estimate volume while LS-SVM models with combination Feature I are suitable for estimate weight.

From all past study it can be seen there are many types and methods that can be used in order to estimate weight and volume of apples. There are advantages and disadvantages to each method. From past study multiple information has been known and some of them can be used in this study.

2.4   3D-Depth Camera

Components that are able to detect and measure distance are very important since they are able to perform many tasks. Components such as IR Sensor, Ultrasonic sensor and many more are very useful especially in order to develop systems that have detection, measurement and many more [23].

3D-Depth camera is a revolutionary which combines vision and distance sensor to perform. The difference between a regular cameras is the output and working function. Standard digital camera output images as 2D grid of pixels. Each pixel has a set of values associated with it, which is frequently referred to as RGB (Red, Green, and Blue). Each characteristic has a value ranging from 0 to 255. The output of the camera is made up of thousands to millions of pixels.

Meanwhile the 3D-Depth camera has additional parameters which refer to depth. The output of 3D-Depth is able to give all four values which refer to RGBD. A depth camera's output could be shown in a variety of ways. Figure below displayed the colour picture alongside the depth image, with each colour in the depth map representing a different distance from the camera. The colour cyan is nearest to the camera, while red is the furthest away.
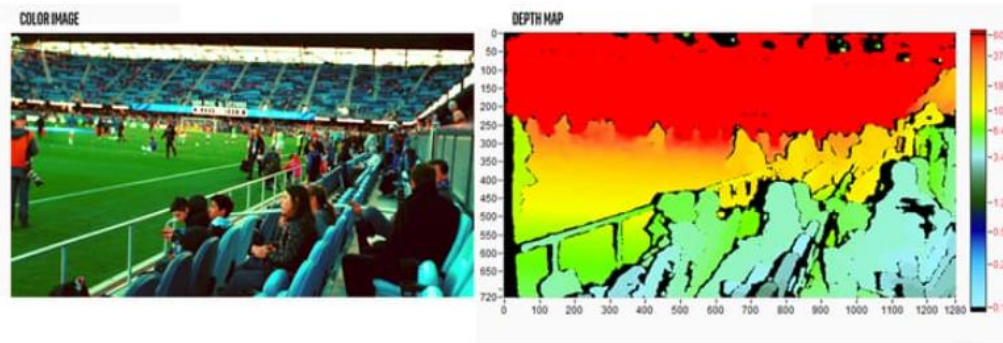
Figure 2.8: Colour Picture and Depth Image Camera Output

Currently there are three types of 3D-Depth camera which are Structured Light, Stereo Depth and Time of Flight (ToF). Each camera has different working function advantages and benefits [24].

2.4.1 Structured Light

Structured light cameras rely on projecting light from emitter onto the scene. The projected lights already determine the specific pattern which enables them to measure the depth information [23]. This information is obtained from sensors that detect the pattern that has been emitted. For example, If the pattern is a sequence of stripes projected onto a ball, the stripes will distort and bend in a precise way around the ball's surface.
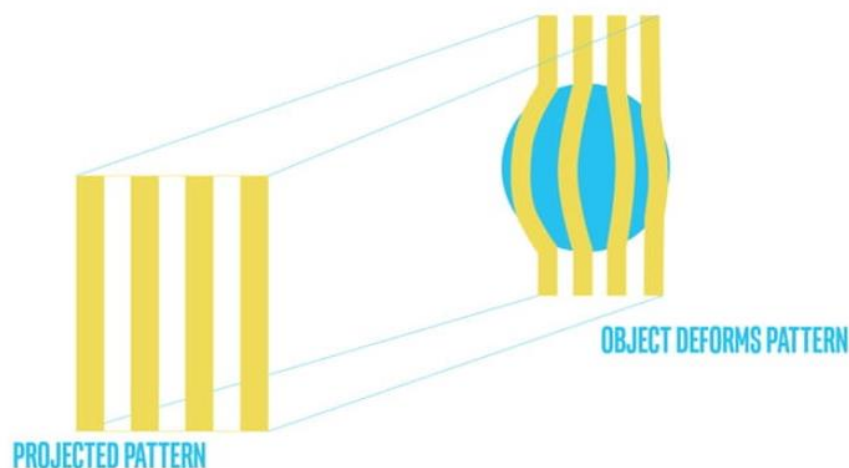
Figure 2.9: 3D-Depth Camera Structured Light

This camera uses light as the main function to measure depth therefore, this camera is able to perform very well in a wide set of lighting conditions, however there is an exception which is sunlight. This is due to avoid the light source overpower the light

that emitted from the sensor. Moreover, it is also sensitive to any other disturbance such as noise since it will disrupt the expected output. Therefore, this camera is only suitable to use in indoor conditions [24]. The best application to use this camera is gesture recognition and background segmentation.

2.4.2 Stereo Depth

Stereo depth cameras are built with two sensors that have a small distance. The distance between these two sensors are the key points to estimate depth value of an image [23]. This camera also emitted light to measure the depth, however, it was able to use any type of light. This camera will take two types of images from each sensor and compare the result of the images. These comparisons provide depth information since the distance between the sensors is known. The working function of this camera is the same with human eyes.
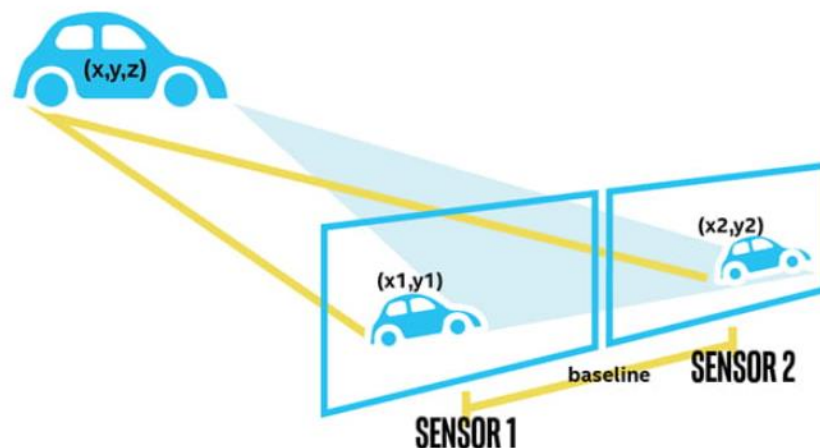
Figure 2.10: 3D-Depth Camera Stereo Depth

As mentioned before, this camera also uses light to measure depth therefore it is able to operate very well in a wide set of lighting conditions. Since this camera is able to use any type of light, sunlight is no longer an obstacle to perform in an outdoor area. Infrared projector also enables the camera to perform well in low light conditions.

2.4.3 Time of Flight (ToF)

The Time of Flight (ToF) camera has the same working function as the IR sensor. This is due to the method of measuring the depth of an image. In this method a constant that already knows is the speed of light. The speed of lights is used to measure the time of flight that is required to finish a cycle. The light is emitted from the sensor and sweeps over the scenes back to the sensor. By using different types of power and wavelength of the light, the sensor is able to measure depth at many different levels. This type of sensor is already widely applied in many aspects such as self-driving cars and map terrain from helicopters.
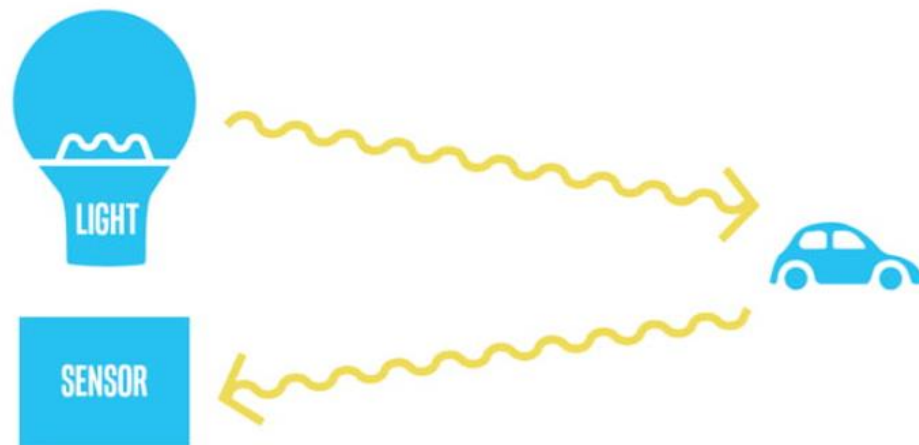
Figure 2.11: 3D-Depth Camera Time of Flight (ToF)

Overall, this type of camera is the best among the other two, however the only primary weakness of this camera is incapability to work in outdoor conditions. This is because any circumstance in which the light contacting the sensor is not from the camera but from another source, such as the sun or another camera, degrades the quality of the depth image. Except for this parameter, this camera has higher accuracy in measuring depth and wider range capabilities.

2.5 Summary

Generally, there is a little research that has been done in this subject. It can be summarized in the table as follows. Multiple results that are obtained from the research depending on the project scope. Most of the result has its own advantages and disadvantages that are able to be learned for this project.

Table 2.1: Literature Review

	Title	Image processing and analysis techniques for estimating weight of Chokanan mangoes		
	Author	C.C. Teoh, A.R. Mohd Syaifudin		
1	**Methodology**	**Achievement**		**Summary**
	Image processing and analysis 1)Image capturing 2)Image analysis filtering and thresholding processes with using PCI software	The mean absolute percentage error of estimating a mango weight using the linear model was 3.76%		the image processing and analysis techniques are practical, feasible and effective in estimating weight of Chokanan mangoes.
	Title	Determination of watermelon volume using ellipsoidapproximation and image processing		
	Author	Ali Bulent Koc		
2	**Methodology**	**Achievement**		**Summary**
	Ellipsoid approximation and image processing methods. 1)Dimensional Calibration 2)Volume evaluation using ellipsoid 3)Volume estimation using projectile	1)ellipsoid approximation and WDM was statistically significant ($P < 0.05$) 2)image processing and WDM were not statistically significant ($P > 0.05$).		As the size of the watermelon increased, the image processing method overestimated the volume.
	Title	Volume estimation of strawberries, mushrooms,and tomatoes with a machine vision system		
	Author	Aníbal Concha-Meyer, Joseph Eifert, Hengjian Wang & Gabriel Sanglay		
3	**Methodology**	**Achievement**		**Summary**
	OMNISURFACE SOFTWARE Volume are determinded using the optical imaging system and compared to volume measurement collect using water displacement method	Machine vision volume measurement technique demonstrated high repeatability for each produce type		For these products, volume determinations based on a single value of density may not be practical since density can be variable and density may not be known or consistently reported.
	Title	2D/3D Vision-Based Mango's Feature Extraction and		
	Author	Thanarat Chalidabhongse, Panitnat Yimyam		
4	**Methodology**	**Achievement**		**Summary**
	2D/3D Vision-Based 1)Image acqustion 2)Camera calibration 3)3D Volume reconstruction 4)2D/3D Visual Feature Extraction	The results shows the average accuracy of the systems are 87.65%		The results show the technique could be a good alternative and more feasible method for sorting mango comparing to human's manual sorting.
	Title	Computer Vision Estimation of the Volume and Weight of Apples by Using 3D Reconstruction and Noncontact Measuring Methods		
	Author	Baohua Zhang, Ning Guo, Jichao Huang, Baoxing Gu, and Jun Zhou		
5	**Methodology**	**Achievement**		**Summary**
	1)**Near-Infrared Linear-Array Structured Light Vision System**(CCD Camera). 2)3D Reconstruction and Height Map Image Acquisition. 3)Feature Extraction and Selection. 4)Partial Least Squares(PLS) 5)Least Squares-Support Vector Machine (LS-SVM). 6)Evaluation of the Performance of the	The 3D upper surface of apples was reconstructed by using the proposed system and triangulation method as moving the fruit on a conveyor belt		The system and method developed in this study provide an alternative to the traditional methods for noncontract measurement of the volume and weight of agricultural products.

Based on the table above there are two things that can be learned which are volume and weight estimation are possible to be done. Second, there is still room for improvement in these studies.

CHAPTER 3

METHODOLOGY

3.1   Introduction

This chapter will explain the detailed method in order to achieve the objective of the project. As stated, in chapter 1, the main objective of this project is to develop a computer vision system that is able to estimate the volume and weight of an apple by using a camera. The methodology of this project can be represented with a block diagram which consists of three main parts which are input, systems and the output. Each part has its own function and objective that needs to be achieved.
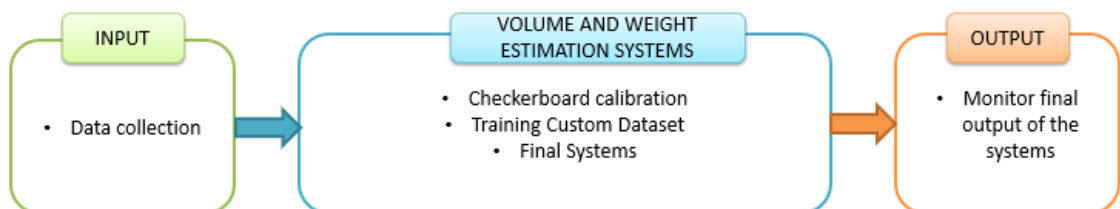
Figure 12: Block Diagram

3.2   Sample Used in Research

There are many cultivators of apples available in Malaysia. In this project, there are three models that will be used to estimate weight and volume which are Koru Apple, Royal Gala Apple and Fuji Apple. The selection of these three models is due to significant size which can be differentiated easily. All types of this apple can be found in the most local market. This project will use 6 Koru Apple, 10 Royal Gala Apple and 4 Fuji Apples. All of this apple will proceed through the same process in obtaining the actual parameters and the set of pictures that will be used on this project.

3.3   Software Used in Develop System

There is multiple software that will be used in order to accomplish this project. The uses of this software are varied and specific for each function and objective. There are three software that will be used in this project which are VGG Image Annotator (VIA), MATLAB and Google Collaboratory.

VGG Image Annotator is free software that is used to manually annotate the images. This software will be used to prepare a custom data set to train the detectron2 network.

MATLAB is a flexible software that can be useful in many situations. In this project, MATLAB is used in checkerboard calibration. This data is important in order to determine pixel values which are included in the formula to estimate the parameters of the apple.

In order to develop a system, one of the most important things that need to be considered is the choice of software. This system will use Python as the main language to develop the systems. Since this project is a computer visual project with multiple integrated development environments (IDE) that can be used such as PyCharm, Microsoft Visual Studio, OpenCV, Google Collaboratory and many more.

In this project Google Collaboration has been chosen as IDE to develop this system. There are a few reasons for the selection of Google Collaboratory compared to other IDEs. Microsoft Visual Studio, PyCharm and OpenCV are certainly more advanced IDEs than Google Collaboratory and more suitable to develop high level systems, however the usage of this IDE is much more complicated. On the other hand, Google Collaboratory is more simple and easier to use. Moreover, it also has a free GPU which is a great advantage to run a high-level program and able to reduce strain on a laptop. Other than that, it also has the capability to run code separately by column. This is very helpful in order to evaluate and try the programs. Furthermore, it also runs online which is able to keep all the saved data in the cloud. This feature is very helpful in order to ensure the file is missing.

3.4   Flowchart

The flowchart of these systems can be represented as in the figure. The flowchart of this system is straightforward since there are not any conditions that will change the program. Each colour of the flowchart represents each phase of the systems. There are five phases of the systems which represent the input, systems and the output of this project.
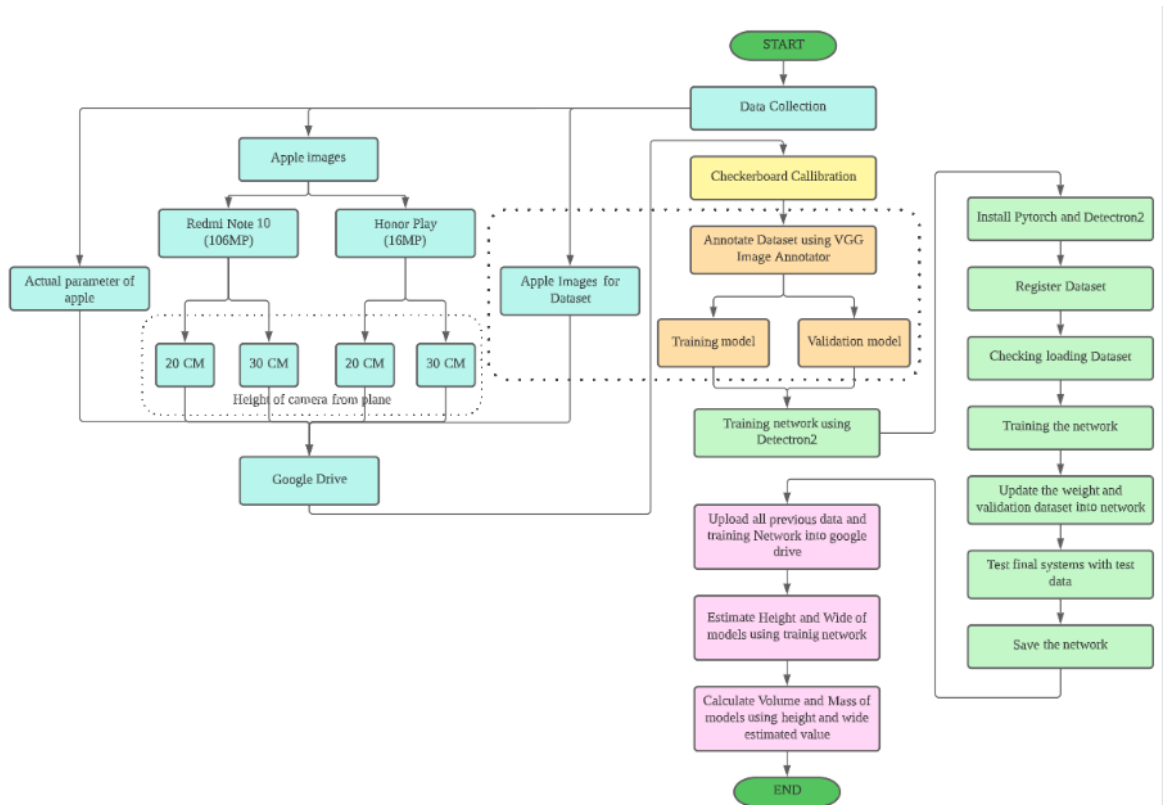
Figure 13: Flowchart of Systems

3.5   Data Collection

Data collection is one the most important parts of this project. A lot of data is required to accomplish this project. There are two types of data that need to be collected which are Measuring the actual parameters of apples and image of apples.

3.5.1   Measuring Actual Parameters of Apples

The main objective of this project is to estimate the weight and volume of apples. However, these parameters are influenced by two main parameters which are the width and height of apples. Through the formula by applying the width and height of apples,

volume can be estimated and the mass of apples also can be calculated. The actual parameters that will measure are width, height, volume and mass of the model.

In this project the actual volume of the apple is gained by using the Water Displacement Method (WDM) [25]. The procedure of WDM is as follows. Firstly, the initial volume of water is taken. After that the apple is submerged into the water using a sinker rod in our project. Then the final volume value after submerged is subtracted from the initial value resulting in the total volume of the apples. It should be mentioned that the sinker rod utilised in our study is extremely thin; however, because the volume of the apple is significantly more than that of the sinker rod, the sinker rod's volume can be regarded inconsequential and thus ignored in the volume measurement [17].

The actual weight of an apple can be measured by using digital balance. Digital balance is more accurate than manual balance. Therefore, this tool can be used in order to determine the actual weight of an apple.

Figure 14: Digital Balance

In order to measure the height and width of the apple, digital vernier calipers are used. Vernier calipers are able to provide precision to 0.0025 mm which is acceptable for this project. Moreover, by using digital vernier calipers, measurement data are more accurate and precise than regular.
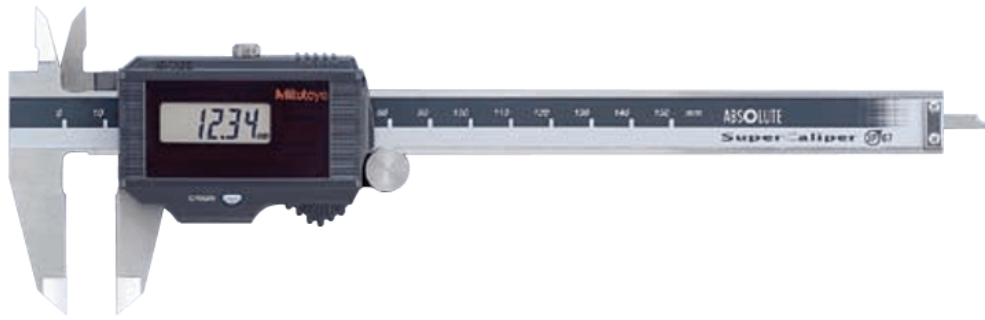
Figure 15: Digital Vernier Caliper

3.5.2 Image of the apple

Image of the apple served two functions in this project. The first function is to create a custom data for the training network that will be used. This set of images are taken from the internet. The selection of the images is based on the cultivator of apple that will be used in this project. A total of 67 images are used in this phase. Which are 51 images of apples being used to create training dataset and 16 images are used to create validation dataset.

The second function is to test the systems. This image is a collection of apples that have taken their actual parameter. The images of these apples will be used to obtain the value that will be produced from the systems. There are several variables for the images such as types of cameras and height of camera from the surface. In this project there are two types of cameras that will be used which are Honor Play (16MP) and Redmi Note 10 (108MP). Each of this camera will take the images from two different heights which are 20 cm above from ground plane and 30 cm above from the ground plane. All this image will be saved in a zip file according to each category and will be uploaded into google drive.

3.6 Checkerboard Calibration

Camera calibration is one of the keys in this project since it aims to figure out the picture generation process of geometric parameters. This is due the subject of this project is an image that is taken from the camera which gives the metric information of the scene. There are many methods in calibrating cameras, however in this project checkerboard

calibration via intersection points are used. This method is chosen due to its capability to obtain accurate data and the simplicity of the algorithms. The checkerboard that will be used in this project is sized 23mm squares with 9x7 squares. MATLAB will be used to calibrate the checkerboard by detecting intersection points. The output of the intersection points will produce coordinates for each point. By subtracting each coordinate, the length and width of each box can be found. The average value of the wide and length can be represented as X-coordinate value and Y-coordinate value. Since the real length of the box is known, the ratio of pixel/cm can be found.
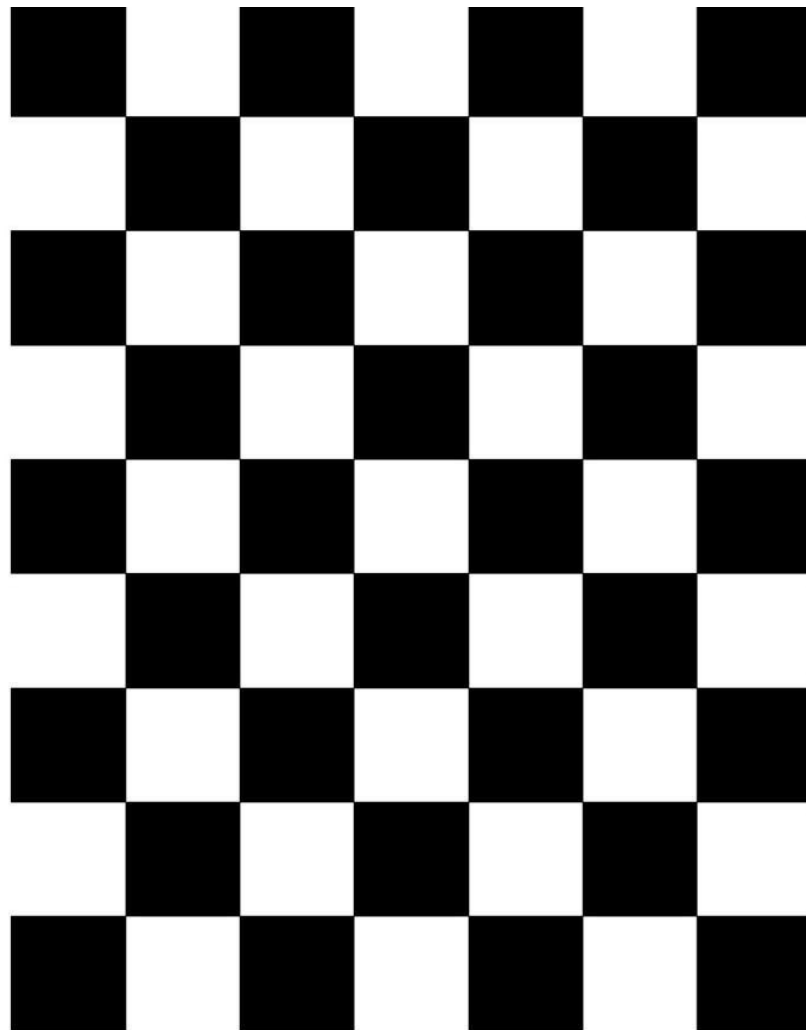
Figure 16: 23mm squares with 9x7 checkerboards

3.7   Annotate Dataset

Algorithms have guided this world without many people noticing this action. Many decisions are decided by algorithms such as GPS, playlist, news feed and many more. The ability of these algorithms to make correct decisions depends on data

annotation. The same with this project, data annotation is very important for training the proposed network to recognize and segment the selected models. Even with current technology, computers are still unable to analyze visual information the way the human brain does. Therefore, human help is needed to analyze visual data to show the computer how to interpret the data. There are many types of data annotation such as semantic annotation, image annotation, video annotation and many more. However, only image annotation is used in this project. Currently there are already library that has collected annotated image dataset which such as Microsoft COCO Dataset (Common Object in Context) which has million labelled instances and hundreds of thousands of images and Google OID (Open Images Database) with million pre-annotated images. However, in this project both these databases are being used since in this project only one model is needed to be recognized by the computer. Therefore, it is important to create this custom database for this project.

Image annotation is utilised to indicate the aspects the system needs to recognise. There are a lot of software and websites that can be used to annotate data. In this project, the annotation is done by using VGG Image Annotator. As mentioned before, VGG Image Annotator is a free software that is used to manually annotate the images.

In this project there are two types of data set that need to be annotated which are the training dataset which consists of 51 images of apples and validation dataset which consist of 16 images of apples. Usually, a huge number of images are needed for this project, however since this task is very time consuming, a minimum number is only used. First and foremost, before labelling the picture, it is important to set region attributes. This acts as a configuration file for the image. Next select polygon region shapes and start labelling. These two steps are repeated to all images in this project. After finishing it, it is crucial to save both files as a json file. By finishing all these steps, the training dataset are ready to be used for the next step. Lastly both files need to be uploaded to google drive for the next use.
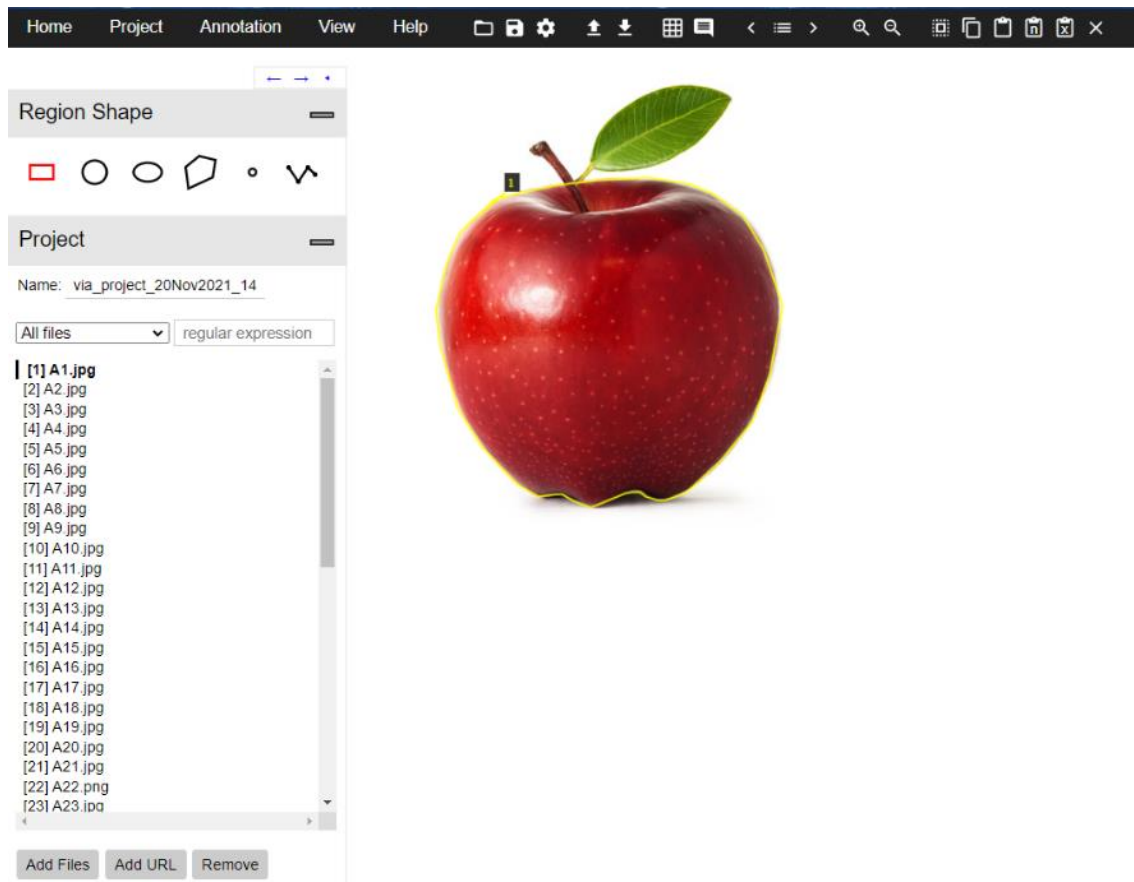
Figure 17: VGG Image Annotator Output

3.8   Detection and Segmentation

In this project the most important component to estimate volume and weight is the capability of systems to detect and segment the Apple models. There are several methods to perform this function and some of it has been provided. One of the methods is to use Objectron. Objectron is one of the machine learning solutions provided in MediaPipe. MediaPipe is a cross-platform framework for building multimodal applied machine learning pipelines. However, Objectron are not used in this system due to several reasons.

The other method is by using Detectron2. Detectron2 is the next-generation library from Facebook AI Research, and it includes cutting-edge detection and segmentation algorithms. In this part, the systems require to recognize the location, mask overlapping and indicating models (apple) exact border. This function is called instance segmentation.

In order to train the desired network, the most important thing that needs to have been the dataset. It is possible to use the provided dataset from Microsoft or google. However, this dataset has multiple types of classes which defeat the main purpose of these systems. The first objective of this project is capability to detect and segment apples only. Therefore, a custom dataset is needed which has been done before this part. The training network of these systems are done using Google Collaboratory and Detectron2. The process of training this network can be simplified with the flow diagram below. As shown in figure, the network is divided into the following steps.
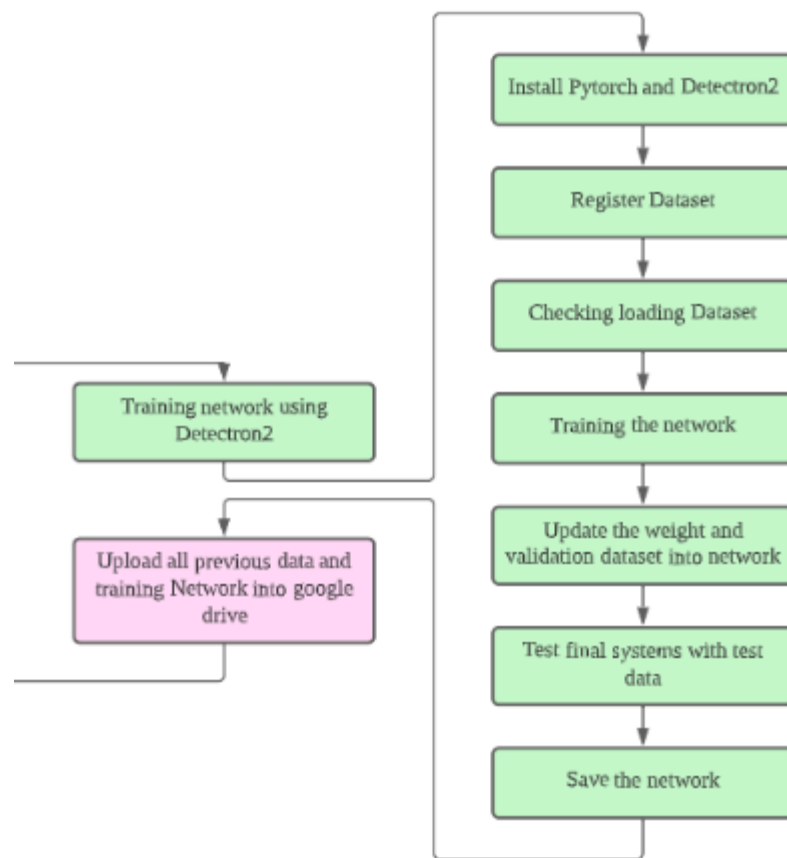
Figure 18: Flowchart of Training Network using Detectron2 with Custom Dataset

As can be shown in the flowchart the process of training the network is started by installing all necessary libraries that will be used such as pytorch and detectron2. The next step is to register the dataset. In this part, all necessary libraries and data that have been prepared before are uploaded here. It is important to check the data that have been uploaded before in order to train the network, therefore before proceeding to the next part, it is must to check the loading dataset that have been uploaded are correct by

displaying some of the data. The samples are loaded randomly from the data that have been uploaded before. Training the network can be done after all data has been uploaded correctly, since the total images of the training dataset are kept to minimum, the time in to proceed this the training are less consuming. After that, the output of the training dataset can be updated into the network. By completing the previous step, the training network is already done. Test the network with the new data to see how it performs. Last but not least, all the output configuration will be saved in google drive for the final system that will be used.

3.9   Estimate Volume and Weight of Apple

In this part the main objective of this project will be to estimate the volume and weight of apples. By combining all the data and the configuration that have been done before, the final systems can be developed. The first and foremost before estimating the volume and weight of apples, is a must to find the width and height of models since the formula to determine the volume depends on the width and height of the models.

$$Volume = \frac{\pi(Height \times Wide \times Wide)}{6}$$ 3.1

By installing all necessary libraries and configuration output before, the systems are able to recognize the apple model and predict the area of the model. From the mask area, the systems will create a boundary box around the mask area. The width and height of the models can be found from the coordinate of the boundary box. The width can be calculated from the difference of the X-coordinate while the height can be calculated from the Y-coordinate. However, the width and length shown by the systems is in pixel value. By using the calibrated value that has been done before, the real value for width and height can be found. By applying the width and height and value in the formula, the volume of models can be found. In order to estimate the weight of models, the volume of the models needs to be multiplied with the density of the models that has been set which is 0.843 g/cm3. The density of apple is obtained by analysis data of apple that has be done in previous study.

$$Mass = Volume \ \times density$$ 3.2

Lastly the comparison between measured data using instruments and calculated data using the systems are compared to find the capability of the systems by determining the error of calculated value between the measured value.

3.10   Summary

To conclude based on the project proposal, this project is able to be done by completing all the steps above. This project consists of multiple stages of programs that relate to each other. The output for each stage will affect the final output for this system. One of the most important parts of this project is collecting the dataset. As explained before there are three types of datasets that were collected for this project. Each dataset will be used for multiple purposes such as to create training and validation dataset for detection and segmentation of apple model. Other than that, camera calibration using point detection of the checkerboard also affects the final result since the ratio value that obtains in this part will be used as conversion value. Therefore, it is very important to proceed each part with utmost cautions in order to obtain the best result.

CHAPTER 4

RESULTS AND DISCUSSION

4.1   Introduction

In this chapter, all the results of this project will be presented and discussion on the obtained result will be shown. There are multiple results that have been obtained since this project has multiple stages and levels. First and foremost, preliminary results that have been obtained before will be discussed. At that moment many improvements and exceptions are being done to satisfy the objective of this project. Other than that, the current result obtained will also discuss such as the camera calibration, the output of the network training and the final result of the system.

4.2   Preliminary Results

Detectron2 has been used as a tool in order to perform detection and segmentation of the models as mentioned before. From a pre-trained model, it can be seen that this method is able to accurately detect and segment the models perfectly. The images below show the result of detection and segmentation of models by using Google Collaboratory, Detectron2 and COCO Dataset.

Figure 19: Output 1 based on Detectron2 and COCO Dataset

However, this result is not good enough since the current libraries are using the provided libraries which have multiple classes. This has made the systems detecting other classes in the images that were provided. It can be seen in this result, even though it perfectly masks the area it is still unable to recognize the true models.
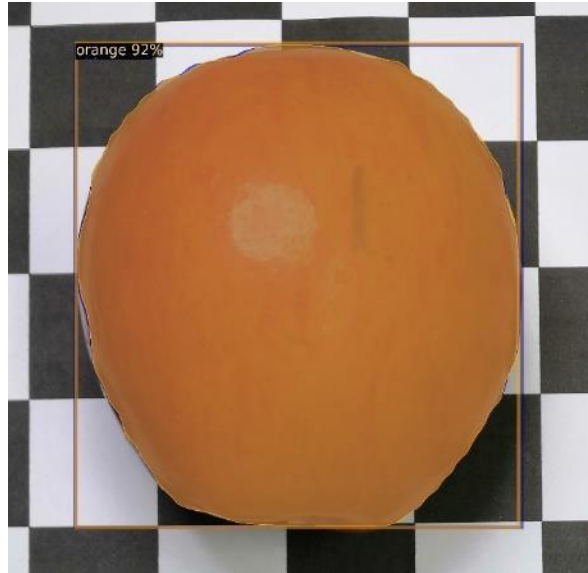
Figure 20: Output 2 based on Detectron2 and COCO Dataset

From this result the systems were able to predict two models in the picture which are orange (class-49) and apple (class-47). Furthermore, the systems also show high confidence values on predicting the images as oranges with 91.71% and apples with 62.98%.

```
tensor([49, 47], device='cuda:0')
Boxes(tensor([[ 609.6826,  567.1762, 1960.6813, 2041.8445],
        [ 608.1311,  572.4471, 1969.5592, 2041.8251]], device='cuda:0'))
tensor([0.9171, 0.6298], device='cuda:0')
```

Figure 21: Configuration data on Output 2

Since in this project the focus model is apple, therefore it is a must to remove other classes in developed systems. There are two methods to remove other classes from libraries which are to filter the classes accordingly or create a new specific dataset from scratch. Even though filtering techniques seem to have simple processes, the procedure to achieve it is much harder than creating a new dataset. Therefore, creating a new dataset has been a choice to improve these systems.

4.3 Checkerboard Calibration

Camera calibration is one of the most important keys in this project which functions as a reference point and parameter. The main objective in checkerboard calibration is to find the pixel/cm ratio. This ratio is important since the final output that these systems will give in the pixel value. Therefore, this ratio is needed to change the output of the pixel value that has been given by the systems to convert in cm.

The checkerboard that was used in these systems is sized 23mm squares with 9x7 squares. The process of this stage is to find the intersection point in the checkerboard. The output from the MATLAB shows the checkerboard that has already been marked as point as the figure below.
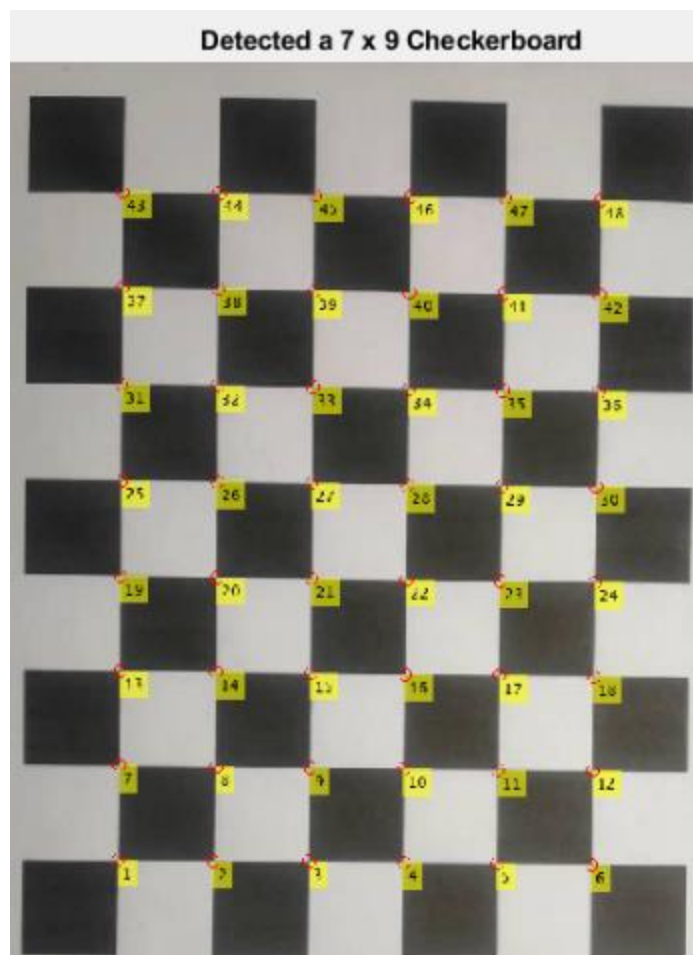
Figure 22: Output of Checkerboard Calibration

The checkerboard in this project is able to mark 48 intersection points which each point has designated coordinates in (X, Y) value. This coordinate will be used to find the size of boxes in the images. The example of coordinate is shown in the figure below.

Table 4.1: Intersection Point Coordinate

no	X-coordinate	Y-coordinate
1	888.406	2087.265
2	1116.391	2089.961
3	1342.872	2092.658
4	1570.090	2095.091
5	1797.451	2098.106
6	2027.150	2100.937
⋮	⋮	⋮
43	909.396	513.889
44	1134.699	514.175
45	1360.414	513.727
46	1586.862	513.855
47	1813.950	514.993
48	2040.995	518.609

The first column indicates the X-coordinate value while the second column indicates the Y-coordinate value. The first row indicates the first intersection points. By subtracting the coordinate of the first points with the second points, the length of the box can be found in pixel value. Since the actual length of the box is known. The ratio of pixel/cm can be known. Usually, calibration of the camera is done only once. However, since this project has few variables which are types of cameras and height of camera from the plane are also included. The calibration needs to be done four times. The result of the calibration can be shown in the table below.

Table 4.2: Ratio of Pixel/cm

Ratio of Pixel/cm				
	20 cm		30 cm	
	X	Y	X	Y
Honor Play (16MP)	169.585	167.701	116.145	116.182
Redmi Note 10 (108MP)	146.955	145.936	98.765	98.07

The table above shows that each setup has a different value of ratio pixel/cm. This difference has been caused due to a lot of reasons such as the resolution of the camera, light around the stop and many more. This value may affect the final results of the systems

when exchanging the pixel value of the final result during estimation length and width of apples. In order to solve this problem fine tune can be done to obtain a much more accurate value of ratio pixel/cm.

4.4   Detection and Segmentron of models

In this part, which make the project is possible which it objective to detect the apple and segment the area of apple accurately. By able detect and segment the apple accurately the next phase of this project will only be possible. Previously, detection using Detectron2 and COCO Dataset has resulting to multiple detection of object. This is due to the excessive number of categories that available in the COCO Dataset. Therefore, a custom dataset needs to be created in order to detect only designated models which is apple. The figure bellows shows the output of the newly designated systems with custom dataset.
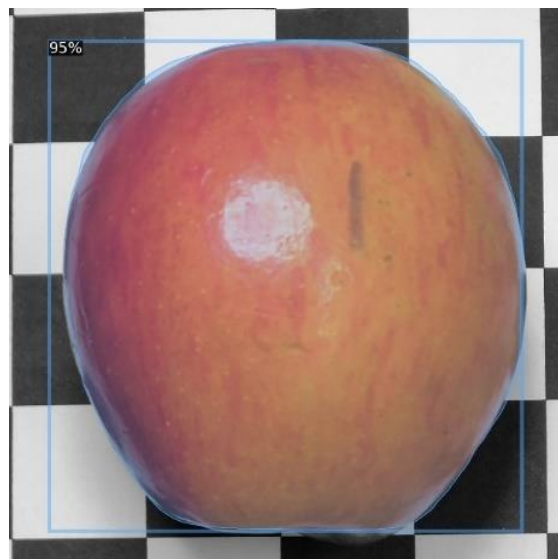
Figure 23: Output based on Detectron2 and Custom Dataset

The output above shows clearly a difference between the previous output that I obtained before. This output only detects one designated object which is apple. Since this is the only class in the dataset that is used, the classes of the apple are set to "0".

```
tensor([0], device='cuda:0')
Boxes(tensor([[ 575.1954,  587.4514, 1956.8206, 2017.2299]], device='cuda:0'))
tensor([0.9528], device='cuda:0')
```

Figure 24: Configuration Data of Output

The configuration data above clearly shows that only one class is detected and only one boundary box is generated around the models. Moreover, the confidence level of apple detection is 95.28% which can be its height rate of detection. This proves that the system is clearly successful in detecting and segmenting apples only.

4.5 Estimate the volume and weight of an apple

In the final stage of this project, the apple picture will be used to analyze in order to obtain the parameters which are the height and width of the apple in the picture. By utilizing both of these parameter volumes of apples can be calculated. The formula has been proved with the apple that will be used in this project. The graph below shows the difference between the actual volume of apples and the calculated volume of apples.
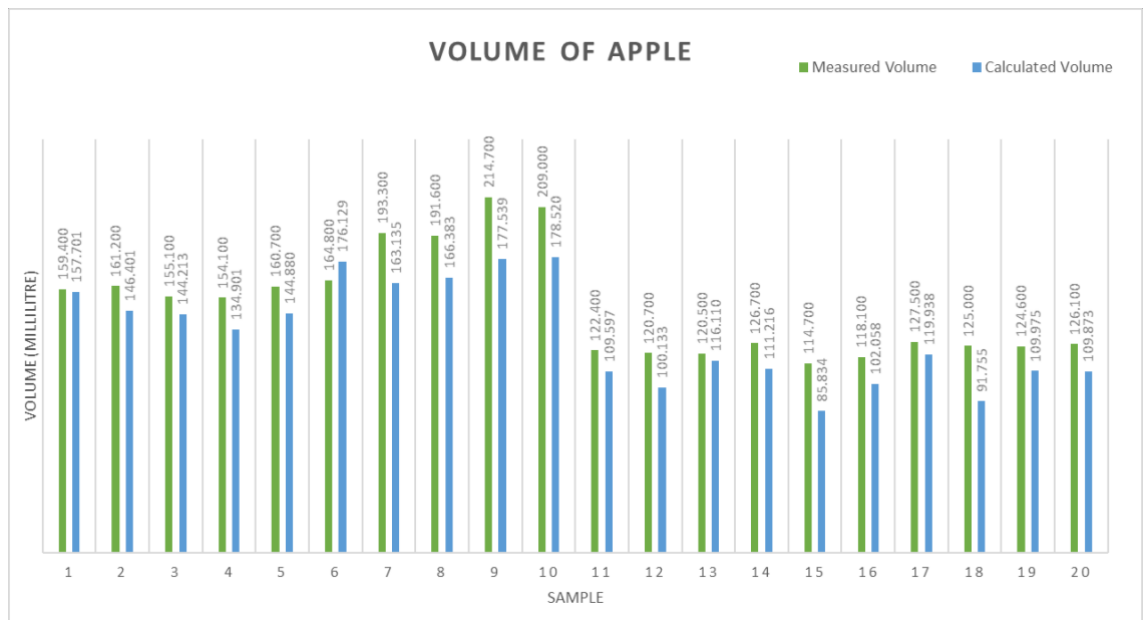
Figure 25: Different volume of apple between theorical value and calculated value

The result shows that the accuracy using the formula in determining apple volume of apples based on width and height is 85.41%. Therefore, by using the formula it has been shown that the volume of apples can be estimated. The mass of the apple is estimated

by using the volume that was obtained previously and the density of the apple. Since the density of each object is constant, the mass of the apple can be calculated. The graph below shows the difference between the actual mass of the apple and the calculated mass of the apple, which has an accuracy of 85.26%.
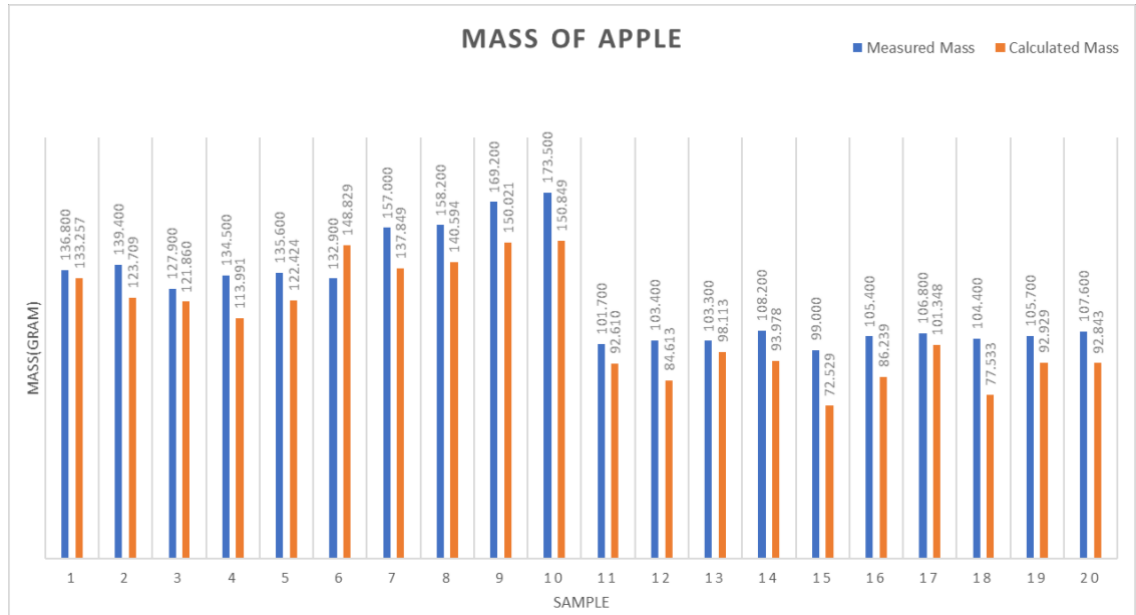
Figure 26: Difference mass of apple between theorical value and calculated value

As all the preparations have been done the final systems can be used to estimate height and width of the apple and calculate the volume and mass. In this part, there are multiple outputs that will be shown such as the detection and the segmentation of the apple in images, the coordinate of the boundary boxes, height, wide volume and mass of the apple. The figure below shows the output of the apple detection and segmentation of the apple.
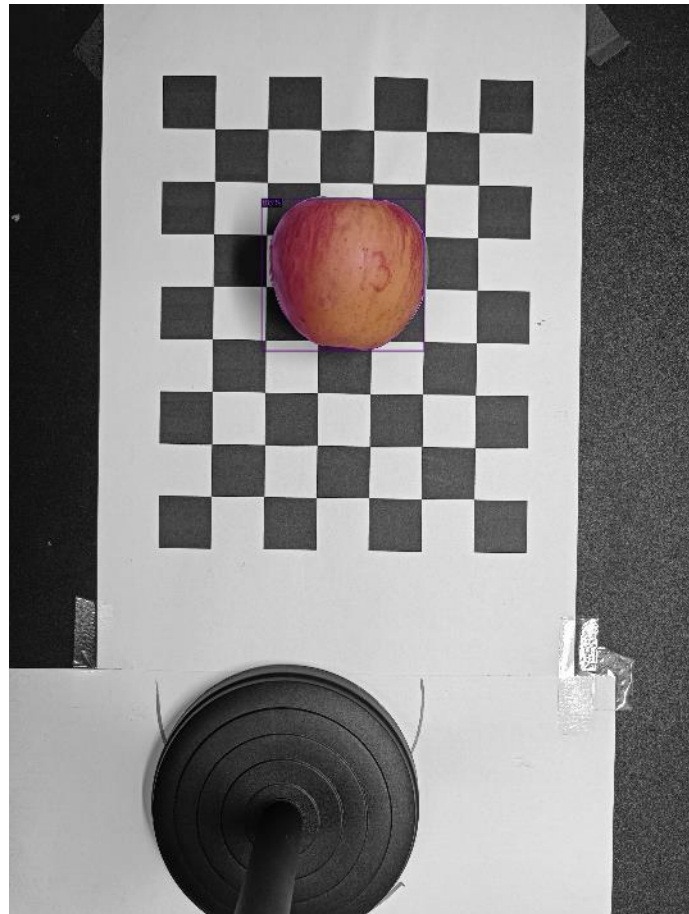
Figure 27: Detection and Segmentation of apple

In the figure above only the mask area is displayed. The parameter in the figure above is displayed separately. The figure below shows the parameter that was obtained from the output previously.


```
Area:375562.5 pixel
Coordinate of X1:1104.767578125
Coordinate of Y1:854.7742919921875
Coordinate of X2:1811.5660400390625
Coordinate of Y2:1520.3739013671875
Height:6.787 cm
Wide  :7.156 cm
Volume:172.616 ml
Mass  :145.515 g
```

Figure 28: Parameter of the Output

By displaying both outputs, the final objective in this project is achieved. This project is done with multiple setups to obtain data for analysis comparison. The total result of this project can be shown in the graph below.
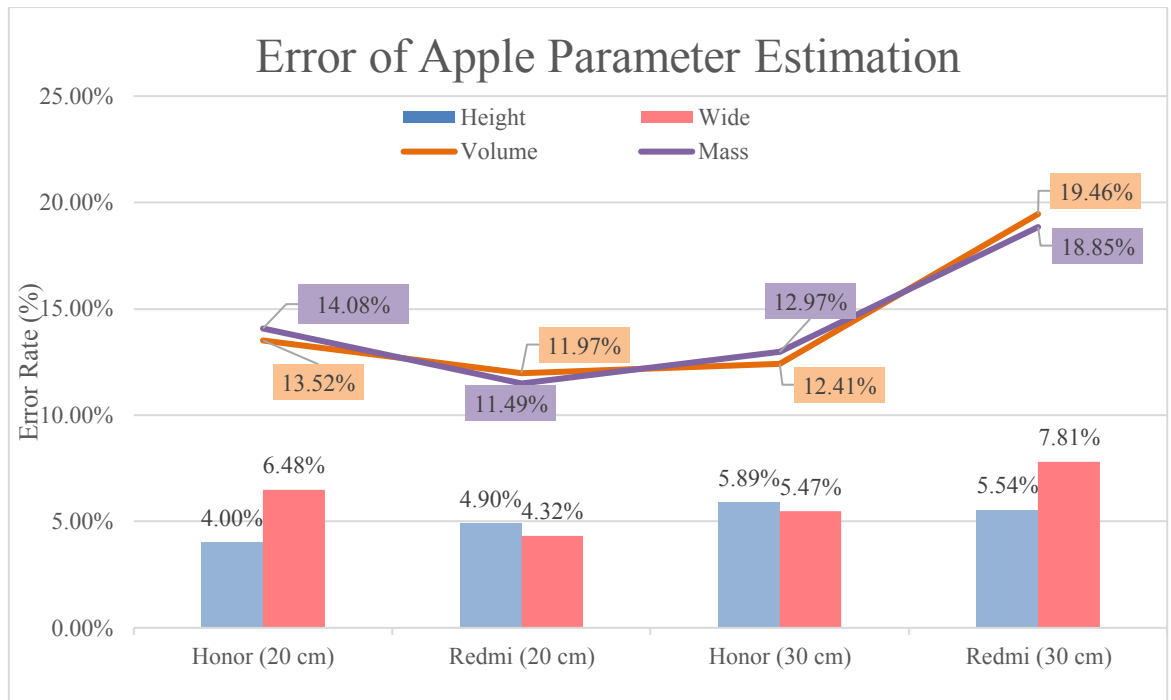
Figure 29: Accuracy Estimation Height and Wide of Apple

In this project the volume and mass result are depending on the width and height of the estimation result. Based on the graph above it can be seen that most of the results have an error of less than 15% except for Volume and mass in the Redmi setup (30 cm). However, the accuracy of All the systems is more than 80% which can be said to be a good result. The error rate that obtains less for height and wide estimation are 4.00% and 4.32%. Meanwhile the highest error rate for height and wide are 5.89% and 7.81%. The average error rate for height and wide are 5.08% and 6.02%. Even though the error rate for height and width is less than 10%. The result is affected for the volume and mass estimation. Therefore, from the result it can be said that there are many factors that will affect the systems. It can be seen that different types of phones and height are able to manipulate the final result.

4.6   Discussion

There are multiple reasons for the error that I obtained from this project. The first reason is the condition of the setup. From the result obtained, it can be seen that image resolution greatly influences the result in this project. In this project, lighting may be one of the reasons that cause the image to be unclear. The lights that were used are small ring lights which are not enough to surround the model. There are also multiple types of lighting setups that can be taken into consideration in order to obtain a much better image.

51

The second reason is due to checkerboard calibration. In this project, checkerboard calibration is done by finding out the ratio/cm with a different value of the intersection point. Theoretical value of pixel/cm is supposedly 37.795 pixel/cm. However, in this project the ratio that is obtained is 98.765 pixel/cm. Even though this value is directly proportional with it, surrounding the outcome still affects the final result. Moreover, the ratio also can be seen as not accurate within the box itself. Where the ratio for X-value and the ratio for Y-value are not the same.

Lastly, lack of data. In this project only 51 images of apples are used to train the system and only 16 images of apples are used for the validation. This lack of data may cause inaccurate detection and segmentation of apples. It can be seen in this result that there are some results that are unable to segment the apple wholly. In other view, there are also certain images that are not created boundary boxes that accurately cover the whole region.

There may be some other factor that caused the inaccuracy of this system which may not go unnoticed. However, all three reasons mentioned above are the main reasons that cause the accuracy of data and need to be overcome.

4.7   Summary

The result above shows that this project is able to achieve all its objectives that have been proposed. The proposed method by creating a custom apple dataset to detect and segment apple models using detectron2 has been proved successful. Other than that, the systems are also able to obtain all the parameters that have been seeked for calculation purposes. Moreover, the result is able to estimate all the parameters with accuracy around 80% based on different setup. All possible errors mentioned above may contribute to the accuracy estimation result. Therefore, it can be sure that there are still rooms of improvement for this project in order to apply this system in industries.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1   Conclusion

To conclude, computer vision is one of the most advanced technologies that is currently available. Exploiting the capabilities of this technology in other sectors can improve in a lot of ways. Gauging is one out of the four main functions of computer vision. In this world there are many things that require precise and accurate measurement. This parameter can affect many things such as market value, customer preference and many more. In the agriculture industry many parameters such as volume and weight are sorted using human judgement which the outcome can be argued. Therefore, this project is proposed for a solution in industry to estimate height, width, volume and mass of apples much faster and more accurately.

The first objective for this project is to detect and segment the apple model. This objective is able to be achieved by creating a custom apple dataset. The dataset is created using VGG Image Annotator while the detection and the segmentation of apple models are accomplished by using Detectron2 library. The result shows that only one model is detected when using the custom dataset as the library.

The second objective is also able to be obtained. At the end of the process a vision-based system that is able to estimate volume and weight of apples is able to be developed. The data which is obtained from the Detectron2 are used to obtain all the parameters of apple from the images. The width and height of the apple are obtained from the boundary box of the apple and are used to calculate the volume and weight of the apple. Based on the result all of the apples are able to estimate the volume and weight of the apple using these systems.

The final objective of this object is to compare the accuracy of the system with the measurement using the instrument. The systems are tested with 20 Apples by using four different setups to estimate all parameters of apple which are wide, height, volume

and mass. The final result is able to estimate all the parameters with error rate less than 20%. The highest accuracy is obtained by using Redmi (20cm) setup with 95.10%, 95.68%, 88.035 and 88.51% for height, width, volume and mass.

5.2   Future Recommendation

In this project, it has been proved that estimation of volume and weight of apples using phone cameras can be done. However, there are still rooms for improvement in these systems in order to apply in industry. The current systems still required human help especially in taking the visual data for analysis. Based on this project one of the most important things that need to have been the orientation of the apple during measurement. The orientation of the apple which enables the camera to obtain height and width of the apple are important for more accurate data. Moreover, the time taken for each apple model is also time consuming. Therefore, a much better camera needs to be used in order to improve the speed of acquiring visual data of apple. This can be done by installing a high-speed camera which takes multiple images of the same apple for data analysis.

Camera calibration is also one of the main concerns in this project. The ratio of pixel/cm is used to convert the visual data into actual data which can be used and understood by many people. The current method is only capable of single data in a fixed position. Moreover, by using checkerboard calibration, many other parameters need to be considered such as the height of the camera from the plane, lighting around the model and many more. Therefore, new methods for camera calibration need to be included in order to improve the flexibility of the current systems.

Lastly, the data for training and validation for apple detection and segmentation. The result that obtains from this part of this project shows that the systems are still unable to accurately segment the apple models. This may be due to lack of training data. The current systems only use 51 images of apples for training and 16 images of apples for validation. The current data is still insufficient for these systems. Therefore, much more data is needed to create these systems which are able to accurately segmented apple models.

REFERENCES

[1]   "What is Computer Vision? | IBM," www.ibm.com. https://www.ibm.com/my-en/topics/computer-vision.

[2]   V. Meel, "56 Most Popular Computer Vision Applications in 2021," viso.ai, Mar. 09, 2021. https://viso.ai/applications/computer-vision-applications/.

[3]   A. Mala, "The Most Popular Fruit in the World," WorldAtlas, Nov. 15, 2017. https://www.worldatlas.com/articles/the-most-popular-fruit-in-the-world.html.

[4]   F. Bridge, "3D Sensing - New Ways of Sensing the Environment," FutureBridge, Jul. 06, 2019. https://www.futurebridge.com/blog/3d-sensing-new-ways-of-sensing-the-environment/.

[5]   S. Vivek, "What Is the Difference Between Mass, Weight and Volume?," Sciencing, Dec. 17, 2021. https://sciencing.com/difference-between-mass-weight-volume-8622796.html (accessed Feb. 18, 2021).

[6]   D. Halliday, Fundamentals Of Physics. WILEY, 2020.

[7]   R. C. Jain and E. Al, Introduction to machine vision. Maidenhead: Mcgraw-Hill, 1995.

[8]   B. Zhang et al., "Computer vision detection of defective apples using automatic lightness correction and weighted RVM classifier," Journal of Food Engineering, vol. 146, pp. 143–151, Feb. 2015, doi: 10.1016/j.jfoodeng.2014.08.024.

[9]   Z. Xiao-bo, Z. Jie-wen, L. Yanxiao, and M. Holmes, "In-line detection of apple defects using three color cameras system," Computers and Electronics in Agriculture, vol. 70, no. 1, pp. 129–134, Jan. 2010, doi: 10.1016/j.compag.2009.09.014.

[10]   J. Blasco, N. Aleixos, and E. Moltó, "Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm," Journal of Food Engineering, vol. 81, no. 3, pp. 535–543, Aug. 2007, doi: 10.1016/j.jfoodeng.2006.12.007.

[11]   S. Cubero, M. P. Diago, J. Blasco, J. Tardáguila, B. Millán, and N. Aleixos, "A new method for pedicel/peduncle detection and size assessment of grapevine berries and other fruits by image analysis," Biosystems Engineering, vol. 117, pp. 62–72, Jan. 2014, doi: 10.1016/j.biosystemseng.2013.06.007.

[12]   M. Hu, Q. Dong, B. Liu, and P. K. Malakar, "The Potential of Double K-Means Clustering for Banana Image Segmentation," Journal of Food Process Engineering, vol. 37, no. 1, pp. 10–18, Nov. 2013, doi: 10.1111/jfpe.12054.

[13]   Y. Zhang and L. Wu, "Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine," Sensors, vol. 12, no. 9, pp. 12489–12505, Sep. 2012, doi: 10.3390/s120912489.

[14]   W. Wang and C. Li, "Size estimation of sweet onions using consumer-grade RGB-depth sensor," Journal of Food Engineering, vol. 142, pp. 153–162, Dec. 2014, doi: 10.1016/j.jfoodeng.2014.06.019.

[15]   A. Vidal, P. Talens, J. M. Prats-Montalbán, S. Cubero, F. Albert, and J. Blasco, "In-Line Estimation of the Standard Colour Index of Citrus Fruits Using a Computer Vision System Developed For a Mobile Platform," Food and Bioprocess Technology, vol. 6, no. 12, pp. 3412–3419, Dec. 2012, doi: 10.1007/s11947-012-1015-2.

[16]   C. Garrido-Novell, D. Pérez-Marin, J. M. Amigo, J. Fernández-Novales, J. E. Guerrero, and A. Garrido-Varo, "Grading and color evolution of apples using RGB and hyperspectral imaging vision cameras," Journal of Food Engineering, vol. 113, no. 2, pp. 281–288, Nov. 2012, doi: 10.1016/j.jfoodeng.2012.05.038.

[17]   A. B. Koc, "Determination of watermelon volume using ellipsoid approximation and image processing," Postharvest Biology and Technology, vol. 45, no. 3, pp. 366–371, Sep. 2007, doi: 10.1016/j.postharvbio.2007.03.010.

[18]   C. C. Teoh and A. R. M. Syaifudin, "Image processing and analysis techniques for estimating weight of Chokanan mangoes (Teknik pemprosesan imej dan analisis imej untuk menganggarkan berat mangga Chokanan)," 2007.

[19]   A. Concha-Meyer, J. Eifert, H. Wang, and G. Sanglay, "Volume estimation of strawberries, mushrooms, and tomatoes with a machine vision system," International Journal of Food Properties, vol. 21, no. 1, pp. 1867–1874, Jan. 2018, doi: 10.1080/10942912.2018.1508156.

[20]   T. Chalidabhongse, P. Yimyam, and P. Sirisomboon, "2D/3D vision-based mango's feature extraction and sorting," 2006. doi: 10.1109/ICARCV.2006.345248.

[21]   M. Omid, M. Khojastehnazhand, and A. Tabatabaeefar, "Estimating volume and mass of citrus fruits by image processing technique," Journal of Food Engineering, vol. 100, no. 2, pp. 315–321, Sep. 2010, doi: 10.1016/j.jfoodeng.2010.04.015.

[22]   B. Zhang, N. Guo, J. Huang, B. Gu, and J. Zhou, "Computer Vision Estimation of the Volume and Weight of Apples by Using 3D Reconstruction and Noncontact Measuring Methods," Journal of Sensors, vol. 2020, p. e5053407, Nov. 2020, doi: 10.1155/2020/5053407.

[23]   "Beginner's guide to depth (Updated)," Intel® RealSenseTM Depth and Tracking Cameras, Jul. 16, 2019. https://www.intelrealsense.com/beginners-guide-to-depth/.

[24]   "3 Common Types of 3D Sensors: Stereo, Structured Light, and ToF | Hacker Noon," hackernoon.com. https://hackernoon.com/3-common-types-of-3d-sensors-stereo-structured-light-and-tof-194033f0.

[25]   U. W. Robert, S. E. Etuk, and O. E. Agbasi, "Modified Water Displacement Method and its Use for Determination of Bulk Density of Porous Materials," Journal of Renewable Energy & Mechanics (REM), vol. 2, no. 01, Feb. 2019, doi: 10.25299/rem.2019.vol1(01).2292.

Task	WEEK — FYP 1														WEEK — FYP 2													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Meeting with a subject coordinator -PSM briefing session	X																											
PSM Titles realaease	X																											
ApprovL FYP title	X																											
PSM Register Title	X																											
Meeting with supervisor		X	X		X		X			X	X				X	X	X	X	X	X	X	X	X	X	X	X	X	X
Specification on the aim of the project's title		X	X																									
Discussion on software application for simulation			X	X																								
Identify the Objective and Scope of Study			X	X																								
Study of Problems				X	X																							
Literature Review			X	X	X	X	X	X																				
Initiating research			X	X	X	X																						
Research progress							X	X	X	X	X																	
Progress Presentation 1													X															
Report Preparation and Presentation 1								X	X	X	X	X	X															
Submission to Supervisor 1														X														
Upload to Coordinator 1														X														
Report and logbook submission 1														X														
Determine algorithms for volume and weight estimation															X	X												
Justify and testing algorithms																X	X											
Reconstruct Dectectron 2 with custom data																		X	X									
Obtain custom data using VGG Image Annotator																				X								
Calibrate checkerboard																				X	X							
Obtain data from camera																						X	X					
Compile and test the systems																			X	X	X							
Fixing error on the systems																						X	X					
Performance Analysis																								X				
Progress Presentation 2																												X
Report Preparation and Presentation 2																								X	X	X	X	
Submission to Supervisor 2																												X
Upload to Coordinator 2																												X
Report and logbook submission 2																												X