

King Saud University
College of Computer and Information Sciences
Computer Engineering Department



A VISION SYSTEM FOR DATE HARVESTING ROBOT

by

Hamdi Taher Altaheri

Advisors

Prof. Mansour Alsulaiman

Prof. Ghulam Muhammad

In partial fulfillment of the requirements for the Master degree in computer engineering, at the College of Computer and Information Sciences, King Saud University

Shaban/1440 H

April/2019 G

THESIS APPROVAL

A Vision System for Date Harvesting Robot


By

Hamdi Taher Altaheri

Supervisor: Prof. Mansour Alsulaiman

Co-supervisor: Prof. Ghulam Muhammad

This thesis has been approved on Monday Shaban 24, 1440 Hijri corresponding to April 29, 2019 G by the examination committee:

Committee member	Signature
Prof. Mansour Alsulaiman (supervisor)	
Prof. Ghulam Muhammad (co-supervisor)	
Dr. Haikel S. Hichri	
Dr. Yakoub Bazi	
Dr. Abdul Wadood Abdul Waheed	

The outputs of this research, including source codes, datasets, models, and real time testing videos, are publically available in the following website:

<https://sites.google.com/view/daterobotic>

©Approval Date
April 29, 2019

Hamdi Altaheri

All Rights Reserved

hamdi.altahery@gmail.com

ABSTRACT

In date cultivation, manual harvesting is the dominant method used, which is inefficient in terms of both time and the economy. Advanced agricultural automation such as robotic harvesting can significantly increase quality and yield as well as reduce production costs and delay. One of the most important aspects of harvesting robots is their ability to interpret and analyze visual data. Accurate vision system to detect, classify, and analyze fruits in real time is critical for the harvesting robot to be cost-effective and efficient. However, practical success in this area remains limited due to the difficulties caused by unstructured and unconstrained agricultural environments. Furthermore, research on machine vision for date fruits in the pre-harvesting and the harvesting stages is scarce. Hence, this research aims to develop intelligent systems for date fruit harvesting robotic in an orchard environment, including date fruit detection and segmentation, variety classification, maturity analysis, and automated harvesting decision system. We propose an efficient deep learning based machine vision framework for date fruit harvesting robots that consists of three classification systems used to classify date fruit images in real time according to their variety, maturity, and harvesting decision. Deep convolutional neural networks (CNNs) are utilized with transfer learning and fine-tuning techniques. In the case of date fruit segmentation, we propose a robust method to detect date fruits and segment them from the background (trunk, stalks, ground, sky, etc.). The method utilizes a superpixel clustering and local binary patterns with support vector machines. This research also creates a comprehensive dataset for date fruits that can be used for multiple tasks including automated harvesting, visual yield estimation, and classification tasks. The dataset has been fully labeled, coded, and released with their associated files to the research community in the IEEE DataPort repository [1] (<http://dx.doi.org/10.21227/x46j-sk98>). The resources of this research, including source codes, datasets, models, and test videos, are available for the benefits of researchers on the website <https://sites.google.com/view/daterobotic>.

KEYWORDS: robotic harvesting; machine vision; date fruit; fruit detection and segmentation; fruit classification; maturity analysis; deep learning, convolutional neural networks, superpixel clustering, local binary patterns, support vector machines.

ACKNOWLEDGMENTS

All praise is due to Allah, Lord of the worlds, The Entirely Merciful, the Especially Merciful, who guided me through all my life and for giving me strength, capability, and opportunity to do this work.

I would like to express my sincere thanks and gratitude to my supervisor **Prof. Mansour Alsulaiman**, professor in Computer Engineering, College of Computer and Information Sciences, King Saud University. I will always be grateful for his motivation, enthusiasm, patience, and continuous support throughout my master study.

In addition, I'd like to thank my co-advisor **Prof. Ghulam Muhammad**, professor in Computer Engineering, College of Computer and Information Sciences, King Saud University, for his valuable guidance, advice and good familiarity.

Furthermore, I would like to express my appreciation for the Center of Smart Robotics Research and the Deanship of Scientific Research, King Saud University for supporting my research. In addition, my sincere appreciation is extended to the Computer Engineering department for their assistance and cooperation.

Finally, I would like to thank my parents and my wife who tolerated my absence during the preparation of this thesis and for making my life so wonderful.

Hamdi Altaheri

TABLE OF CONTENTS

ABSTRACT	4
TABLE OF CONTENTS.....	7
LIST OF TABLES.....	9
LIST OF FIGURES	11
LIST OF ACRONYMS AND ABBREVIATIONS.....	15
CHAPTER 1	16
INTRODUCTION	16
1.1 PROBLEM STATEMENT	17
1.2 RESEARCH GOAL AND OBJECTIVES	18
1.3 CONTRIBUTIONS OF THE STUDY	18
1.4 ORGANIZATION	20
CHAPTER 2	21
BACKGROUND AND LITERATURE REVIEW	21
2.1 DATE FRUIT CHARACTERISTICS AND THEIR IMPORTANCE.....	21
2.2 DATE FRUIT HARVESTING.....	23
2.3 MACHINE VISION FOR HARVESTING ROBOTS	26
2.3.1 Imaging Sensors.....	26
2.3.2 Visual Features.....	28
2.4 RESEARCH IN MACHINE VISION FOR DATE FRUIT	29
CHAPTER 3	32
DATE FRUIT DATASET FOR INTELLIGENT HARVESTING	32
3.1 INTRODUCTION	32
3.2 DATASET DESCRIPTION	33
3.2.1 Dataset-1	34
3.2.2 Dataset-2	37
3.3 COLLECTION PROTOCOL AND EQUIPMENT.....	40
3.4 DATASET CODING.....	44
3.5 GENERAL STATISTIC OF THE DATASET	45
3.6 DATASET LABELING	49
3.7 DIVIDING DATASET-1 INTO TRAINING AND TESTING SETS	54
3.8 CONCLUSION.....	56

CHAPTER 4	58
DATE FRUIT CLASSIFICATION FRAMEWORK FOR ROBOTIC HARVESTING	58
4.1 INTRODUCTION	58
4.2 CHALLENGES OF DATE FRUIT CLASSIFICATION IN A NATURAL ENVIRONMENT	59
4.3 PROPOSED FRAMEWORK	60
4.3.1 Convolutional Neural Network Deep Learning	60
4.3.2 AlexNet Architecture	63
4.3.3 VGGNet Architecture	65
4.3.4 Transfer Learning.....	65
4.4 RESULTS AND DISCUSSIONS.....	68
4.4.1 Performance Evaluation.....	69
4.4.2 Visualization Results of the CNN models	70
4.4.3 Date Type Classification System	71
4.4.4 Date Maturity Classification System	74
4.4.5 Harvesting Decision System	82
4.4.6 Comparison with Other Classification Methods.....	85
4.5 CONCLUSION.....	89
CHAPTER 5	90
DATE FRUIT SEGMENTATION IN AN ORCHARD	90
5.1 INTRODUCTION	90
5.2 PROPOSED APPROACH.....	91
5.2.1 Pre-processing.....	91
5.2.2 Superpixels Clustering	93
5.2.3 Color Components	95
5.2.4 Local Binary Pattern	97
5.2.5 Support Vector Machine	97
5.3 RESULTS AND DISCUSSIONS.....	98
5.4 CONCLUSION.....	104
CHAPTER 6	105
CONCLUSION AND FUTURE WORK	105
PUBLICATIONS	107
REFERENCES	108

LIST OF TABLES

TABLE 1. The variations in the dataset-1. _____	36
TABLE 2. The number of images taken per session for all date varieties in the dataset-1. _	46
TABLE 3. List of all date palms in the dataset-2. _____	47
TABLE 4. An example of data for a date palm in the dataset-2. _____	48
TABLE 5. The number of individual dates in dataset-2. _____	48
TABLE 6. Samples of single date weights for the Barhi variety in three maturity stages: Khalal, Rutab, and Tamar. _____	48
TABLE 7. The length and weight measurements of Barhi date bunch stalks. _____	49
TABLE 8. Distribution of the dataset-1 images between the seven maturity classes and five type classes _____	50
TABLE 9. Description of the proposed maturity classes (phases). _____	52
TABLE 10. Dataset-1 images in the training and testing sets. (a) The number of training and testing images for the date type classification task showing their distribution among the six imaging sessions. (b) The number of training and testing images for the date maturity classification task showing their distribution among the five date types. _____	56
TABLE 11. The confusion matrix for the multiclass classification problem, with classes C_1, C_2, \dots, C_n . TP, TPR, and PPV refer to the true positive, true positive rate, and positive predictive value, respectively _____	70
TABLE 12. The confusion matrices (as illustrated in TABLE 11) of the type classification models on the testing dataset. _____	72
TABLE 13. Results of the proposed date type classification system on the testing dataset. _	73
TABLE 14. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the seven maturity classes. _____	76
TABLE 15. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the five maturity classes _____	79

TABLE 16. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the five maturity classes with doubling the number of training examples of the Khalal and Khalal-with-Rutab classes. _____	81
TABLE 17. Results of the proposed maturity classification system on the testing dataset using the different settings of the maturity classes. _____	82
TABLE 18. The confusion matrix of the harvesting decision model on the testing dataset. _	85
TABLE 19. Performance of the proposed classification systems (speed and accuracy measures). _____	85
TABLE 20. Performance comparison between the methods. _____	86
TABLE 21. Performance comparison between the methods using the same dataset adopted in this work. _____	88
TABLE 22. Cross-validation accuracy of the proposed date fruit segmentation system. __	100
TABLE 23. The performance of the proposed date fruit segmentation system. _____	102

LIST OF FIGURES

FIGURE 1. The main blocks of an agriculture robot showing the scope of this study.	17
FIGURE 2. Top ten date fruit producer countries in 2016, According to the Food & Agricultural Organization (FAO) [2].	22
FIGURE 3. Samples of Barhi date fruit in immature and three mature stages (Khalal, Rutab, and Tamar).	23
FIGURE 4. Harvesting the dates manually in a medium orchard (around 800 palms) in Al-Ammaria, Riyadh.	24
FIGURE 5. Date harvesting mechanism. (a) Ladders attached to date palm trees (b) Mechanical ladder.	25
FIGURE 6. A brief description of the date fruit datasets.	33
FIGURE 7. Sample images of a Sullaj date captured in six imaging sessions and covered all date maturity stages (immature, Khalal, Rutab, and Tamar).	35
FIGURE 8. Sample images of the dataset-1 showing large variation in scales, angles, and illumination.	35
FIGURE 9. Sample images of the five date types in dataset-1 showing different maturity stages. Some bunches are covered with bags for their protection.	36
FIGURE 10. A sample of the Barhi date bunch captured in front of graph paper.	38
FIGURE 11. A sample of date bunches marked in the Barhi date palm.	38
FIGURE 12. Date bunches from dataset-2 showing different levels of date compactness.	39
FIGURE 13. Sample images of one date bunch captured in ideal front view (with uniform background) and from other angles on the palm.	39
FIGURE 14. 64-individual Barhi dates at the four maturity stages (16 dates per stage): immature, Khalal, Rutab, and Tamar.	40
FIGURE 15. (a) Camera position setup for captured videos. The orientation of the camera depends on the palm height. Figure (a) shows a camera orientation of 45 degrees for a three-	

meter high palm tree. Figure (b) shows a real example for the palm-3 (B3.K.BW); the palm height is 2.85 meters, so the camera orientation was less than 45 degrees.	42
FIGURE 16. The Dimensions of the graph paper, showing the distances between the centers of the reference shapes.	43
FIGURE 17. Calculate the weight of a date bunch after harvesting. The bunch is placed in a container and weighted.	43
FIGURE 18. The template of reference paper used in the individual date imaging.	44
FIGURE 19. The camera position setup for the individual date imaging.	44
FIGURE 20. Date bunch stalk showing the ends we used to measure the length.	44
FIGURE 21. A description of the symbols used in the dataset coding.	45
FIGURE 22. The state of each palm during the six imaging sessions: harvested, partially harvested, or not harvested.	46
FIGURE 23. Sample images of Sullaj dates labeled into the seven maturity classes based on the threshold values described in TABLE 9.	50
FIGURE 24. Sample images of individual dates in the four maturity stages of the five date types. The description of date's colors at each stage is according to [45].	51
FIGURE 25. The proposed maturity classes of date bunches showing the harvesting decision for each class according to the bunch-based harvesting method, and the approximate thresholding based on the number of individual dates.	52
FIGURE 26. Distribution of the images captured at each imaging session between the seven maturity classes. Images captured in one imaging session are distributed between many classes. For instance, the images captured in session-4 are distributed between four classes for Meneifi dates (Immature-2, Pre-Khalal, Khalal, and Khalal-with-Rutab) and Sullaj dates (Pre-Khalal, Khalal, Khalal-with-Rutab, and Pre-Tamar).	53
FIGURE 27. Demonstration of the labeling process into the maturity classes. The process depended on the visual estimation of the number of individual dates in date bunches that belonged to the four maturity stages.	54

FIGURE 28. Sample images of three date types at the Khalal stage that are difficult to distinguish.	59
FIGURE 29. Sample images of date bunches that have individual dates at different maturity levels, which makes them hard to label or classify.	59
FIGURE 30. Block diagram of the proposed date fruit classification framework. (a) Training phase, (b) real-time employment.	61
FIGURE 31. Deep learning architectures of the proposed date fruit classification framework based on AlexNet pre-trained model.	64
FIGURE 32. Deep learning architectures of the proposed date fruit classification framework based on VGG-16 pre-trained model.	66
FIGURE 33. Visualization of the type classification models based on AlexNet (1) and VGG-16 (2). The figure shows first 36 features (filters) of the first and last convolutional layers in the top, and the activations of these features are shown in the bottom.	71
FIGURE 34. The behavior of the training and validation accuracies and losses of the date type classification models during the training phase.	73
FIGURE 35. The behavior of training and validation accuracies and losses of the date maturity classification models based on the seven maturity classes.	75
FIGURE 36. Samples of misclassified images using the maturity classification CNN models on the seven maturity classes.	77
FIGURE 37. The behavior of the training and validation accuracies and losses of the date maturity classification models based on the five classes.	78
FIGURE 38. The behavior of the training and validation accuracies and losses of the date maturity classification models based on the five maturity classes with doubling the number of training examples of the Khalal and Khalal-with-Rutab classes.	80
FIGURE 39. Samples of misclassified images using the maturity classification CNN models on the five maturity classes	81
FIGURE 40. The harvesting decision of dates depending on the harvesting methods during the development period.	83

FIGURE 41. Possible harvesting stages of the five date types in the dataset. _____	83
FIGURE 42. Labeling each date type in the testing dataset according to the harvesting decision. Decisions are based on the assumption that the required harvesting stage of Barhi is Khalal, Meneifi and Sullaj is Rutab, and Khalas and Naboot Saif is Tamar. _____	84
FIGURE 43. Comparison between the accuracy obtained by the proposed approach for date classification problems and the approaches in [27], [31], and [33]. _____	88
FIGURE 44. Average classification time of the date harvesting decision system based on the proposed method and the methods in [27], [31], and [33], using a machine with a GPU and CPU. _____	89
FIGURE 45. Block diagram of the date fruit segmentation system. _____	92
FIGURE 46. Standard k-mean searches the entire image (a), while SLIC searches a limited region (b). _____	93
FIGURE 47. Date palm image clustered using the SLIC algorithm with a different number of superpixels. _____	95
FIGURE 48. Samples of date fruit and background regions. _____	99
FIGURE 49. Label an image into fruit and background regions using image Labeler app in Matlab. _____	99
FIGURE 50. Graphical interface (UI) to help segmenting the images into fruit and background. _____	100
FIGURE 51. Samples of the color components of images that can discriminate the date bunches from the background. (a) Original image; (b) R image component; (c) Cb image component; (d) H image component. _____	101
FIGURE 52. Samples segmented images using our proposed fruit segmentation method. _	103

LIST OF ACRONYMS AND ABBREVIATIONS

<i>FAO</i>	<i>Food and Agricultural Organization of the United Nations</i>
<i>CCD</i>	<i>Charge-Coupled Device</i>
<i>SLIC</i>	<i>Simple linear iterative clustering</i>
<i>CMOS</i>	<i>Complementary Metal-Oxide-Semiconductor</i>
<i>MPEG</i>	<i>Moving Picture Experts Group</i>
<i>MTS</i>	<i>MPEG Transport Stream</i>
<i>HD</i>	<i>High Definition resolution</i>
<i>SVM</i>	<i>Support Vector Machines</i>
<i>CNN</i>	<i>Convolutional Neural Networks</i>
<i>FC</i>	<i>Fully Connected layer</i>
<i>ReLU</i>	<i>Rectified Linear Unit</i>
<i>LRN</i>	<i>Local Response Normalization</i>
<i>VGG</i>	<i>Visual Geometry Group, Oxford university</i>
<i>CPU</i>	<i>Central Processing Unit</i>
<i>GPU</i>	<i>Graphics Processing Unit</i>
<i>WLD</i>	<i>Weber Local Descriptor</i>
<i>LBP</i>	<i>Local Binary Pattern</i>
<i>PPV</i>	<i>Positive Predictive Value</i>
<i>TPR</i>	<i>True Positive Rate</i>
<i>TP</i>	<i>True Positive</i>
<i>ACC</i>	<i>Accuracy</i>
<i>fps</i>	<i>Frames per second</i>
<i>HSV</i>	<i>Hue, Saturation, and Value color space</i>
<i>YCbCr</i>	<i>Luminance (Y), Chrominance (Cb and Cr) color space</i>
<i>RGB</i>	<i>Red, Green, and Blue color space</i>

CHAPTER 1

INTRODUCTION

Global date fruit production was 8.5 million tons in 2016 according to the Food & Agriculture Organization (FAO) [2]. Date fruit cultivation is a major strategic agricultural industry in Middle East and North Africa countries, which produce 91% of the world's dates [2]. In date cultivation, manual harvesting is the dominant method used, which requires skilled workers to climb palm trees to reach date bunches. However, manual harvesting is dangerous and labor-intensive as well as inefficient in terms of both time and cost. Such methods are the major cause of delays in the date production cycle and account for more than 45% of the date production cost [3]. Recently, due to the increase in date palm cultivation and shortage of skilled workers, the cost of date harvesting has increased significantly, necessitating a change to automated harvesting. Advanced agricultural automation such as robotic harvesting can significantly increase quality and yield as well as reduce production costs and delays [4]. One of the most important aspects of harvesting robots is their ability to interpret and analyze visual data. Using an accurate vision system to detect and classify fruits in real time is critical for the harvesting robot to be cost-effective and efficient.

The goal of this study is to develop a vision system for date fruit harvesting robot in an unstructured natural environment. Agriculture robot consist in general of three main parts: a vision system, a control system, and a mechanical system, as shown in **FIGURE 1**. The vision system provides information to the control system that gives orders to control the mechanical system. Our scope in this study is focused on the vision system of the harvesting part of the agriculture robot.

The main problems of the machine vision of a date fruit harvesting robot are demonstrated in section 1.1. The objectives described in section 1.2 constitute a roadmap for research in this study. The contributions of the study based on these objectives are reported in section 1.3

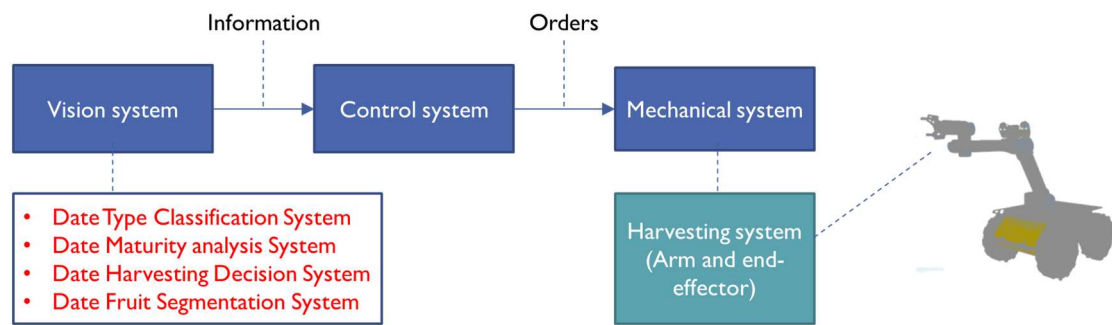


FIGURE 1. The main blocks of an agriculture robot showing the scope of this study.

1.1 PROBLEM STATEMENT

Practical success in the area of machine vision for harvesting robots is still limited due to the difficulties caused by unstructured and unconstrained agricultural environments [4], [5]. Furthermore, to the best of our knowledge, there is no research in the area of machine vision for date fruits in an orchard environment. Automatic fruit detection, classification, and maturity analysis in a natural environment are challenging machine vision tasks due to the difference in size, shape, color, and texture properties of various fruits; the large degree of uncertainty with unstructured orchard scenes; harsh occlusions; and highly variable illumination and shadow states. However, these tasks are more complicated in date fruits due to several aspects. First, date orchards usually have many types (varieties) of dates with numerous similarities in their visual appearance. Second, different date types in the same orchard are harvested in different maturity stages. Third, individual dates in one bunch do not mature uniformly (i.e. a date bunch usually has individual dates at different maturity stages), which complicates

labeling or classifying date bunches into a specific maturity class. Fourth, date bunches can be covered with net bags, which distort their visual features. In this study, we address all these challenges.

1.2 RESEARCH OBJECTIVES

The objectives of the study are:

- Build a database for date bunches' types and weight in different maturity stages (before, during, and after maturing). This database is intended to provide the requirements needed to automate date palm agricultural applications in the orchard, including crop yield estimation and robotic harvesting.
- Develop a vision system capable of automatically:
 - detecting the presence of date fruits in orchards and discriminating them from the rest of the scene (trunk, fronds, leaves, ground, and sky),
 - recognizing the variety of date fruits in an orchard environment,
 - classifying date fruit bunches in an orchard environment according to their maturity stages,
 - deciding the harvesting decision based on the type of the date palm and the ripeness of the date fruits.
- Evaluate the system in a real and unstructured natural environment.

1.3 CONTRIBUTIONS OF THE STUDY

The main contribution of this study is building a machine vision system for date fruit harvesting robots to detect date fruits bunches, classify them according to their types, and analyze their maturity and harvesting stages. To the best of our knowledge, there

is no research on date detection, classification, or maturity analysis before harvesting in an orchard environment. In this study, we highlight the following contributions:

- (1) We built a comprehensive dataset for date fruits that can be used for multiple tasks including automated harvesting, visual yield estimation, and classification tasks. The dataset has been fully labeled, coded, and released with their associated files to the research community in [1], (<http://dx.doi.org/10.21227/x46j-sk98>). To the best of our knowledge, this is the only publicly available dataset for date fruit pre-harvesting and harvesting applications.
- (2) We analyzed the maturity stages of date fruits and proposed labeling date fruit bunches in orchards into seven maturity classes based on the harvesting decision of date bunches and the maturity index of individual dates.
- (3) We proposed three classification systems for date fruit harvesting robots in an orchard environment based on deep learning. The systems classify date fruit bunches in real time according to their type, maturity, and harvesting decision. Deep convolutional neural networks (CNNs) with transfer learning and fine-tuning were applied in the classification tasks to achieve high performance with minimum training time using a relatively small dataset.
- (4) We proposed a robust method to segment date fruits from trees and orchard's background. Simple linear iterative clustering (SLIC) was used to initially group pixels in an image into superpixels, and then, these superpixels were classified into fruit or background using local binary pattern (LBP), as a feature descriptor, and support vector machine (SVM), as a classifier.
- (5) The source codes, datasets, systems, and videos of this study are publically available for the benefits of researchers on: <https://sites.google.com/view/daterobotic>.

1.4 ORGANIZATION

Chapter 1 discusses the problems and challenges of automatic date fruit detection, classification, and maturity analysis in a natural environment. The objectives and research achievements of the study are summarized in this chapter.

Chapter 2 explores the research background and related works. In this chapter, we introduce date fruits and their harvesting stages and methods and demonstrate the need for automatic harvesting. Then, we survey the machine vision for harvesting robots and the current research in the vision system of date fruits.

Chapter 3 presents our date fruit dataset for automated harvesting and visual yield estimation. The chapter describes the dataset and its collection protocol, coding scheme, statistics, labeling methods, and the rules used to divide the dataset into training and testing sets.

Chapter 4 presents the proposed date fruit classification framework for robotic harvesting. The chapter first illustrates the challenges of date fruit classification in a natural environment and then describes the proposed framework and discusses the results.

Chapter 5 presents the proposed date fruit segmentation system. The chapter describes the proposed method and then presents the experimental results with the discussion.

Finally, in **chapter 6**, we draw the conclusions and propose future directions.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

This chapter explores the research background and related works. Section 2.1 describes the characteristics of date fruit and their commercial significance. The harvesting stages and methods of date fruit are presented in section 2.2. Section 2.3 discusses the machine vision for harvesting robots. Finally, we investigate the current research in the vision system of date fruit in section 2.4.

2.1 DATE FRUIT CHARACTERISTICS AND THEIR IMPORTANCE

Date palm is one of the oldest and most valuable fruit trees in the world. This is valuable because of health benefits, ritual importance in human cultures, productive capacity in dry and harsh environments, and the variety of subsistence products from its fruits and other parts of the palm [6]. Date fruit is considered a health mine because it is very rich in minerals, like magnesium, phosphorous, iron, etc. Dates are also a good source of vitamins A, B1 (thiamine), B2 (riboflavin), and B7 (nicotinic acid); natural sugar, protein, and fiber [7].

Date production is one of the major agricultural industries, producing approximately 8.5 million tons of fruit in 2016 according to the Food & Agricultural Organization (FAO) [2]. Countries in the Middle-East and North Africa are the largest date producers in the world; they produce approximately 91% of the world's dates [2]. **FIGURE 2** shows the dates production in tons of the top ten producing countries. Date production is an important and strategic agriculture industry in these countries.

Date fruit is the main and strategic agriculture industry in Saudi Arabia, covering approximately 72% of the total area of permanent crops. According to the Ministry of

Agriculture - Saudi Arabia, an estimated 24 to 25 million date palm trees produce nearly a million tone of dates annually accounting for about 11.4% of the global world dates in 2016 (FAO) [2]. According to FAO, the production of dates of Saudi Arabia was around half million tons in 1985 and it became around one million tons by 2016 [2].

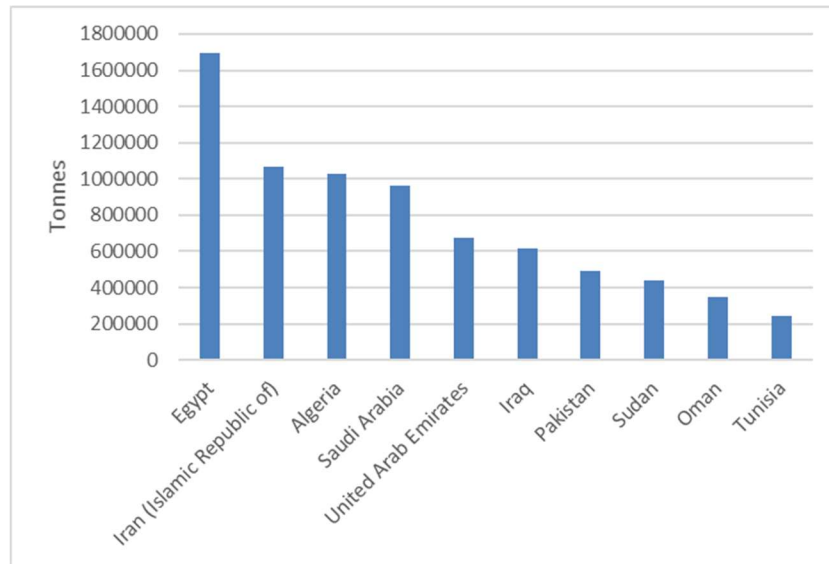


FIGURE 2. Top ten date fruit producer countries in 2016, According to the Food & Agricultural Organization (FAO) [2].

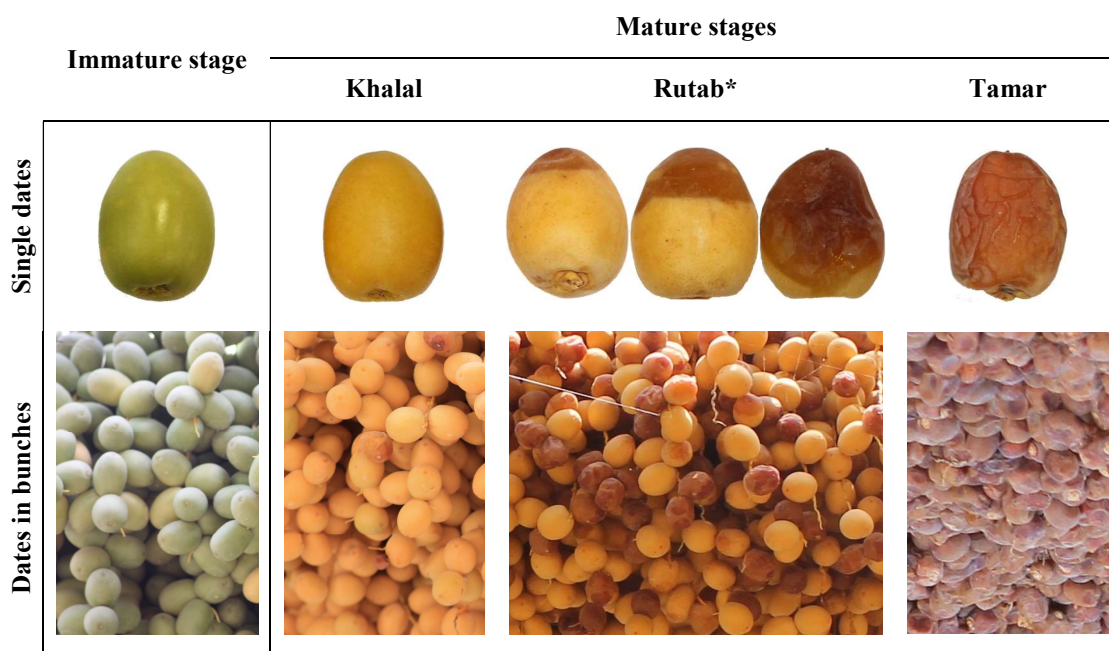
Date fruits are considered one of the most suitable crops for exportation and consumption over the year, which can be harvested and marketed as fresh ripe fruits at three stages of their development—Khalal, Rutab, and Tamar, as shown in **FIGURE 3**.

The characteristics of these stages are [7]:

Khalal: Bright yellow or red, physiological mature, crisp and hard, perishable, moisture range from 50% to 85%.

Rutab: Partially browned in color, fibers softened, perishable, moisture from 30% to 45%.

Tamar: Texture from soft pliable to firm to hard, color from amber to dark brown, moisture content below 25%, protected from insects, and can be stored for a long time.



* The Rutab stage starts with the appearance of the color at the date's tip and then spread gradually to the whole date.

FIGURE 3. Samples of Barhi date fruit in immature and three mature stages (Khalal, Rutab, and Tamar).

2.2 DATE FRUIT HARVESTING

Dates are harvested in August at the Khalal stage or in September to December at the Rutab and Tamar stages [8]. Dates in Khalal are the first in the harvesting season where dates are regarded as ready for sale as "fresh" fruit. The Rutab stage is the second in the harvesting season. Rutab is delicate, highly perishable, and produced in a short period of time, which makes handling and transporting costly and difficult. Tamar is the final development stage of dates. In this stage, dates are non-perishable, so they can be stored for a long period of time and can be consumed throughout the year. It also can be used for the production of various types of products, e.g. sauces, cakes, and components of date honey or granules [7].

The time of harvesting the date fruits depends on fruit's texture and appearance (related to sugar content and moisture). The choice for harvesting at one or another stage depends on several factors, including climatic conditions, date type (fruit

characteristics), and market demand [7]. The decision to harvest date fruit depends on both fruit type (variety) and maturity stage. Date type is a key factor in determining the stage at which dates are harvested, unlike most fruits that only depend on the maturity stage of an individual fruit (i.e. ripe or not). For instance, some types of dates are harvested in the first maturity stage (Khalal) such as Barhi, whereas other date types are usually harvested later in the Rutab or Tamar stages such as like Sullaj and Khalas.

The whole bunches of Rutab and Tamar dates are harvested when the most dates are mature. Then, bunches are lowered to the ground and shaken into a box to take out the ripe dates [8]. Alternatively, individual mature dates are picked directly from the bunches while they are on the palm tree, hence in this way three pickings sessions, on average, are required over several days. Pickers use several types of harvesting aids and containers to lower the dates to the ground, as shown in **FIGURE 4**.



FIGURE 4. Harvesting the dates manually in a medium orchard (around 800 palms) in Al-Ammaria, Riyadh.

As the palm tree grows taller, harvesting the dates becomes costlier and more difficult [5]. To facilitate harvesting, ladders may be mounted on the palm, as shown in **FIGURE 5-a**. Various types of lifts, such as self-propelled elevating and tree squirrel, are also used to facilitate harvesting by elevating the laborers to the top of the palms [6], as shown in **FIGURE 5-b**.

Manual harvesting method is widely used for date fruits. It requires skilled labors to climb the palm tree to reach date bunches [6], as shown in **FIGURE 4**. Manual harvesting is inefficient in terms of both time and economy. It is highly labor intensive, dangerous, body demanding, and exhausting task. The study in [3] reported that date harvesting is performed manually and this is the major reason for delays in the date production cycle. The study also pointed out that 45% of the total production cost is spent on the harvesting phase. In Middle East countries such as Saudi Arabia, a significant number of farmers are suffering from a lack of skilled labor and are forced to sell their produce on the trees to investors at a low price.

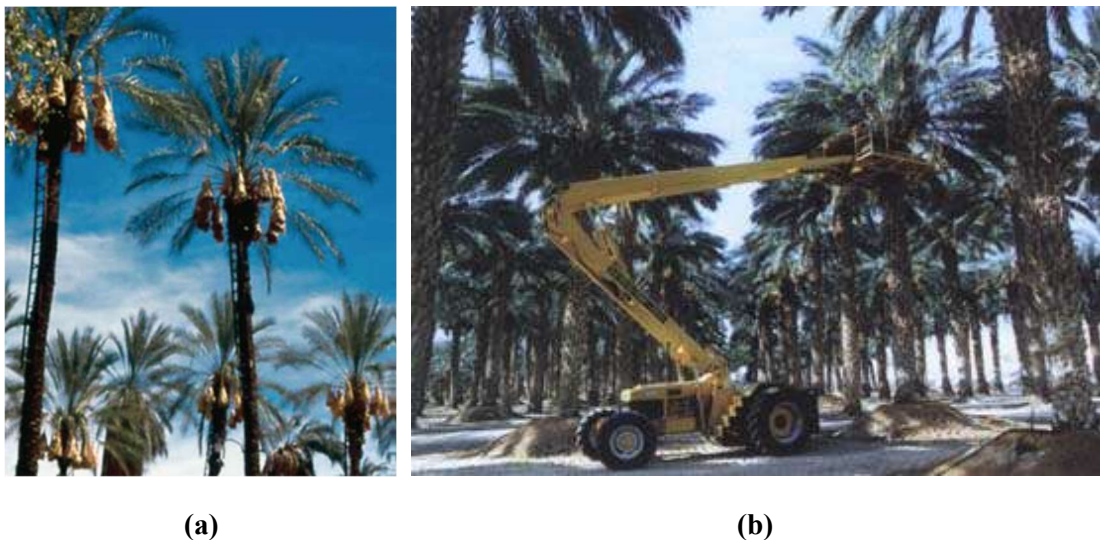


FIGURE 5. Date harvesting mechanism. (a) Ladders attached to date palm trees (b) Mechanical ladder.

Machine harvesting system can overcome the limitations of manual harvesting. Machine harvesting removes fruits from the trees efficiently, hence reducing the

harvesting cost to about 45% of the total production cost [3]. For fruit in general, two streams of machine harvesting systems have been researched since early 60's. The first stream is mechanical harvesting methods; such as air blasting, limb shaking, trunk shaking, canopy shaking, etc. [9]. However, the mechanical harvesting system cannot ensure the quality and size selection that the human vision can. The second stream is the automatic harvesting system, which is considered as a better alternative method to mechanical harvesting [9]. Automated harvesting has increased productivity many folds by reducing production costs and manual labor, increasing quality and yield [4]. Our work is a contribution to the automatic harvesting.

2.3 MACHINE VISION FOR HARVESTING ROBOTS

The interpretation and analysis of visual information are one of the most important aspects in successful harvesting robots. Indeed, the coupling of proper machine vision algorithms with visual data will advance agricultural automation to new levels [4]. Accurate machine vision is critical for achieving high detection rates of mature fruit in real time; this is a compulsory precondition for a harvesting robot to be cost-effective and efficient. The accuracy of machine vision systems is influenced by the complex and variable canopy structures, variable and uncertain lighting conditions, and varying size shape, and color of the fruit [10]. Many studies have been conducted in the past to detect or classify fruit in outdoor environments. Various types of imaging sensors and visual features employed in harvesting environment are reviewed in the following subsections.

2.3.1 Imaging Sensors

The variability of sensor types and configurations is large, ranging from a single grey level camera to the hyperspectral cameras. However, the most widely used sensor in

the machine vision system is a standard color camera. Imaging sensors are usually positioned on the main platform or on the body of the robot to provide a single view to the scene.

Black and white (BW) cameras were used in some of the earliest studies for fruit detection and classification. For instance, the study in [11] used BW camera to detect melons based on geometric, texture, and reflectance features and obtained 82-88% accuracies. BW cameras were replaced later with color cameras. The major limitation of BW cameras is the lack of color information, which is one of the most important features of fruit. Color cameras with CMOS (Complementary Metal-Oxide-Semiconductor) or CCD (Charged Coupled Device) sensors have been extensively used by researchers in computer vision systems for harvesting robots [12]–[16].

Stereo camera can be used to extract the depth information on the imaged objects [17]. However, the depth information that is obtained only from visual stereo sensors may not be accurate. Some researchers have merged range sensors with visual input in order to obtain accurate depth information.

Thermal cameras can be used to detect or classify fruits based on their thermal response. For instance, [18], [19] used thermal cameras to detect citrus and apple fruits based on the temperature difference between the fruit and background. The fruit absorbs more heat and radiates more temperature in comparison with stems and leaves, which allows for distinguishing between the fruits and their background items.

With the improvement of spectroscopy technology, spectral imaging has become more common for recognition of objects based on their different reflectance in selected wavelengths. Hyperspectral sensor provides the complete spectral signature for each pixel in the visual field of the camera. Clearly, the hyperspectral sensor offers a massive

amount of information that can be used to enhance the accuracy of fruit detection and classification. However, this comes at a costly price, both in processing and acquisition time, nearly some minutes per image [4]. Hence, it is generally used for offline processing. The studies in [20]–[22] used spectral imaging to detect fruits based on their reflectance at different wavelengths.

2.3.2 Visual Features

Computer vision algorithms are characterized by many visual cues, such as color, spectral reflectance, thermal reflectance, texture, and geometric. Machine vision algorithms in harvesting robots have tried to use and integrate many of these visual cues in order to take advantage of all possible information in the harvesting environment.

Fruit color is considered one of the most important visual cues used in machine vision in the harvesting robots. However, the color feature is sensitive to lighting variation. Also, it is not effective when the fruits have a color that is less distinct from the rest of the plant materials, e.g. green apples or immature fruits [4]. A color-based classification was used by many researchers to detect fruits with distinct colors; including apples [14], citrus [15], grapes [16], almonds, and mangoes [23].

Texture is another significant visual cue that can be useful in classifying different fruits and separating the fruit from the background. Different fruits have distinct texture patterns and fruits generally have smoother surfaces than background items such as stems and leaves [5]. The texture is the repeated visual pattern that covers regions and surfaces either randomly or regularly. Texture features are extracted using several texture descriptors such as local binary patterns (LBP) and Weber local descriptors (WLD). Many studies have used texture feature to detect or classify fruits, such as [16],

[24]–[27]. These studies obtained accuracies of 75-98% when texture features were used.

Shape and size features are global visual cues that could be less sensitive to illumination. Fruit shape is becoming more common in harvesting robots, although it is more computationally demanding to analyze and extract [4]. Shape and size provide another distinct set of features of some fruits such as dates, citrus, apple, and pear. The work in [27] used support vector machine (SVM) to classify date fruits in a lab environment using WLD texture descriptor combined with size and shape features. The study reported accuracies of 95% and 87% for texture and size-shape features respectively, and 98.1% by fusing all these features.

2.4 RESEARCH IN MACHINE VISION FOR DATE FRUIT

Many studies have been performed to classify and harvest fruits other than dates. Compared with the research on other fruits, especially in the area of automated harvesting, research on date fruits is negligible. Most previous research has focused on dates after the harvesting phase, with research on date fruit detection, classification, or maturity analysis in an orchard environment lacking. Though there were some research efforts by Al-Janobi et al. in the way of automated date palm harvesting [3], [28], the work did not address the problem of the computer vision for the harvesting robot. Current studies in date fruits can be categorized into three groups: type classification [27], [29]–[34], maturity analysis [35]–[37], and quality grading [38]–[46].

Most research on dates focuses on grading fruit quality. Date quality can be evaluated by several factors such as moisture and sugar content, hardness, and surface defects. For instance, in [39], the authors proposed an electronic sensor to measure the moisture content of dates to classify them into moist, semi-moist, and dry. In [40], researchers

adopted the hardness of dates as a quality indicator, using stepwise and linear discriminate analysis to grade dates into soft, semi-hard, and hard. Grading dates according to surface defects has been investigated in many studies. For example, [41], [42] proposed techniques using a co-occurrence matrix and color machine vision for date grading. The study in [43] sorted dates as defected or good using image analysis techniques. In [44], a back-propagation neural network was used to classify dates into three quality grades using size, shape, intensity, and surface defect features. In the more recent automatic date grading approach proposed by [46], fuzzy inference was used to measure the quality of dates based on length and freshness features.

Unlike quality grading, few studies have analyzed the maturity of date fruits. The study by [35] investigated date maturity using infrared spectrometry. In [36], the authors used a color distribution analysis and back projection to classify one date type into four maturity classes. In another study [37], a taxonomy classification method with RGB color and texture features including contrast, entropy, and homogeneity was used to classify one date type into four maturity classes.

The classification of date fruits according to their type also has limited research. In [29], the researchers developed a date fruit classification system using neural networks, and in [30] they used probabilistic neural networks for a similar task. In [31], 15 size, shape, color, and texture features were used to classify single date images into seven classes according to their type. In another study [27], a technique was proposed based on shape and size features and local texture descriptors to classify four classes of dates using 800 single date images. In a more recent study [32], the authors classified single date images based on their types using a dataset containing 5000 images of 10 date types. They used an RGB color histogram, a gray-level co-occurrence matrix (GLCM), and four shape features including area and eccentricity, combined with Gaussian mixture model

(GMM). All these studies used single date images with a uniform background. However, the handcrafted feature-based approaches used in previous research are unsuitable for constructing a robust vision system that works in unstructured natural environments such as that of date orchards. In the most recent approach [33], the researchers reported an accuracy of 99.2% using a deep learning based technique for date fruit classification. They built a dataset of four date types by acquiring date images from the Google search engine.

None of the previous works has investigated the problem of date fruit detection and segmentation, variety classification, or maturity analysis in an orchard environment. All previous approaches have used datasets with images from a limited perspective. Further, most such images were for individual dates in the post-harvesting or post-production stages, so they were solving simpler problems than this work, which solves a real-life problem that has many difficulties.

CHAPTER 3

DATE FRUIT DATASET FOR INTELLIGENT

HARVESTING

3.1 INTRODUCTION

Research on automated date fruit harvesting is limited as there is no public dataset for date fruits to aid in this. Hence, in this work, we built a comprehensive dataset for date fruits that can be used by the research community for multiple tasks including automated harvesting, visual yield estimation, and classification tasks. This dataset contains images of date fruit bunches of different date types (varieties), captured at different pre-maturity and maturity stages. These images cover multiple sets of variations such as multi-scale images, variable illumination, and different bagging states. We also marked date bunches for selected palms and measured the weights of the bunches, captured their images on a graph paper, and recorded 360° video of the palms. Our dataset can help in advancing research and automating date palm agricultural applications, including robotic harvesting, fruit detection and classification, maturity analysis, and weight/yield estimation. The dataset has been fully labeled, coded, and released with their associated files freely to the research community in the IEEE DataPort repository [1].

Referred papers:

- H. Altaheri, M. Alsulaiman, M. Faisal, and G. Muhammed, "Date Fruit Dataset for Automated Harvesting and Visual Yield Estimation", IEEE DataPort, v1, 2019. [Online]. Available: <http://dx.doi.org/10.21227/x46j-sk98>.
- H. Altaheri, M. Alsulaiman, G. Muhammad, S. U. Amin, M. Bencherif, and M. Mekhtiche, "Date fruit dataset for intelligent harvesting," Data Br., vol. 26, p. 104514, Oct. 2019.

Refer to the following websites to download the related data:

<https://sites.google.com/view/daterobotic/date-database>

<http://dx.doi.org/10.21227/x46j-sk98>

3.2 DATASET DESCRIPTION

The date fruit dataset was created to address the requirements of many applications in the pre-harvesting and harvesting stages. The two most important applications are automatic harvesting and visual yield estimation. Since date applications require different types of data with different characteristics, we built two separate datasets for the benefits of researchers in different date applications. **FIGURE 6** gives a brief overview of the two datasets.

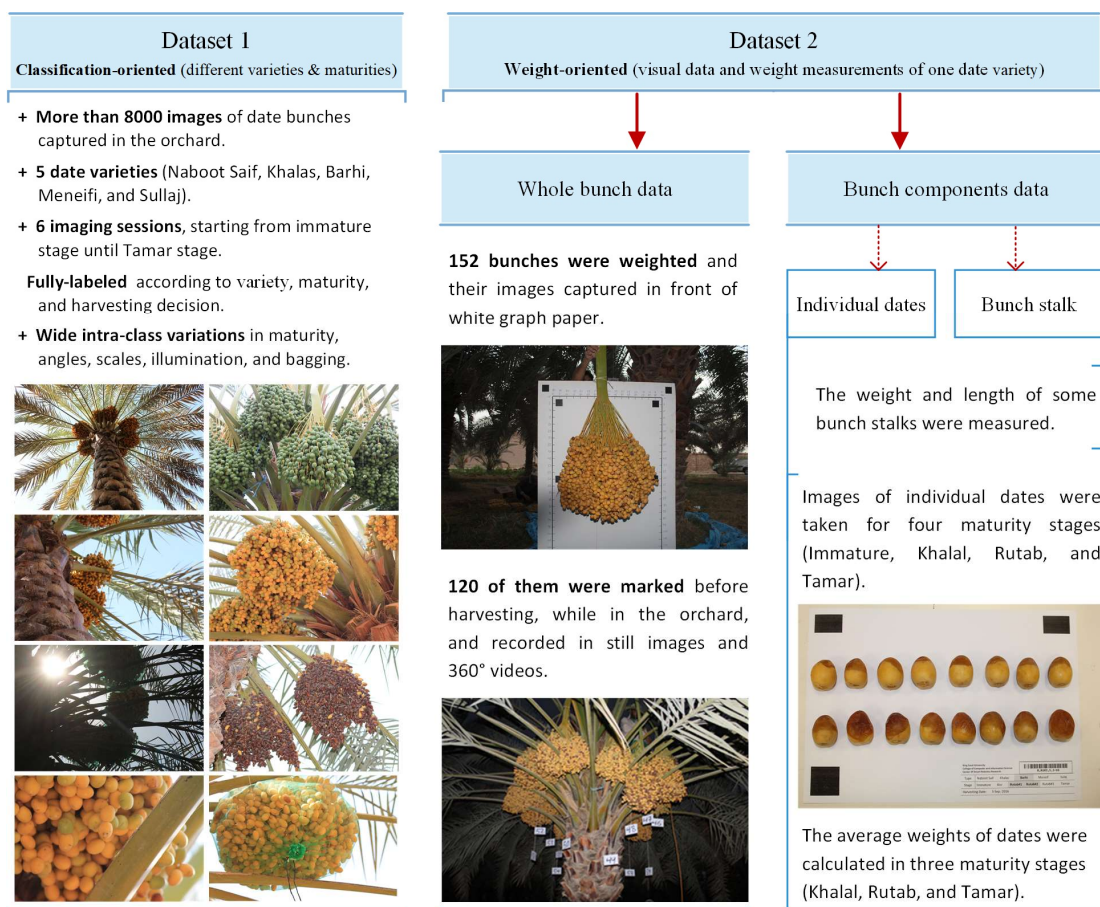


FIGURE 6. A brief description of the date fruit datasets.

The first dataset (dataset-1) contains images of date fruit bunches for different date varieties, showing various pre-maturity and maturity stages. These images cover a large degree of variabilities in order to address the challenges in the date orchard environment, such as multi-scale images, variable illumination, different angles, and diverse bagging

states. Dataset-1 can be used in many applications such as automatic harvesting tasks, fruit detection, fruit recognition, maturity analysis, etc.

The second dataset (dataset-2) contains images, videos, and weight measurements to help in many applications such as yield estimation. In this dataset, we marked date bunches for selected palms, recorded 360° video for each palm, and measured their data (height, trunk circumference, total yield, number of bunches, and weight of bunches). We also captured images of each bunch from different angles before harvesting and on a graph paper after harvesting. Due to the enormous work needed and the huge amount of data, this dataset built for only one date variety. However, the proposed method for yield estimation (as a major application of dataset-2) based on one date variety can be generalized for other varieties because visual weight features, e.g. bunch area, perimeter, compactness, volume, etc., are almost invariable in different date varieties.

3.2.1 Dataset-1

Dataset-1 consists of 8079 images of more than 350 date bunches captured from 29 date palms. The date bunches belong to five date varieties: Naboot Saif, Khalas, Barhi, Meneifi, and Sullaj. The images of date bunches were captured using two color cameras, their models and settings are described in section 3.3, in six imaging sessions (recording times). The imaging sessions covered all date maturity stages: immature, Khalal, Rutab, and Tamar. **FIGURE 7** shows sample images of the Sullaj date fruit over the six sessions. The dates in one session may have multiple maturity stages because individual dates do not mature uniformly. The first session was when the dates are usually immature, and the last session was when the dates were in the Tamar stage. Dataset-1 was labeled according

to the variety, maturity, and harvesting decision. The annotation files and labeling instructions are available in [1].

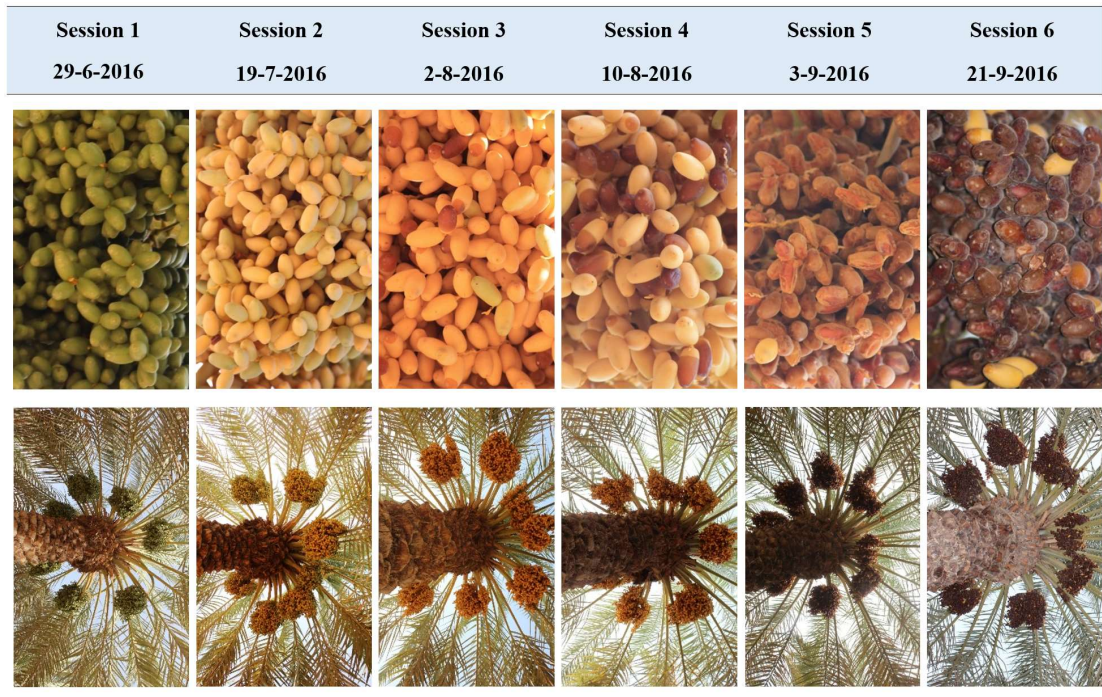


FIGURE 7. Sample images of a Sullaj date captured in six imaging sessions and covered all date maturity stages (immature, Khalal, Rutab, and Tamar).

In order to build a comprehensive dataset that can address various machine vision applications and challenges for date fruits in an orchard environment, the dataset was provided with multiple sets of variations: multi-scale images, variable illumination, different date varieties, multiple maturity stages, and different bagging states, as shown in **FIGURE 8** and **FIGURE 9**. **TABLE 1** highlights these variations.



FIGURE 8. Sample images of the dataset-1 showing large variation in scales, angles, and illumination.

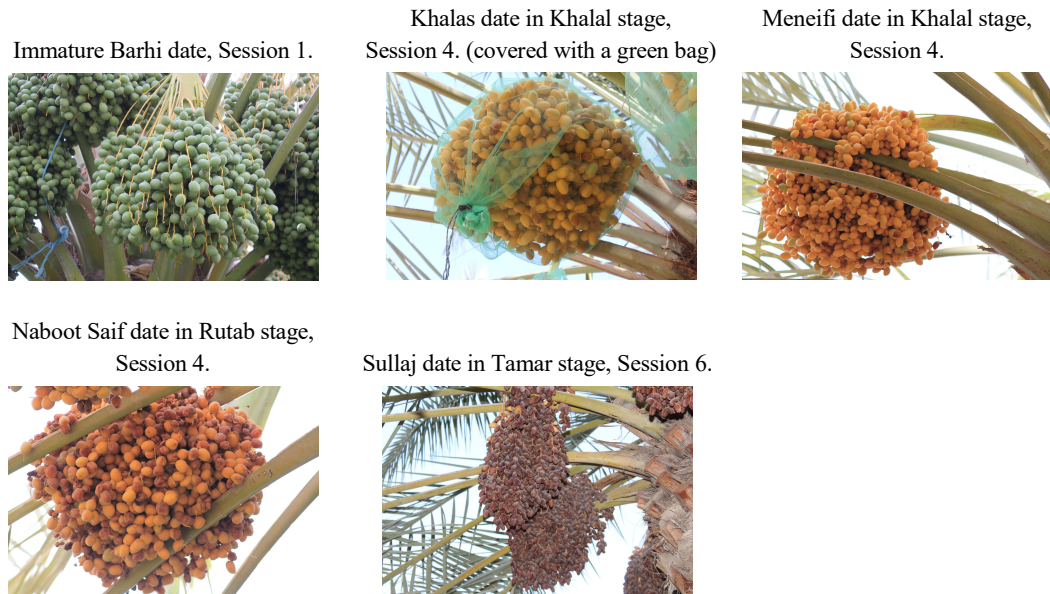


FIGURE 9. Sample images of the five date types in dataset-1 showing different maturity stages. Some bunches are covered with bags for their protection.

TABLE 1. The variations in the dataset-1.

Variation Type	Description
Multi-scale images	Images containing part of the bunch, the whole bunch, or many bunches.
Variable illumination	The images were taken under different natural daylight conditions: in the morning (9:00-11:00) and afternoon (3:00-5:00). The images also were captured in poor illumination conditions from different camera positions relative to the sun.
Different date varieties	Five varieties: Naboot Saif, Khalas, Barhi, Meneifi, and Sullaj.
Multiple maturity stages	The images were taken in six imaging sessions, which covered all date maturity stages: immature, Khalal, Rutab, and Tamar.
Different bagging states	Bagging is performed for some high-quality date varieties for their protection. The images include some bagged date bunches.

The dataset includes multi-scale images of date bunches, as shown in [FIGURE 8](#), to increase the variation in images, which helps make the vision system robust and versatile for different harvesting methods. Some images contain part of the bunch, which helps in individual-date-based harvesting, where the vision system must detect individual dates and harvest them based on maturity. Images that contain one or many bunches can be used for bunch-based harvesting, where the vision system would detect and harvest a

complete bunch based on its maturity. Some images contain whole date palms, and these can be used for applications such as palm recognition and maturity classification. The dataset also includes images for different date varieties in different maturity stages. This diversity is particularly essential for automated date fruit harvesting because date orchards usually contain different varieties of dates that are harvested and marketed at different maturity stages, e.g., Barhi dates are harvested early in the Khalal stage, whereas Khalas dates are usually left until the Tamar stage. This dataset also includes images of bagged bunches for dates of Khalas variety. These bunches were covered with green bags in the Khalal maturity stage, as shown in **FIGURE 9**. In date orchards, bunches of some date varieties are covered (usually at the beginning of August) with a bag to protect them from dust, pests, and rain. The images of bagged bunches can be used to increase the robustness of the classification system by training the system to recognize the variety or maturity state of dates even if the bunch is bagged. It will also allow the classification of date bunches as bagged or not bagged, hence can be used for developing automatic bagging machine.

3.2.2 Dataset-2

This dataset consists of 152 Barhi date bunches belonging to 13 palms. All bunches were weighted after harvesting, and their images were captured in front of a white graph paper, as shown in **FIGURE 10**. Subset of these bunches (120 bunches from nine palms) were provided with more inclusive data: we marked the whole bunches in the palm (as in **FIGURE 11**), captured their images from different angles before and during harvesting, recorded 360° video for each palm, and registered their characteristics (height, trunk circumference, total yield, and number of bunches).

The selected 13 date palms had different heights and ages (1.7 – 4.85 m and 10 – 25 years) to maintain the variation in bunch area, perimeter, and date size as much as possible. The dataset also has variations in bunch compactness and the angle of view (FIGURE 12 and FIGURE 13). Including these variations are important for building an effective and reliable machine vision system for weight and yield estimation. The date compactness in a bunch is affected by the thinning process, the operation of removing some stands or dates from a bunch to enhance its quality. Therefore, different date bunches with different degrees of thinning levels were selected, as shown in FIGURE 12. We also captured several images for each date bunch from various angles of view. One image was captured from a constant distance and front view-point, and other images from different variable angles, sample images are shown in FIGURE 13.



FIGURE 10. A sample of the Barhi date bunch captured in front of graph paper.



FIGURE 11. A sample of date bunches marked in the Barhi date palm.



FIGURE 12. Date bunches from dataset-2 showing different levels of date compactness.



FIGURE 13. Sample images of one date bunch captured in ideal front view (with uniform background) and from other angles on the palm.

The uniform background in the graph paper is used to achieve better image segmentation, where we can get a high contrast between the background and the bunch. The graph paper has six reference objects, which are uniquely identifiable by shape (square), color (black), and position (corners). The graph paper with the reference objects can be used to estimate geometric information of the bunch, i.e., area, dimensions, and perimeter.

In addition to the whole bunch data, other information was collected: images and weights of the bunch components (individual dates and bunch stalk). The primary purpose of this information is to estimate the bunch weight based on its components, i.e., by using dates' dimensions and weights and estimating the number of dates in the bunch. In the current dataset, images of 256 individual dates of Sullaj, Barhi, and Meneifi dates were recorded.

FIGURE 14 shows 64 individual Barhi dates at the four maturity stages (Immature, Khalal, Rutab, and Tamar). We also calculated the average weight of Barhi and Sullaj dates at three maturity stages (Khalal, Rutab, and Tamar). For bunch stalk, the length and weight of ten stalks were measured.



FIGURE 14. 64-individual Barhi dates at the four maturity stages (16 dates per stage): immature, Khalal, Rutab, and Tamar.

3.3 COLLECTION PROTOCOL AND EQUIPMENT

The datasets were recorded in an orchard in Al-Ammaria, located 25 km Northwest of Riyadh, Saudi Arabia. The orchard contained more than 800 date palms of different varieties and ages.

The images of dataset-1 were taken during one season in six imaging sessions over the period of June–September 2016. The usual period between imaging sessions was 20 days, but because the Rutab stage has a short development time, we made the period during the Rutab stage 10 days (between sessions 2, 3, and 4). The images in dataset-1 were taken under different natural daylight conditions: in the morning (9:00-11:00) or afternoon (3:00-5:00). These images were captured using two RGB (Red, Green, and Blue) cameras: Canon, EOS-600D (camera-1) and EOS-1100D (camera-2), Tokyo, Japan, with variable focal length and in automatic mode. The resolutions of the cameras were 5184 x

3456 and 4272 x 2848, respectively. The orientation of the camera and the distance between the camera and the fruits depend on the height of the palm. Because the selected date palms were with different heights (1.7 – 6 meters), these parameters were variable. For short palms, the orientation was approximately 0° and the distance approximately three meters, but for tall palms, the orientation was from 45° to 90° and the distance was from three to six meters.

In dataset-2, the images and videos were acquired in the morning (5:00 – 7:00 am) during the harvesting period of Barhi dates (at the Khalal stage), 16 – 26 August. At the beginning of this period (16 August), each date bunch was marked as shown in **FIGURE 11**. Then, several images and videos were collected for date bunches from different angles. The images were captured using camera-1. The videos were recorded using an RGB camera (camera-3: Sony, HDR-CX405, Tokyo, Japan), in full HD resolution (1920x1080) and were stored in MPEG Transport Stream format (.MTS). Each video was recorded for all bunches on the palm as one continuous 360° video. The camera was moved around the palm in a circular manner at an approximate radius of three meters and a height of 1.5 meters from the ground, as shown in **FIGURE 15**. The orientation of the camera (the angle from the horizon) ranged from zero to 80 degrees depending on the palm heights. The characteristics of the selected palms were also registered, including palm height and trunk circumference. Palm height is the distance between the ground and the fruit, as shown in **FIGURE 15**. The circumference of the palm trunk was measured at a height of one meter from the ground.

Immediately after harvesting, the date bunches in dataset-2 were placed in front of the white graph paper, as in **FIGURE 10**, and their images were captured. Camera-1 was directed to the center of the date bunch and placed approximately one meter away from the bunch. The graph paper was 1.52 x 1.02 meters in dimensions. The dimensions of

the reference objects (black squares) were 50 x 50 mm. The distances between the reference objects are illustrated in **FIGURE 16**. Finally, the weights of these date bunches were recorded using a weight scale (Citizen, CTB 300, Metuchen, United States). The scale was placed on an even surface and date bunches were placed in a container and weighted, as shown in **FIGURE 17**.

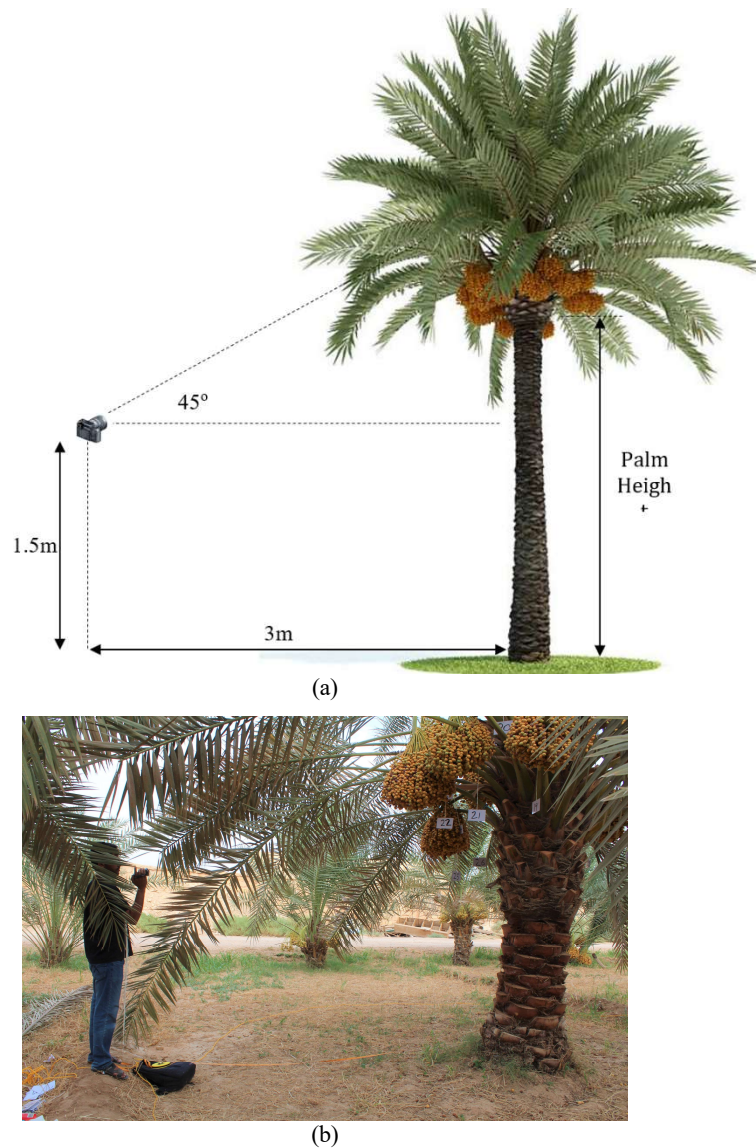


FIGURE 15. (a) Camera position setup for captured videos. The orientation of the camera depends on the palm height. Figure (a) shows a camera orientation of 45 degrees for a three-meter high palm tree. Figure (b) shows a real example for the palm-3 (B3.K.BW); the palm height is 2.85 meters, so the camera orientation was less than 45 degrees.

The images of individual dates were acquired under laboratory conditions at 18° C. For each image, 16 dates in one maturity stage were placed above a white reference paper

numbered from the top left side, as illustrated in **FIGURE 18**. Camera-1 was directed to the center of the reference paper at a distance of 320 mm, as shown in **FIGURE 19**. The uniform background was used to facilitate image segmentation, and the reference shapes at the paper corners can be used to estimate the dimensions and area of the dates. The dimensions of the two upper shapes were 30 x 20 mm and that of the bottom shape was 30 x 30 mm. The date information, i.e., date variety, maturity stage, harvesting date, and data code, were positioned in the bottom right of the paper. The average weights of individual dates were recorded at three maturity stages (Khalal, Rutab, and Tamar) based on around 200 dates for each stage using a Bizerba scale (model BS 200, Balingen, Germany). The average bunch stalk weight per centimeter was calculated by measuring the length and weight of ten bunch stalks. The physical length of the stalk (along the stalk curve) was measured from the inner end of the date bunch to the endpoint of the stalk, as in **FIGURE 20**. However, the endpoint of the stalk depends on the cutting position of the bunch, and it is not the ultimate end to the stalk.

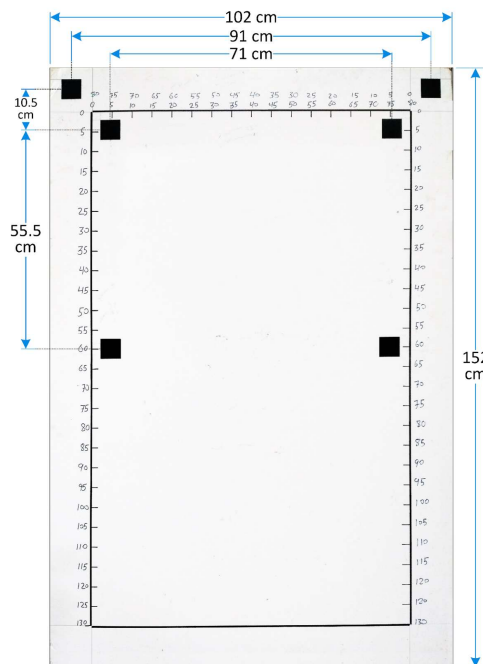


FIGURE 16. The Dimensions of the graph paper, showing the distances between the centers of the reference shapes.



FIGURE 17. Calculate the weight of a date bunch after harvesting. The bunch is placed in a container and weighed.

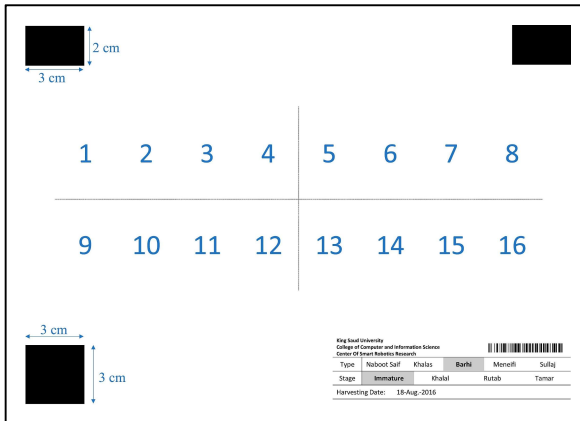


FIGURE 18. The template of reference paper used in the individual date imaging.



FIGURE 19. The camera position setup for the individual date imaging.

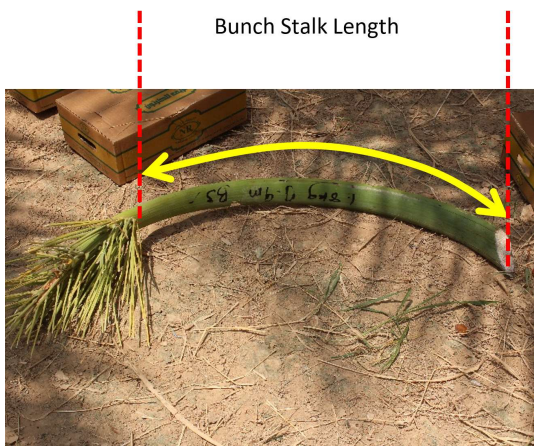


FIGURE 20. Date bunch stalk showing the ends we used to measure the length.

3.4 DATASET CODING

In this dataset, we arranged the data (dates, bunches, and palms and their related visual data, i.e., image/video files, or numerical data, e.g., weight), with a coding scheme to simplify referring, linking, and facilitating future extensions of the dataset.

The coding scheme consists of four parts separated by dots, as shown in **FIGURE 21**. The first part is the code for the palm tree, which consists of two parts: palm type (variety) and palm number. The palm number is omitted if the data is not associated with a specific palm (e.g. in the case of individual dates). The second part usage depends on the dataset. In dataset-2, this part consists of a character that refers to the maturity stage of the dates

(I: Immature, K: Khalal, R: Rutab, and T: Tamar). In the dataset-1, this part refers to the imaging session code, which consists of a session symbol [S] and session number, e.g., S1 refers to the first imaging session. The imaging session was used in dataset-1 instead of referring to the maturity stage explicitly because this dataset was collected over several periods of time during the maturity development of the fruit. Therefore, in one period (imaging session) the dates may have multiple maturity stages (some dates ripen before others). The third part contains one or more symbols that refer to the type of data, as stated in **FIGURE 21**. Multiple symbols in this part refer to multi-data, e.g., BW refers to a visual bunch data that also has a weight associated with it. The combination of these symbols should be in the same order as given in **FIGURE 21**, starting with visual data symbols (B, S, T) followed by numerical data symbols (W, D). The fourth part consists of a numerical value that indicates the data sequence. In the single date images, where the image contains 16 dates (**FIGURE 14**), this part refers to the range sequence, e.g., 1-16.

<u>Palm code</u> <u>Palm Type + Palm no.</u>	<u>Maturity Stage/ Imaging Session</u>	<u>Data type</u>	<u>Number</u>
<p>Palm types symbols:</p> <p>N: Naboot Saif K: Khalas B: Barhi M: Meneifi S: Sullaj</p> <p>The Palm number is omitted if the data is not associated with a specific palm, like in single dates.</p>	<p>Maturity stages symbols:</p> <p>I: Immature K: Khalal R: Rutab T: Tamar</p> <p>Imaging Session code: Session symbol [S] + Session no.</p>	<p>Data types symbols:</p> <p>B: Date bunch(s), visual data S: Single date(s), visual data T: Bunch stalk, visual data W: Weight, numerical data D: Dimension, numerical data</p> <p>Multi-symbols (at the same order above) indicate multi-data, e.g. : BW: bunch visual data associated with weight numerical data</p>	<p>A numerical value(s) (or range) to indicate the data sequence</p>

FIGURE 21. A description of the symbols used in the dataset coding.




3.5 GENERAL STATISTIC OF THE DATASET

In dataset-1, **FIGURE 22** gives details of the state of each palm during the six imaging sessions: harvested (whole palm bunches were cut), partially harvested (some individual Rutab dates were picked), or not harvested yet. The figure presents 29 date palms, 25

palms (5 palms per variety) were studied during all imaging sessions and four palms of Sullaj dates (S6, S7, S8, and S9) were inserted to the dataset starting from session four. This figure shows that some date palms were completely harvested before the Tamar stage (palms B3, B5, K2, M1, M2, M3, M4, and M5), and some date palms were gradually harvested over several sessions, by picking individual mature dates over several days, e.g. B1 and B2. **TABLE 2** shows the details of the number of images taken per session for all date varieties.

FIGURE 22. The state of each palm during the six imaging sessions: harvested, partially harvested, or not harvested.

Session no.	Date type / Session date	Barhi					Khalas					Meneifi					Naboot Saif					Sullaj								
		B1	B2	B3	B4	B5	K1	K2	K3	K4	K5	M1	M2	M3	M4	M5	N1	N2	N3	N4	N5	S1	S2	S3	S4	S5	S6	S7	S8	S9
1	29-6-2016																													
2	19-7-2016																													
3	2-8-2016																													
4	10-8-2016 13-8-2016*																					PH	PH	PH			*	*	*	*
5	3-9-2016	PH	PH			H									PH	PH	PH	PH	PH											
6	21-9-2016	PH	PH	H											H	H	H	H	H											

 The palm was not harvested yet.
 The palm was partially harvested (some dates were picked).
 The palm was harvested (whole palm bunches were cut).

* Four new Sullaj palms (S6, S7, S8, and S9) are inserted in the dataset (starting from session four) instead of the partially-harvested ones (S1, S2, and S3). After three days from session four, we captured images for the new palms (S6, S7, S8, and S9), and also for the two unharvested palms (S4 and S5).

TABLE 2. The number of images taken per session for all date varieties in the dataset-1.

Session no.	Session date	Barhi	Khalas	Meneifi	Naboot Saif	Sullaj	Total per session
1	29-6-2016	219	190	152	124	188	873
2	19-7-2016	193	161	183	245	262	1044
3	2-8-2016	330	256	298	156	351	1391
4	10-8-2016	719	425	445	401	530	2973
	13-8-2016	-	-	-	-	453	
5	3-9-2016	284	263	217	286	232	1282
6	21-9-2016	68	91	4	212	141	516
Total per date type		1813	1386	1299	1424	2157	8079

In dataset-2, the information of the date palms and their bunches are included in an excel file. This file consists of several tables. Each table provides information about one palm tree, which includes palm variety, code, height, trunk circumference, number of bunches, weight of bunches, harvesting date, and the state of the recording. **TABLE 3** presents overall information for all date palms in the dataset-2. Thirteen palms of the Barhi date were included in this dataset (152 bunches). Nine of them (palms 3 to 11) have full information about the palms and their bunches (all bunches in the palm). **TABLE 4** shows sample data for one palm (B5.K.BW).

In the current dataset, images of 256 individual dates of Sullaj, Barhi, and Meneifi dates were captured, as shown in **TABLE 5**. The average weights of individual dates were recorded for Sullaj and Barhi dates based on 1740 dates, **TABLE 6** presents samples of Barhi dates. The length and weight of the ten bunch stalks for Barhi dates were measured as presented in **TABLE 7**.

TABLE 3. List of all date palms in the dataset-2.

Palm no.	Palm code	Harvesting date	Total palm yield (kg)	Palm height (m)	Palm trunk circumference (m)	Total no. of bunches in the palm	Number of bunches				
							weighted	Recorded still image		Marked on palm	Recorded video
								On graph paper	On Palm		
1	B1.K.BW	16/8/2016	-	-	-	12	3	3	3	0	0
2	B2.K.BW	16/8/2016	373.5	-	-	15	15	15	14	0	0
3	B3.K.BW	18/8/2016	396.1	2.85	1.95	15	15	15	15	15	15
4	B4.K.BW	19/8/2016	222.2	2.6	1.77	12	12	12	12	12	12
5	B5.K.BW	21/8/2016	287.5	2.75	1.74	14	14	14	14	14	14
6	B6.K.BW	26/8/2016	361.0	3.7	1.98	15	15	15	15	15	15
7	B7.K.BW	20/8/2016	312	3.3	1.78	13	13	13	13	13	13
8	B8.K.BW	19/8/2016	307.8	3.45	1.72	14	14	14	14	14	14
9	B9.K.BW	20/8/2016	375.9	4.85	2.15	14	14	14	14	14	14
10	B10.K.BW	26/8/2016	203.0	1.72	1.85	11	11	11	11	11	11
11	B11.K.BW	21/8/2016	227.1	1.72	1.84	12	12	12	12	12	12
12	B12.K.BW	17/8/2016	-	-	-	-	4	4	2	0	0
13	B13.K.BW	18/8/2016	100.8	-	-	10	10	10	6	0	0
Total (Number of bunches)							152	152	145	120	120

TABLE 4. An example of data for a date palm in the dataset-2.

Palm type:	Barhi					
Palm code	B5.K.BW					
Palm height	2.75 m					
Palm trunk circumference	1.74 m					
No. of bunches in the palm	14					
Harvesting date	21/8/2016					

No.	Bunch code	Bunch weight (kg)	Recording state of the bunch			
			Image on palm	Image on graph paper	Video	Marked on palm
1	B5.K.BW.46	9.68	yes	yes	yes	yes
2	B5.K.BW.47	17.66	yes	yes	yes	yes
3	B5.K.BW.48	11.36	yes	yes	yes	yes
4	B5.K.BW.49	13.98	yes	yes	yes	yes
5	B5.K.BW.50	27.52	yes	yes	yes	yes
6	B5.K.BW.51	20.52	yes	yes	yes	yes
7	B5.K.BW.52	29.90	yes	yes	yes	yes
8	B5.K.BW.53	27.50	yes	yes	yes	yes
9	B5.K.BW.54	13.44	yes	yes	yes	yes
10	B5.K.BW.55	26.22	yes	yes	yes	yes
11	B5.K.BW.56	30.86	yes	yes	yes	yes
12	B5.K.BW.57	14.08	yes	yes	yes	yes
13	B5.K.BW.58	28.60	yes	yes	yes	yes
14	B5.K.BW.59	16.16	yes	yes	yes	yes

TABLE 5. The number of individual dates in dataset-2.

		Maturity stage				
		Immature	Khalal	Rutab	Tamar	Total
No. of dates	Sullaj	16	32	80*	32	160
	Barhi	16	16	16	16	64
	Meneifi	-	-	16	16	32

* The horizontal and vertical diameters were measured manually for 16 dates of this variety (S.R1.SD.17-32) using caliper.

TABLE 6. Samples of single date weights for the Barhi variety in three maturity stages: Khalal, Rutab, and Tamar.




Maturity Stage	Khalal	Rutab	Tamar
Harvesting date	18-Aug.-2016	3-Sep.-2016	18-Aug.-2016
No. of dates	180	160	200
Weight (kg)	1.855	1.885	1.345
image			

TABLE 7. The length and weight measurements of Barhi date bunch stalks.

Bunch stalk code: B.K.TWD.	1	2	3	4	5	6	7	8	9	10
Weight (kg)	2.90	2.42	1.80	1.56	1.84	1.40	3.44	2.78	2.96	3.64
Length (cm)	125	103	90	94	94	82	130	136	114	167
Weight per length (g/cm)	23.2	23.5	20	16.6	19.58	17.07	26.46	20.44	25.97	21.8

3.6 DATASET LABELING

Each image in the dataset-1 has two labels: one for type and one for maturity phase. The dataset-1 was labeled into five type classes and seven maturity classes (with three intermediate). **TABLE 8** shows the distribution of the dataset images between these classes. Type-based labeling was performed during dataset collection by an expert marking the selected date trees in the orchard. For maturity-based labeling, images were labeled according to the decision to harvest date bunches and maturity index of individual dates. The maturity index typically indicates when an individual fruit is ready to harvest. However, for date fruits, several aspects should be taken into consideration when making the harvesting decision. First, dates grow as bunches and the individual dates in one bunch do not mature uniformly. Second, dates are harvested by different methods: either by cutting the whole bunch (when most of the dates are ripe) or by selecting and picking individual ripe dates. Third, dates are harvested in different maturity stages (Khalal, Rutab, and Tamar). Based on these factors, we categorized the date bunches into seven classes (phases): Immature-1, Immature-2, pre-Khalal, Khalal, Khalal-with-Rutab, pre-Tamar, and Tamar. **FIGURE 23** shows sample images of date bunches labeled based on these maturity classes. These maturity classes are based on the color and texture of dates, as shown in **FIGURE 24**, as well as the harvesting decisions and methods described by experts and farmers, as demonstrated in **TABLE 9**. Hence, in

this study, maturity stages refer to the maturity status of date fruits (immature, Khalal, Rutab, and Tamar), whereas maturity phases refer to the seven maturity classes.

TABLE 8. Distribution of the dataset-1 images between the seven maturity classes and five type classes

	Type class					
	Barhi	Khalas	Meneifi	Naboot Saif	Sullaj	
1 Immature-1	412	232	332	332	261	1569
2 Immature-2	261	113	135	45	184	738
<i>intermediate-class-1</i>	300	24	176	190	22	712
3 Pre-Khalal	280	156	92	16	121	665
4 Khalal	187	90	83	-	299	659
<i>intermediate-class-2</i>	49	-	223	146	210	628
5 Khalal-with-Rutab	266	416	247	197	534	1660
<i>intermediate-class-3</i>	1	-	-	5	62	68
6 Pre-Tamar	42	263	-	281	109	695
7 Tamar	6	90	-	212	353	661
<i>no class</i>	7	1	7	-	2	17
	1811	1385	1295	1424	2157	8072

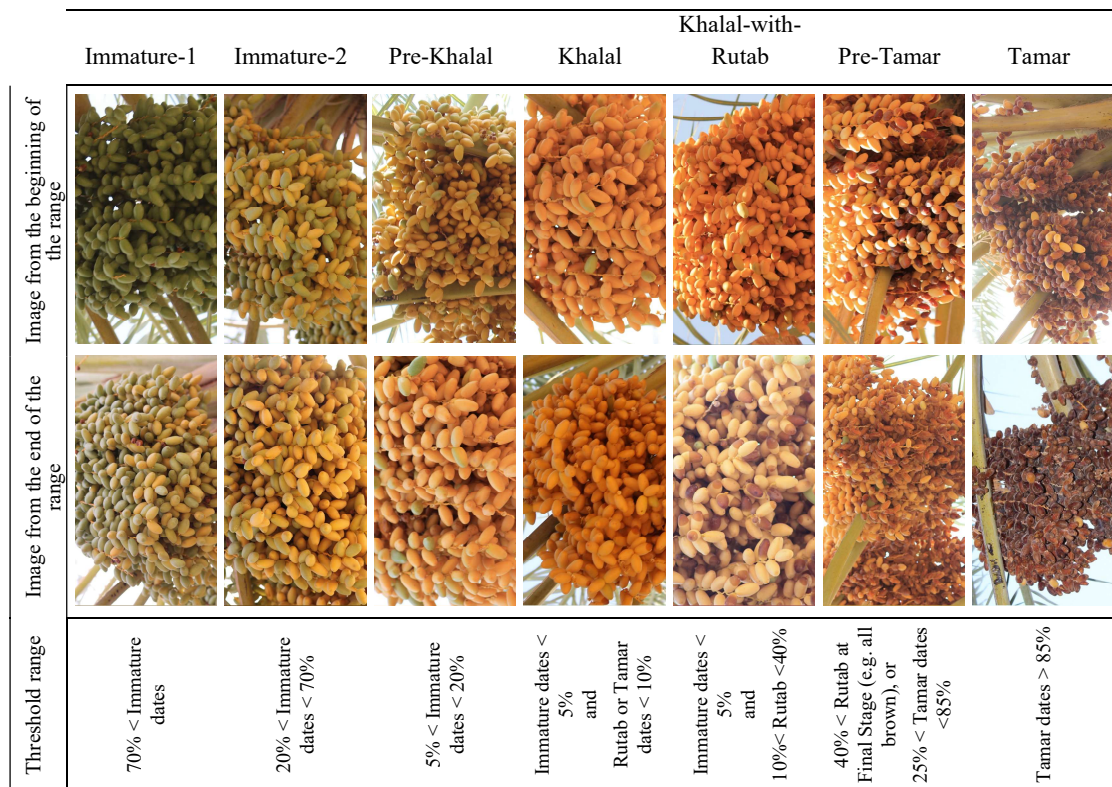


FIGURE 23. Sample images of Sullaj dates labeled into the seven maturity classes based on the threshold values described in TABLE 9.





















		Maturity stage				
		Immature stage (Green to greenish-yellow)	Khalal	Rutab*		Tamer
Date fruit type	Sullaj		yellow 	Golden yellow 		maroon 
	Barhi		yellow 	Apricot yellow 		Light brown 
	Khalas		Apricot yellow 	Yellowish brown 		Brown 
	Meneifi		yellow 	Apricot brown 		Light brown 
	Naboot Saif		yellow 	Golden 		Golden brown 

FIGURE 24. Sample images of individual dates in the four maturity stages of the five date types. The description of date's colors at each stage is according to [47].

TABLE 9. Description of the proposed maturity classes (phases).

Maturity phase/class	Harvesting decision	Description by experts	Approximate thresholding according to experts	More information about the maturity phase
Immature-1	No harvest action	Most dates are in the immature stage	70% < Immature dates	
Immature-2	No harvest action , but the farmers may pick individual Khalal dates	A large amount of immature dates	20% < Immature dates < 70%	Sometimes farmers start to collect individual mature Khalal dates in this phase (selective harvesting)
Pre-Khalal	(Transition phase) Usually no harvest action, but the farmer may harvest date bunches as Khalal	Most dates are in the Khalal stage, but some are in the immature stage	5% < Immature dates < 20%	Usually no harvest action in this phase, but sometimes farmers harvest the whole palm even if some bunches have immature dates, which are then removed after harvesting
Khalal	Date bunches are ready to be harvested as Khalal	Most dates are in the Khalal stage. A small amount of immature or Rutab dates is acceptable	Immature dates < 5% and Rutab or Tamar dates < 10%	Usually, Khalal dates are harvested by cutting the whole bunch (bunch-based harvesting)
Khalal-with-Rutab	Dates are ready to be harvested as Rutab , and date bunches can be harvested as Khalal	A sufficient amount of dates have reached the Rutab stage. A small amount of immature or Tamar dates is acceptable. A large amount of Khalal dates is acceptable	Immature dates < 5% and 10% < Rutab < 40%	Usually, Rutab dates are harvested individually (selective harvesting) by picking the ripe dates over several days. Alternatively, the whole bunch can be harvested and Rutab are separated from Khalal and each sold alone
Pre-Tamar	(Transition phase) Usually no harvest action, but farmers may pick individual Rutab or Tamar dates	A large amount of dates have reached the final Rutab stage (e.g. all brown) or Tamar stage	40% < Rutab at Final Stage (e.g. all brown) or 25% < Tamar dates < 85%	When date bunches reach this phase, usually they are left until the Tamar phase, i.e. when the majority of dates in the bunch are in the Tamar stage. However, sometimes farmers collect Rutab or Tamar dates individually (selective harvesting)
Tamar	Date bunches are ready to be harvested as Tamar	Most dates are in the Tamar stage. A small amount of Khalal or Rutab dates is acceptabl.	Tamar dates > 85%	In this phase, most dates are in the Tamar stage and they are usually harvested by cutting the whole bunch (bunch-based harvesting)

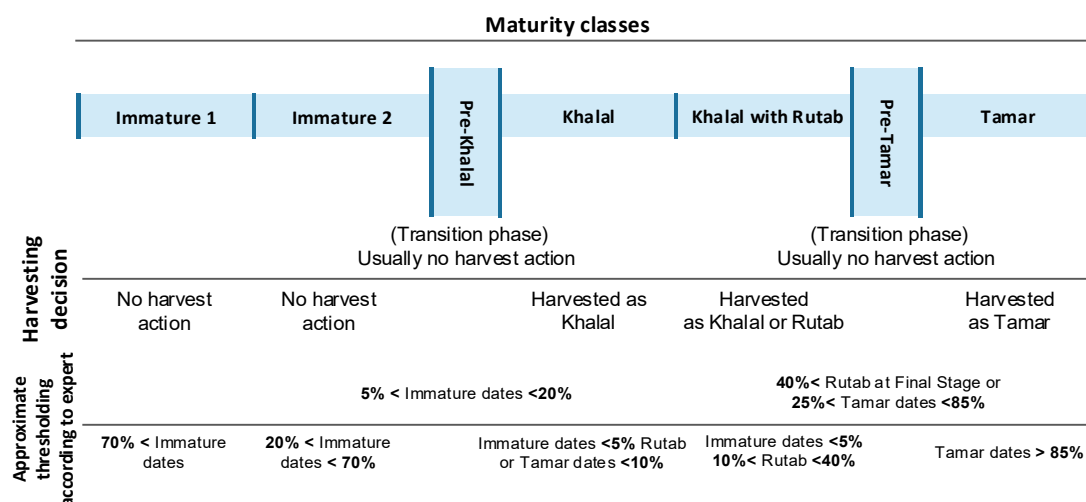


FIGURE 25. The proposed maturity classes of date bunches showing the harvesting decision for each class according to the bunch-based harvesting method, and the approximate thresholding based on the number of individual dates.

Since individual dates in the same a bunch do not mature uniformly, this lead to overlaps between the seven maturity classes, making it difficult to categorize all the images in the dataset into these seven classes. In some cases when date images have a high degree of confusion (e.g. an equal chance of being classified into two classes), we labeled them as an intermediate class between their two confused adjacent classes. Therefore, there are three intermediate classes between the seven main classes, as shown in **TABLE 8**. Intermediate classes were not used to train the system; however, during the testing phase, the images of these classes were be classified by the system into one of the confused adjacent classes. **FIGURE 26** shows how far images captured in each imaging session are distributed among the seven maturity classes.

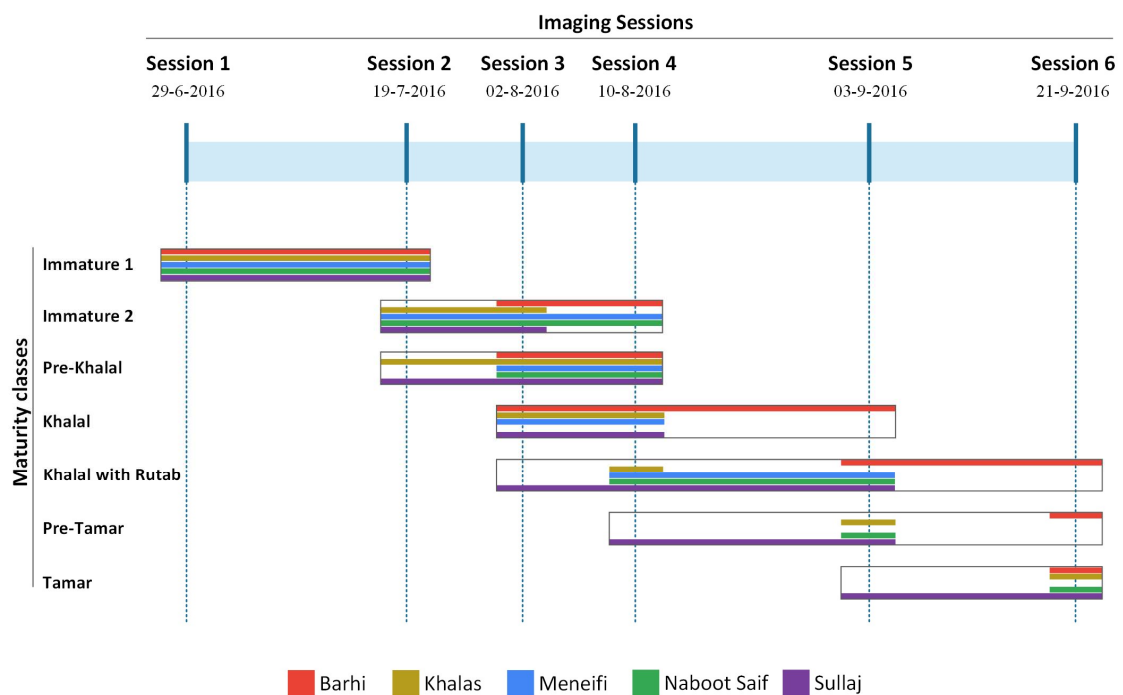


FIGURE 26. Distribution of the images captured at each imaging session between the seven maturity classes. Images captured in one imaging session are distributed between many classes. For instance, the images captured in session-4 are distributed between four classes for Meneifi dates (Immature-2, Pre-Khalal, Khalal, and Khalal-with-Rutab) and Sullaj dates (Pre-Khalal, Khalal, Khalal-with-Rutab, and Pre-Tamar).

The process of labeling the images between the maturity classes was performed visually based on the knowledge we have obtained through continuous consultation and

discussion with experts. We set an approximate thresholding for each class, as shown in **FIGURE 25** and described in **TABLE 9**, to make this process more accurate. The threshold values depended on the number of individual dates in bunches that belong to the four maturity stages, as shown in **FIGURE 27**. The threshold ranges were approximations based on information gathered from experts and farmers, with no distinct boundaries between the adjacent classes.



Immature dates = **1%**
 Khalal dates = **59%**
 Rutab dates = **39%**
 Tamar dates = **1%**

Since immature dates < 5%, and Rutab dates < 40%, this image is labeled in the **Khalal-with-Rutab** class



Immature dates = **3%**
 Khalal dates = **95%**
 Rutab dates = **2%**
 Tamar dates = **0%**

Since immature dates < 5%, and Rutab dates < 10%, this image is labeled in the **Khalal** class

FIGURE 27. Demonstration of the labeling process into the maturity classes. The process depended on the visual estimation of the number of individual dates in date bunches that belonged to the four maturity stages.

3.7 DIVIDING DATASET-1 INTO TRAINING AND TESTING SETS

Since the date dataset has several types in different maturity stages and the number of images per type and maturity stage differ, dividing the dataset randomly into training and testing sets will affect the classification accuracy due to bias. Therefore, we divided dataset-1 into training and testing sets based on rules to ensure the robustness of the classification models and sufficient distributed data to validate the system correctly:

1. The number of training images for all classes should be equal. This gives the classification model an equal chance of learning each class fairly.
2. At least 30% of the images should be included in the testing set for each date type as well as for each maturity stage or imaging session. This will help validate the models by testing them using enough images distributed over all date types, maturity stages, and imaging sessions.
3. The selected training images (after image equalization, rule 1) should be distributed equally between all imaging sessions, for the type classification model, and between all date types, for the maturity classification model. This will help build robust classification models independent of the maturity stage or date type.
4. In case of conflict, priority is given to the rules in sequence from rule 1 to rule 3

TABLE 10 shows the number of training and testing images after applying these rules. The training dataset for date type classification consisted of 4530 images (906 per class). As Meneifi had the lowest number of images, we took 30% for testing (389 images) and 70% for training (906 images). This number of training images was fixed for all the other types, and the rest of the images were used for testing, as shown in **TABLE 10-a**. Hence, the testing set amounted to 44% of the overall dataset (3542 images: Barhi 905, Khalas 479, Meneifi 389, Naboot Saif 518, and Sullaj 1251).

The same distribution rules applied to date maturity classification, which had 3227 training images, 461 per class (40% of overall dataset) distributed among all date types, as shown in **TABLE 10-b**. The remaining images were used as the testing dataset.

TABLE 10. Dataset-1 images in the training and testing sets. (a) The number of training and testing images for the date type classification task showing their distribution among the six imaging sessions. (b) The number of training and testing images for the date maturity classification task showing their distribution among the five date types.

(a) Date type classes

		Date type class (Training images: Testing images)						
		Barhi	Khalas	Meneifi	Naboot Saif	Sullaj		
Imaging session	1	29-6-2016	153 : 66	133 : 57	106 : 46	87 : 37	132 : 56	Total 4530 : 3542
	2	19-7-2016	135 : 58	113 : 48	128 : 55	172 : 73	171: 91	
	3	2-8-2016	190 :140	179 : 77	209 : 89	109 :47	171:180	
	4	10-8-2016	191 : 528	234 : 191	311: 134	195 : 206	171:812	
	5	3-9-2016	190 : 93	184 : 79	152 : 65	195 : 91	162 : 70	
	6	21-9-2016	47 : 20	63 : 27	-	148 : 64	99 : 42	
			906 : 905	906 : 479	906 : 389	906:518	906:1251	

(b) Date maturity classes

		Date type class (Training images: Testing images)					Total	
		Barhi	Khalas	Meneifi	Naboot Saif	Sullaj		
Date maturity class	1	Immature-1	93 : 319	92 : 140	92 : 240	92 : 240	92 : 169	461 : 1108
	2	Immature-2	128 : 133	79 : 34	95 : 40	32 : 13	127 : 57	461 : 277
	3	Pre-Khalal	192 : 88	109 : 47	64 : 28	11 : 5	85 : 36	461 : 204
	4	Khalal	131 : 56	63 : 27	58 : 25	-	209 : 90	461 : 198
	5	Khalal-with-Rutab	92 : 174	92 : 324	92 : 155	92 : 105	93 : 441	461 : 1199
	6	Pre-Tamar	29 : 13	178 : 85	-	178 : 103	76 : 33	461 : 234
	7	Tamar	4 : 2	63 : 27	-	148 : 64	246 : 107	461 : 200
								3227 : 3420

3.8 CONCLUSION

In this chapter, we presented a comprehensive dataset of date fruits for automated harvesting and visual yield estimation. To the best of our knowledge, this is the only publicly available dataset for date fruit pre-harvesting and harvesting applications. Our dataset consists of images, videos, and weights covering multiple sets of variations, which can be used for various machine vision applications relating to date fruits. It can be used in automated harvesting, fruit detection and segmentation, classification, maturity analysis, and weight/yield estimation. The dataset can help to advance automated harvesting research related to date fruits and thereby reduce dependence on manual methods. This would make date production more efficient and cost-effective.

The dataset is fully labeled according to type, maturity, and harvesting decision. We released the dataset with its associated files freely to the research community in the IEEE DataPort repository [1] (<http://dx.doi.org/10.21227/x46j-sk98>).

CHAPTER 4

DATE FRUIT CLASSIFICATION FRAMEWORK

FOR ROBOTIC HARVESTING

4.1 INTRODUCTION

An accurate vision system to classify and analyze fruits in real time is critical for harvesting robots to be cost-effective and efficient. However, practical success in this area is still limited, and to the best of our knowledge, there is no research in the area of machine vision for date fruits in an orchard environment. In this chapter, we propose an efficient machine vision framework for date fruit harvesting robots. The framework consists of three classification systems used to classify date fruit images in real time according to their type, maturity, and harvesting decision. In the classification systems, deep convolutional neural networks (CNNs) are utilized with transfer learning and fine-tuning on pre-trained models.

Referred paper:

H. Altaheri, M. Alsulaiman, and G. Muhammad, "Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning," IEEE Access, vol. 7, no. 1, pp. 117115–117133, Aug. 2019.

The models, codes, data, and test videos of the developed systems are available in the following website:

- Date variety classification system:
<https://sites.google.com/view/daterobotic/variety-classification>
- Date maturity classification system:
<https://sites.google.com/view/daterobotic/maturity-analysis>
- Date harvesting decision system:
<https://sites.google.com/view/daterobotic/harvesting-decision-system>

4.2 CHALLENGES OF DATE FRUIT CLASSIFICATION IN A NATURAL ENVIRONMENT

Automatic date fruit classification in an orchard environment is a difficult task. In the case of type classification, some date types can easily be distinguished based on their visual appearance, whereas it is difficult, for even a specialist, to distinguish some types, as shown in **FIGURE 28**. In addition, there are large differences between date bunches of the same type, as they vary in terms of maturity level, illumination, bagging state, scale, and angle (**FIGURE 7, 8, and 9**). Maturity classification of date fruits is also considered to be challenging because dates grow in large clusters (bunches) and can be harvested in different maturity stages that overlap because the dates in one bunch do not mature at the same time. This fact makes them hard to classify or even label by an expert, as shown in **FIGURE 29**. Furthermore, many external effects make classification based on maturity difficult. This includes images of date bunches covered by bags and images with poor illumination. All these variations were included in the dataset to help build a robust vision system that provides accurate results in natural environments.

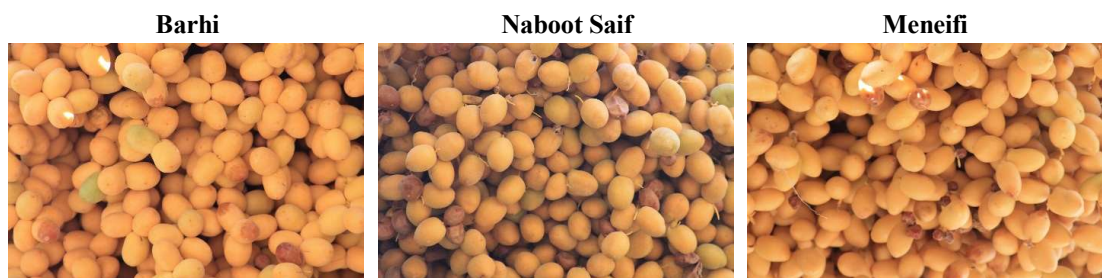


FIGURE 28. Sample images of three date types at the Khalal stage that are difficult to distinguish.



FIGURE 29. Sample images of date bunches that have individual dates at different maturity levels, which makes them hard to label or classify.

4.3 PROPOSED FRAMEWORK

The proposed framework consists of three classification systems for type, maturity, and harvesting decision. The input to the framework is a stream of frames (images) from an RGB video camera in a date orchard. The framework has three outputs to determine the type, the maturity stage, and the harvesting decision of dates in each image. Type and maturity models are multiclass classifiers that use fine-tuning with transfer learning based on pre-trained CNN models. We investigated two popular and widely used deep learning CNN architectures: AlexNet [48] and VGGNet [49]. A harvesting decision model, on the contrary, is a binary classifier that uses the output of the two previous models with the harvesting rules, which are entered manually by a user, and then suggests the appropriate harvesting decision based on all the previous information. The block diagram in [FIGURE 30](#) provides an overview of the proposed framework in the training phase (a) and real-time implementation (b).

4.3.1 Convolutional Neural Network Deep Learning

Deep learning has become the dominant approach in many computer vision tasks including object detection, recognition, and classification [50], [51]. Deep learning achieves high levels of success in these tasks due to the availability of a large number of labeled images, such as ImageNet [52], and substantial computing power devices such as GPUs or distributed large-scale clusters using cloud computing [53]. The success of deep learning also goes beyond images and video to speech and audio. Earlier, its success was restricted by its need for large databases and long training times. However, these problems have been solved using transfer learning and fine-tuning techniques.

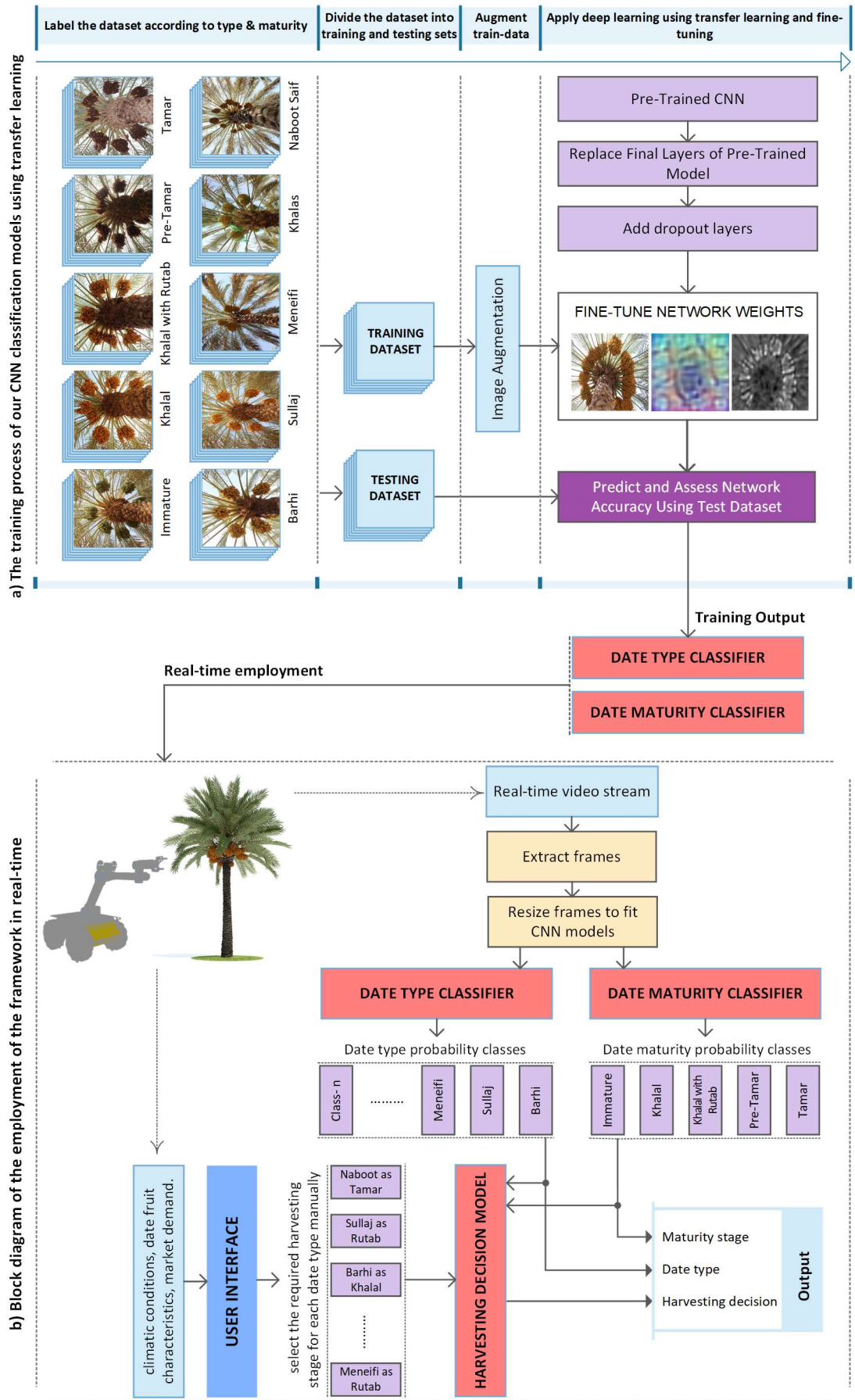


FIGURE 30. Block diagram of the proposed date fruit classification framework. (a) Training phase, (b) real-time employment.

A CNN is one of the most successful types of deep learning. It uses deep convolutional networks and non-linearity to learn local and spatial features and patterns directly from raw data such as images, video, text, or sound. A CNN thus learns the features from data automatically, eliminating the need to extract them manually. It can build complex features as an integration of simple features through a series of successive convolutional layers. The earlier layers learn low-level features such as edges and curves and the deeper layers learn to recognize complex high-level features such as entire objects in an image [50].

CNN models can be trained from scratch (with random initialization) or using transfer learning. Training from scratch needs a large dataset to learn millions of parameters in CNN models. In many tasks such as fruit classification, it is rare for CNNs to learn from scratch due to the limitation of a labeled dataset of sufficient size. Instead, it is common to pre-train a CNN using a large-scale dataset (e.g. ImageNet, which has 1.2 million images with 1000 classes), and then use the CNN model either as a fixed feature extractor or as an initialization for other specific tasks. In transfer learning, the CNN learns generic mid- and low-level features from a large dataset that can be fine-tuned by other target datasets [54]. Therefore, in this work, two CNN architectures, well-trained over ImageNet dataset [52], are utilized and fine-tuned for the date fruit dataset. This approach learns features automatically with high accuracy and short training times.

4.3.2 AlexNet Architecture

The AlexNet model was the first deep architecture that popularized convolutional networks in computer vision and significantly improved the classification accuracy of the ImageNet ILSVRC2012 challenge compared with the traditional methods. The AlexNet architecture (**FIGURE 31**) consists of eight learnable layers, five convolutions (conv) layers followed by three fully connected (FC) layers. Every convolutional and fully connected layer is attached by rectified linear units (ReLU) as non-linear activations. Adding non-linearity using ReLU helps CNNs train much faster. The first and second convolutional layers are followed by both local response normalization (LRN) and max-pooling layers, but only max-pooling layer is used after the fifth convolutional layer. The first convolutional layer has 96 kernels (filters) of dimension 11×11 with a step (stride) of four pixels. The strides of the remaining convolutional layers are set to one pixel. The second layer has 256 kernels of dimension 5×5 . The third, fourth, and fifth layers have 384, 384, and 256 kernels of dimension 3×3 , respectively. Max-pooling layers use non-linear down-sampling to abstract the network. They thus retain the most important features and reduce the number of parameters that the network needs to learn, which diminishes network computation. The first two fully connected layers have 4096 neurons and the last one has 1000 neurons, which is the same as in the ImageNet classes. The neurons in fully connected layers are connected to all the neurons in the previous layer. The last layer of the model provides the classification output using the softmax classifier.

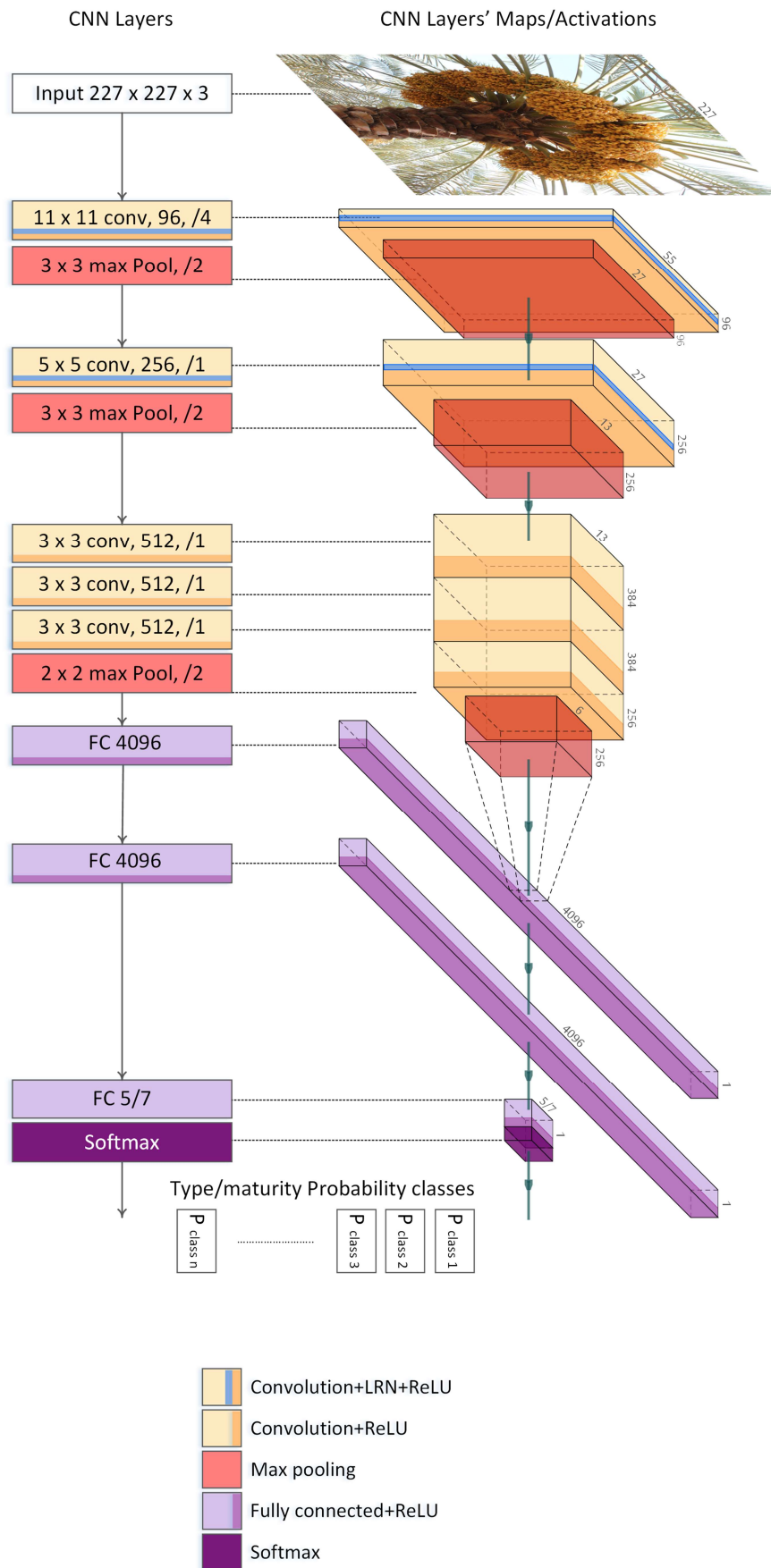


FIGURE 31. Deep learning architectures of the proposed date fruit classification framework based on AlexNet pre-trained model.

4.3.3 VGGNet Architecture

VGG-16, which is the D-configuration in [49], also known as VGG-D, is the best network configuration among the different VGGNets built by the VGG group [49]. VGG-16 (VGGNet) achieved 92.7% top-five accuracy in the ImageNet ILSVRC2014 challenge. Its architecture is popular in the literature due to its uniform structure and high accuracy in classification tasks. VGG-16 has a deeper network than AlexNet and consists of 16 trainable layers including 13 convolutional layers and three fully connected layers (FIGURE 32). The model features a homogeneous and smooth architecture that only uses filters of size 3×3 with a stride of one for convolutions and 2×2 pooling with a stride of two in all layers. The convolutional layers are grouped into five blocks. Adjoining blocks are linked through a max-pooling layer, which performs down-sampling by half along the spatial dimensions. Max-pooling reduces the dimensions of the layers from 224×224 in the first block to 7×7 after the last one. The number of convolution filters remains fixed within one block and doubles after each max-pooling layer from 64 in the first block to 512 in the last block. As in AlexNet, the ReLU layer is tailed after each convolutional and fully connected layer. The fully connected layers in VGG-16 have the same configurations as in AlexNet. VGG-16 does not use normalization layers due to their minimal contribution to enhancing accuracy.

4.3.4 Transfer Learning

Transfer learning makes CNNs work very effectively in many visual classification tasks, even if their datasets have insufficient or limited size [55]. In transfer learning and fine-tuning techniques, the CNN network of a pre-trained model, trained on a large-scale dataset, can be used either as a feature extractor [56] or as a weight initializer rather than generating weight values randomly [57]. The work in [55], discussed the best practices for transfer learning with fine-tuning.

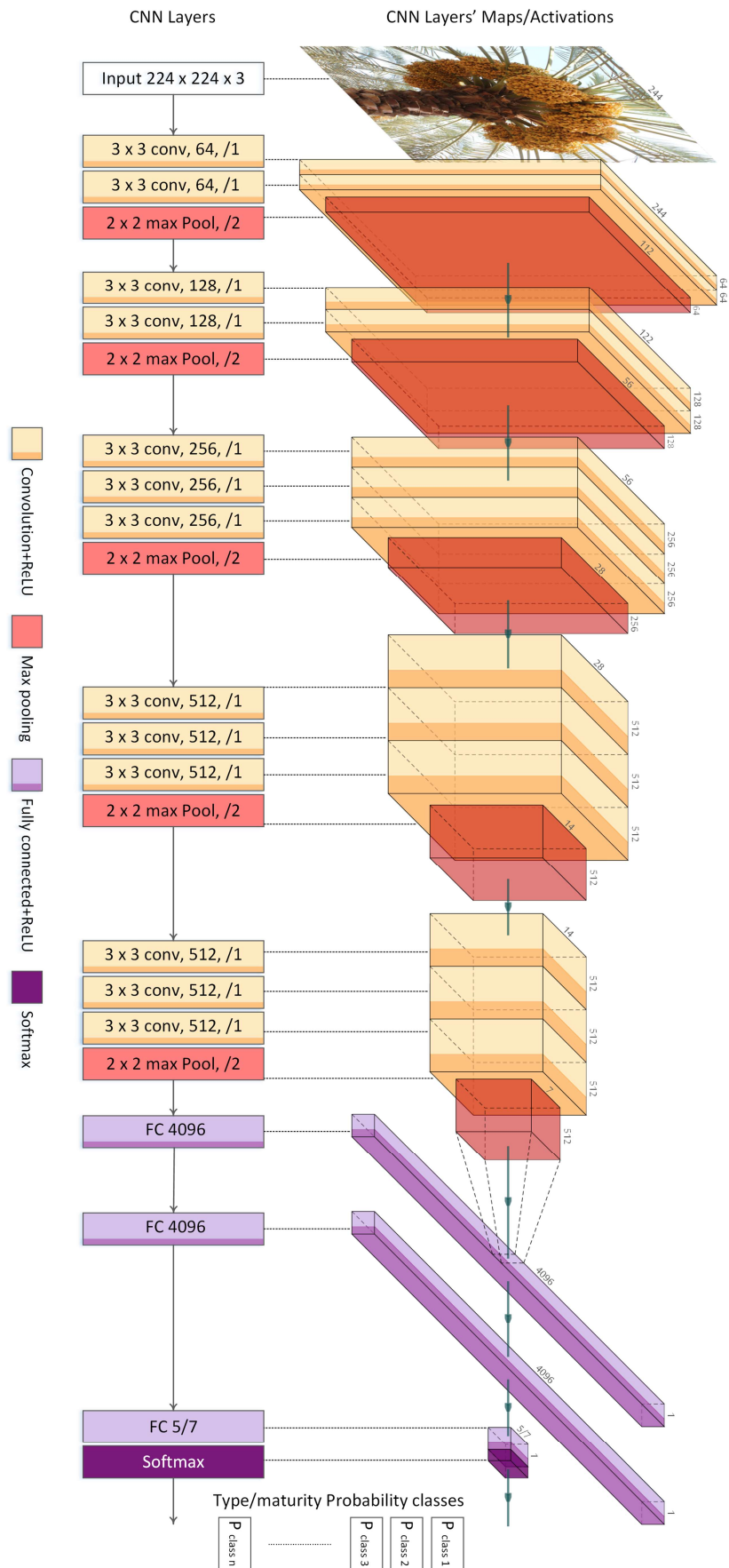


FIGURE 32. Deep learning architectures of the proposed date fruit classification framework based on VGG-16 pre-trained model.

CNN models have millions of trainable parameters; for example, AlexNet has 60 million and VGG-16 has 138 million. Training these parameters from scratch is difficult using relatively small datasets such as the one in this work. CNN networks can easily memorize small datasets which lead to overfitting. Hence, to prevent overfitting, transfer learning was used in this study. Moreover, the size of the dataset was increased using image augmentation performed on the training datasets randomly using different values. During the training phase, the data images in each batch were randomly subjected to the following operations: horizontal reflection, horizontal and vertical translation with a random value in the range $[-30\ 30]$ pixels, and horizontal and vertical scaling with a random rate in the range $[0.9\ 1.1]$.

We used transfer learning with fine-tuning as demonstrated in [58]. In the fine-tuning technique, the weights of the CNN models were trained starting from the transferred values for all the layers excluding the last fully connected layer. Because this layer was trained to classify 1000 categories (ImageNet classes), we replaced it with a new one to classify the new tasks classes, as shown in [FIGURE 31](#) and [FIGURE 32](#). The weights of the new layer were initialized with random values and its learning rates were raised compared with the rest of the CNN model. In fine-tuning, learning rates are reduced for original layers and boosted for new layers. This changes the CNN model by only a small amount, as the weights were optimized using a large dataset and only need to be modified slightly. However, the new layers change their weights quickly and therefore learn much faster. We also used dropout layers with a 0.5 probability after the first two fully connected layers. The dropout layer generalizes the network and prevents overfitting [59]. At the end of the last fully connected layer, a softmax layer was used to produce the classification outputs. The images in the training and testing sets were resized to $227 \times 227 \times 3$ and $224 \times 224 \times 3$ before feeding them into the AlexNet and

VGG-16 networks, respectively. Finally, the models were trained using the training datasets.

We experimented with different hyper-parameter settings to make the CNN models generalize well. For both CNN models (AlexNet and VGG-16), the learning rates of the original and new layers were adjusted to 0.0001 and 0.002, respectively. The weights of the last fully connected layer were initialized using a Gaussian distribution with zero mean and 0.01 standard deviation. The weights of the remaining layers were initialized using the pre-trained models. A stochastic gradient descent optimizer with a momentum of 0.9 was used to train the CNN models. A weight decay (regularization) of 0.0001 was added to the cross-entropy loss function to help reduce overfitting.

4.4 RESULTS AND DISCUSSIONS

We performed many experiments for the three classification tasks. In all the experiments, the CNN models were trained by one GPU, Nvidia GeForce GTX-1060 6 GB, with Intel Xeon E5-2600 CPU and 28 GB RAM, using Matlab2018b. The batch sizes were set to 128 and 32 training images for the AlexNet and VGG-16 models, respectively. The batch sizes were set depending on the hardware capabilities (GPU) and depth of CNN models used (number of trainable parameters). The improvement of the CNN models was tracked during the training phase by testing them on the testing datasets every 50 iterations. Using the testing datasets, the proposed systems achieved final accuracies of 99.01%, 97.25%, and 98.59% for date type classification, date maturity classification, and harvesting decision, respectively.

4.4.1 Performance Evaluation

The performance of the classification models was evaluated using speed and accuracy measures. The speed of the model was measured by the classification rate, number of classified frames (images) per second (fps), and the average classification time (i.e. the average time spent by the model classifying one frame (seconds/frame)). The accuracy of the model was evaluated using the positive predictive value (PPV) or precision, true positive rate (TPR) or recall, *f*score, and accuracy (ACC). These measures were calculated by Equations (1), (2), (3), and (4), respectively.

$$PPV_x = \frac{TP_x}{\text{TotalPredicted}_x} \quad (1)$$

$$TPR_x = \frac{TP_x}{\text{TotalActual}_x} \quad (2)$$

$$f\text{score}_x = \frac{1}{\alpha/TPR_x + \alpha/PPV_x} \quad (3)$$

where: $\alpha = 0.5$ gives equal weight to TPR and PPV

$$ACC = \frac{\sum_{i=1}^n TP_i / I_i}{n} \quad (4)$$

where: $n = \text{no. of classes}$, $I_i = \text{no. of images in classe } i$

For class x , if TP_x is the true positive (i.e. the number of images correctly predicted as belonging to class x), the PPV is the number of true positives divided by the total number of images predicted as belonging to class x . The TPR is defined as the number of true positives divided by the actual number of images in class x . The *f*-score is used to combine PPV and TPR into one measure using the harmonic mean. The overall accuracy in Eq. (4) was calculated using balanced accuracy, which normalizes the true positive for each class by the number of images in the class and divides their sum by the number of classes. Balanced accuracy ensures that all classes contribute equally to

the calculation of overall accuracy even if the number of samples in the classes is unequal. In our experiments, we demonstrated the confusion matrices of the classification models using the same format as in TABLE 11.

TABLE 11. The confusion matrix for the multiclass classification problem, with classes C_1, C_2, \dots, C_n . TP, TPR, and PPV refer to the true positive, true positive rate, and positive predictive value, respectively

		Predicted class					
		C_1	C_2	C_n	
Actual class	C_1	TP ₁					TPR ₁
	C_2		TP ₂				TPR ₂

	C_n					TP _n	TPR _n
		PPV ₁	PPV ₂	PPV _n	ACC

4.4.2 Visualization Results of the CNN models

FIGURE 33 presents some of the visualization results for the fine-tuned AlexNet and VGG-16 CNN models. The first row in the figure (a to d) visualizes the first 30 features (filters) of the first and last convolutional layers for AlexNet and VGG-16, respectively. The second row visualizes their 30 activations. The filters in the first convolutional layers (a and c) mostly contain colors and edges, which indicates that they are color filters and edge detectors. In the last layers (b and d), the filters represent more detailed and compacted features. We can also see that the filters in all the layers are well formed, smooth, and have no noisy patterns, indicating a well-converged network. Noisy patterns usually indicate insufficient training time or a very low regularization, which may lead to overfitting. The corresponding activations (feature maps of these 30 features) are shown in the bottom (e, f, g, and h). In these feature maps, the light (white) pixels in some location indicate strong positive activations at that position, whereas the

dark (black) pixels indicate strong negative activations. Mostly gray feature maps indicate weak activation on the input image. Since we employ ReLUs, the initial activations appear relatively dense; however, as the training progresses, they look more localized and sparser.

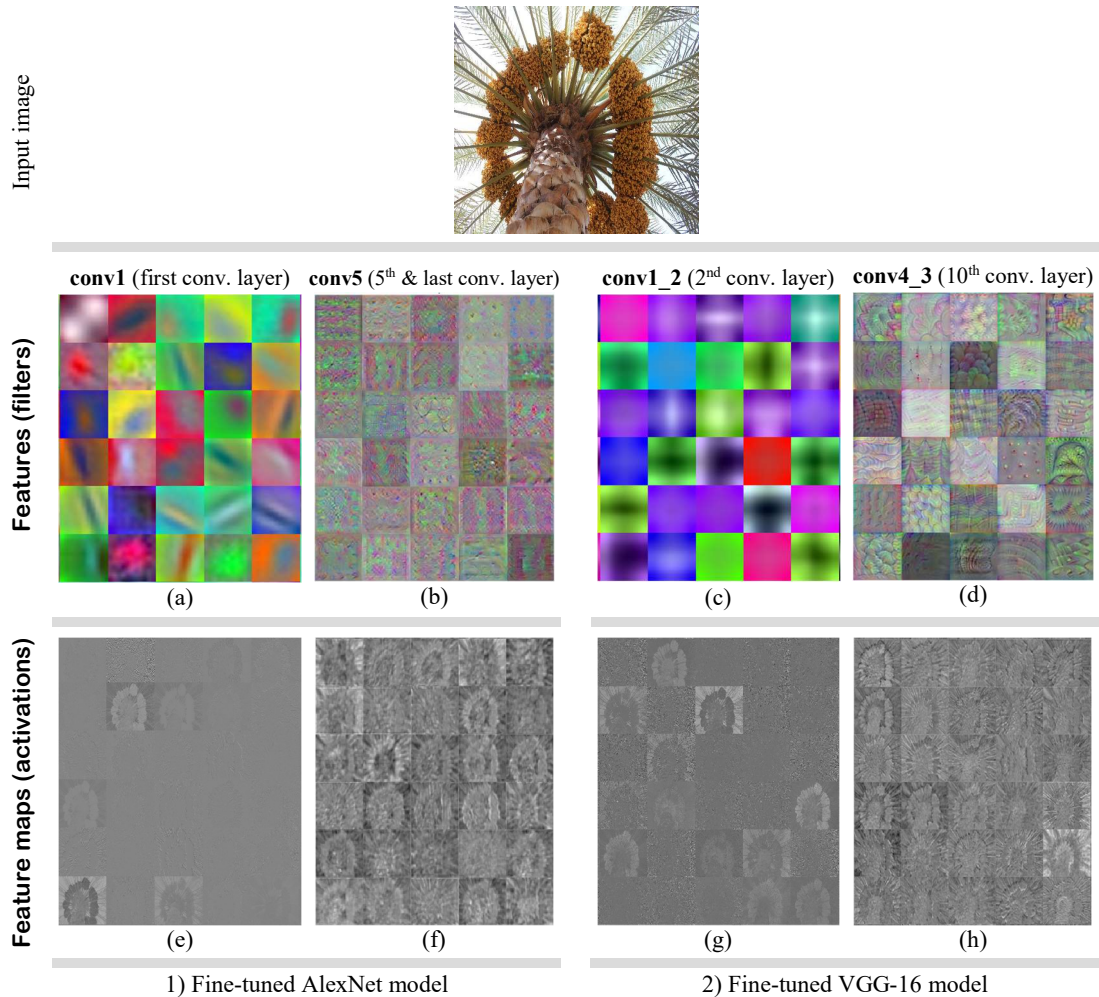


FIGURE 33. Visualization of the type classification models based on AlexNet (1) and VGG-16 (2). The figure shows first 36 features (filters) of the first and last convolutional layers in the top, and the activations of these features are shown in the bottom.

4.4.3 Date Type Classification System

In both the pre-trained CNN networks, the last fully connected layer was set to five neurons, equal to the number of date type classes. The AlexNet and VGG-16 CNN networks were trained using 4530 training images (TABLE 10-a) for 7000 iterations. The VGG-16 model needs more iterations than AlexNet to pass all the training examples in

the dataset because it has more learnable parameters. In 7000 iterations, VGG-16 and AlexNet adjusted their weights using all the training images during 50 and 200 epochs, respectively.

After fine-tuning, the CNN models were tested on the testing dataset containing 3542 images (TABLE 10_a) and they achieved accuracies of 96.51% and 99.01% for AlexNet and VGG-16 models, respectively. The VGG-16 model achieved better accuracy than AlexNet due to its deeper network. The multiple stacked smaller filters in VGG-16 are better than the larger ones in AlexNet because a higher number of non-linear layers increases the network depth. The deep architecture of VGG-16 enables it to learn more complex features and therefore it is more accurate. The training and validation accuracies and losses during the training phase are shown in FIGURE 34. The losses of the pre-trained models decrease slowly. After around 4000 iterations, both the CNN models reach almost their highest accuracies and stabilize in further iterations with little improvement. The performance of the fine-tuned CNN models on the testing dataset is illustrated by the confusion matrices in TABLE 12 and summarized in TABLE 13.

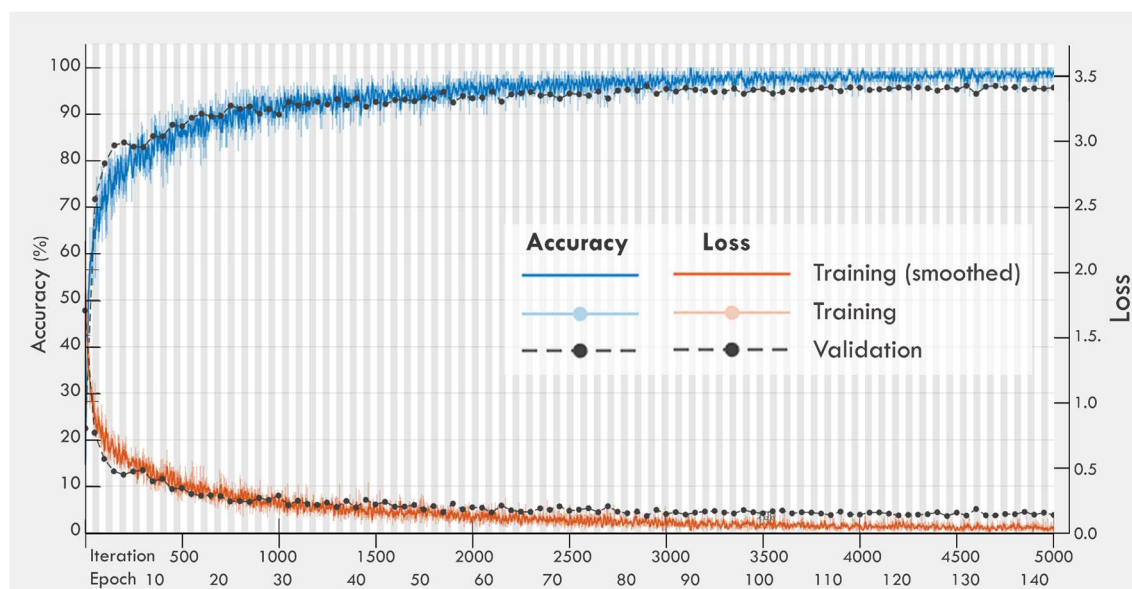
TABLE 12. The confusion matrices (as illustrated in TABLE 11) of the type classification models on the testing dataset.

		(a) Predicted class based on AlexNet					(b) Predicted class based on VGG-16						
		Barhi	Khalas	Meneifi	Naboot Saif	Sullaj	Barhi	Khalas	Meneifi	Naboot Saif	Sullaj		
Actual class	Barhi	881	2	6	11	5	97.3	896	1	4	2	2	99.0
	Khalas	2	473	2	1	1	98.7	0	478	1	0	0	99.8
	Meneifi	14	1	367	3	4	94.3	2	5	381	1	0	97.9
	Naboot Saif	9	1	4	500	4	96.5	5	0	0	513	0	99.0
	Sullaj	19	2	24	10	1196	95.6	3	0	6	0	1242	99.3
		95.2	98.7	91.1	95.2	98.8	96.5	98.9	98.8	97.2	99.4	99.8	99.0

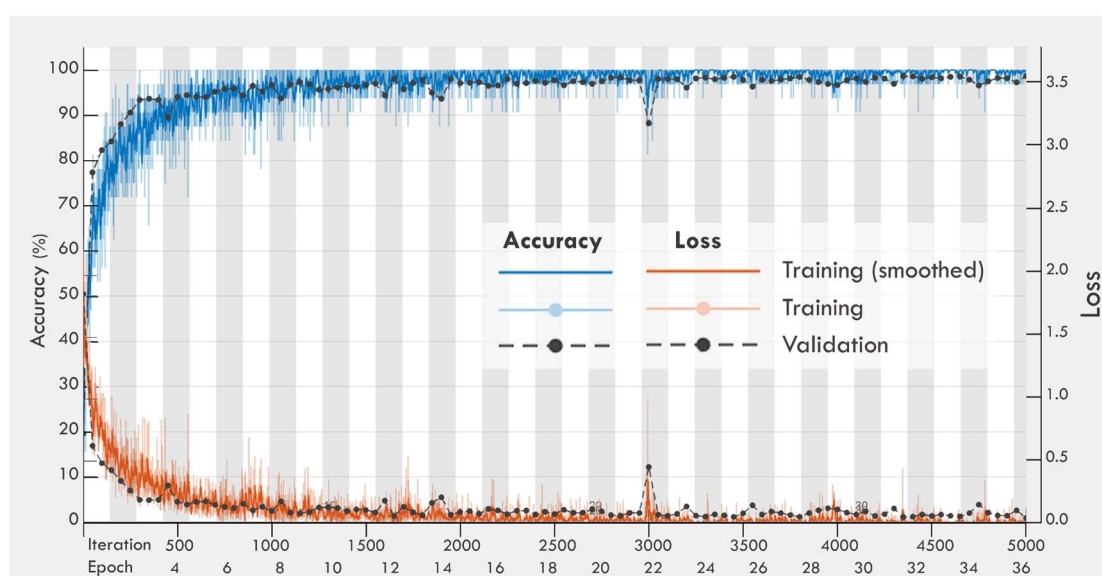
TABLE 13. Results of the proposed date type classification system on the testing dataset.

Method	Accuracy (%)	PPV (%)	TPR (%)	<i>f</i> score (%)	Classification rate (fps)	
					GPU*	CPU*
Transfer learning based on VGG-16	99.01	98.82	99.01	98.92	48.33	1.35
Transfer learning based on AlexNet	96.51	95.83	96.51	96.17	83.96	17.65

* GPU: Nvidia GeForce GTX 1060 6 GB, CPU: Intel Xeon E5-2600 with 28 GB RAM.



(a) AlexNet



(b) VGG-16

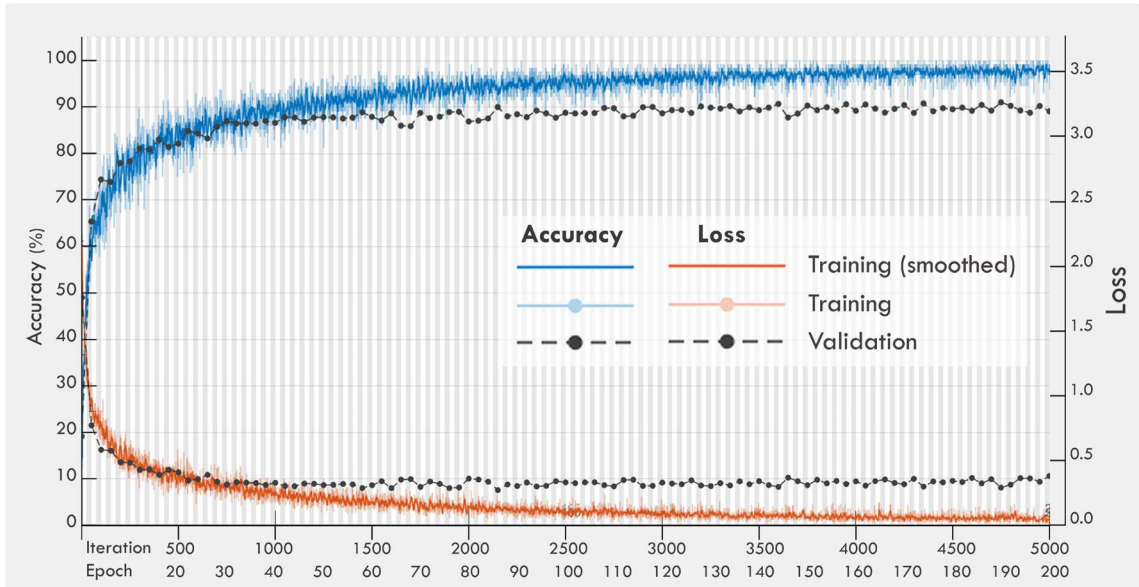
FIGURE 34. The behavior of the training and validation accuracies and losses of the date type classification models during the training phase.

4.4.4 Date Maturity Classification System

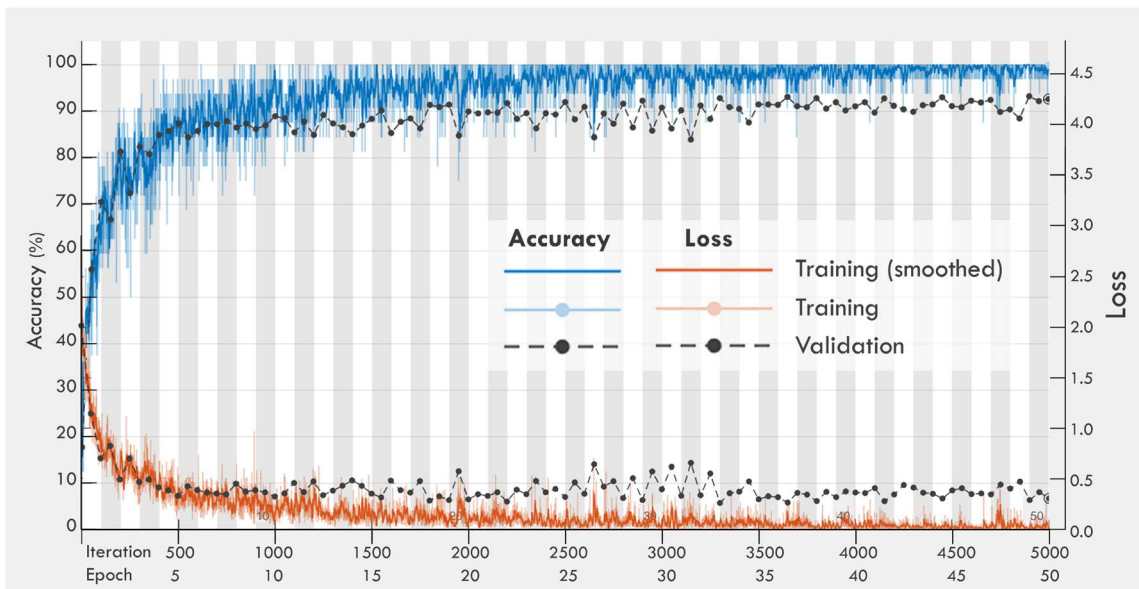
The date dataset was labeled into seven maturity classes according to the maturity state of the date fruit bunches, as illustrated in the dataset labeling subsection. These maturity classes represent the four maturity stages of date fruits (immature, Khalal, Rutab, and Tamar). The objective of the maturity classification model was to classify the images of date fruit bunches into the four maturity stages using the proposed seven classes. We conducted many experiments to determine the best selection of maturity classes and their training examples to train the CNN networks. Here, we present three experiments for training the CNN models using three selections of the maturity classes.

1 - In the first experiment, the CNN models were trained on the seven maturity classes using 3227 training images (TABLE 10-b) for 5000 iterations. The CNN models achieved accuracies of 90.1% and 92.3% on the testing dataset containing 3420 images for the AlexNet and VGG-16 models, respectively. FIGURE 35 shows that both the pre-trained models reached almost their highest accuracies after around 3000 iterations and did not show significant improvement in further iterations.

The achieved accuracies are relatively low because adjacent classes significantly overlap, which is clearly noticeable in FIGURE 23. The confusion matrices in TABLE 14 show that in both CNN models, most of the misclassified images were predicted between the prior or subsequent classes of the actual class. For example, as shown in TABLE 14-b, 13 of 14 Khalal sample images were predicted either as pre-Khalal (prior class) or Khalal-with-Rutab (subsequent class). For the images misclassified as Tamar (last columns in TABLE 14), most had poor illumination, as shown in FIGURE 36. This can be justified by the fact that dark images lost most of their color features and thus were closer to the features of the Tamar class that contain dates with dark colors (e.g. dark brown).



(a) AlexNet



(b) VGG-16

FIGURE 35. The behavior of training and validation accuracies and losses of the date maturity classification models based on the seven maturity classes.

Due to the small inter-class variation in the adjacent classes, it is difficult for the CNN models to learn distinct features for these classes. Therefore, to increase the accuracy of the maturity classification model, we need to train the CNN models with fewer number of maturity classes that have more distinct features and directly represent the required maturity stages of date fruits. This will help the classification model learn the

clear and strong features related to date maturity stages. Then, in the test process, it will automatically classify other confused images to the appropriate nearest classes. This was done in the next experiment.

TABLE 14. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the seven maturity classes.

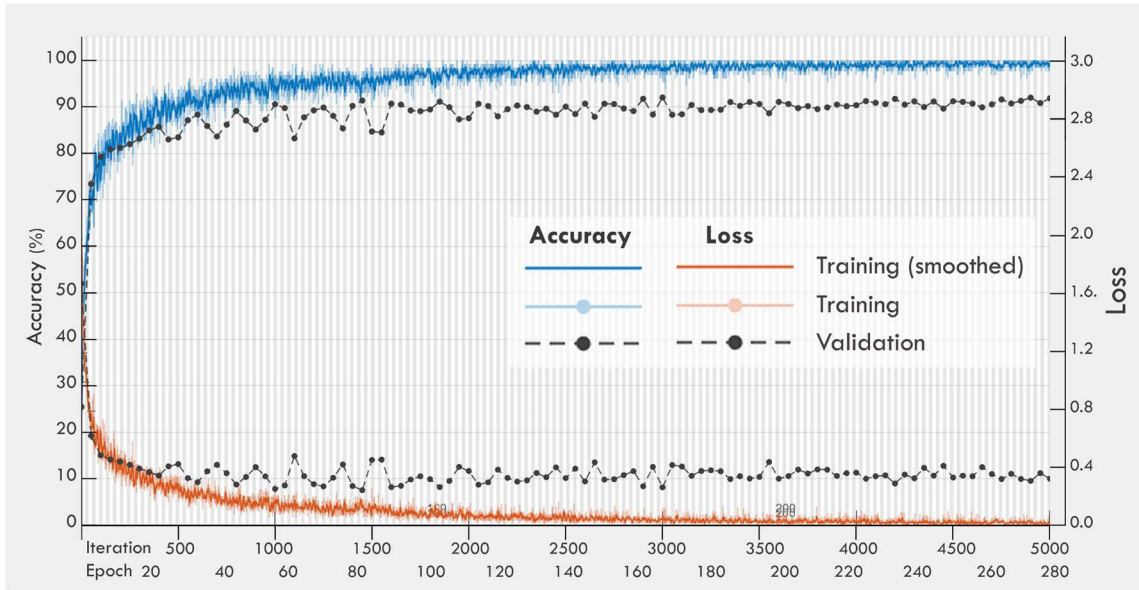
		(a) Predicted class based on AlexNet							
		Immature-1	Immature-2	Pre-Khalal	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar	
Actual class	Immature-1	1063	34	0	1	1	0	9	95.9
	Immature-2	5	266	6	0	0	0	0	96.0
	Pre-Khalal	0	7	165	12	17	0	3	80.9
	Khalal	0	3	12	163	15	4	1	82.3
	Khalal-with-Rutab	0	5	41	31	956	149	17	79.7
	Pre-Tamar	0	0	0	0	1	229	4	97.9
	Tamar	1	0	0	0	0	3	196	98.0
		99.4	84.4	73.7	78.7	96.6	59.5	85.2	90.11

		(b) Predicted class based on VGG-16							
		Immature-1	Immature-2	Pre-Khalal	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar	
Actual class	Immature-1	1086	13	0	2	1	0	6	98.0
	Immature-2	11	251	13	1	1	0	0	90.6
	Pre-Khalal	0	1	165	27	9	0	2	80.9
	Khalal	0	0	8	184	5	0	1	92.9
	Khalal-with-Rutab	0	0	18	46	1056	74	5	88.1
	Pre-Tamar	0	0	0	1	0	226	7	96.6
	Tamar	0	0	0	0	0	2	198	99.0
		99.0	94.7	80.9	70.5	98.5	74.8	90.4	92.3

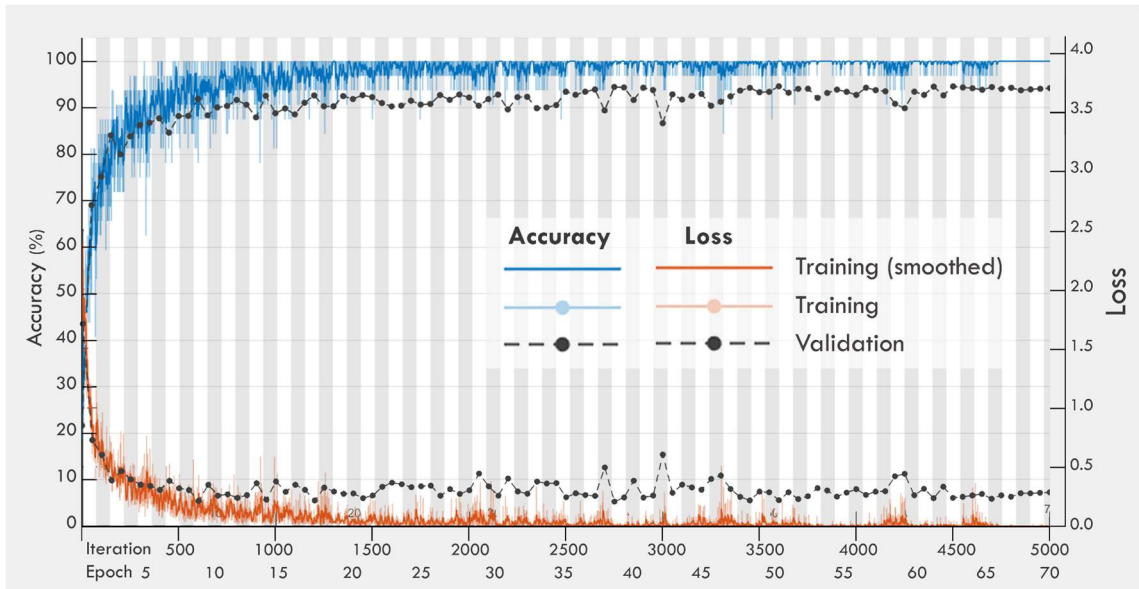


FIGURE 36. Samples of misclassified images using the maturity classification CNN models on the seven maturity classes.

2 - In the second experiment, we excluded the immature-2 class and merged the pre-Khalal and Khalal classes. For the harvesting objective, date bunches in the pre-Khalal and Khalal phases are considered to be in the Khalal stage. We only separated these two classes in the dataset labeling section to allow the farmer to give date bunches in the Khalal phase priority to be harvested before those in the pre-Khalal phase.



(a) AlexNet



(b) VGG-16

FIGURE 37. The behavior of the training and validation accuracies and losses of the date maturity classification models based on the five classes.

The CNN models were thus trained to classify five maturity classes using 2305 training images for 5000 iterations. We achieved accuracies of 93.36% and 95.8% on the testing dataset containing 3604 images for the AlexNet and VGG-16 models, respectively. The behavior of the training and validation accuracies and losses during fine-tuning are shown in **FIGURE 37**. The confusion matrices in **TABLE 15** show that most of the misclassified images (270 of 292 and 176 of 192 in **TABLE 15 a and b**, respectively)

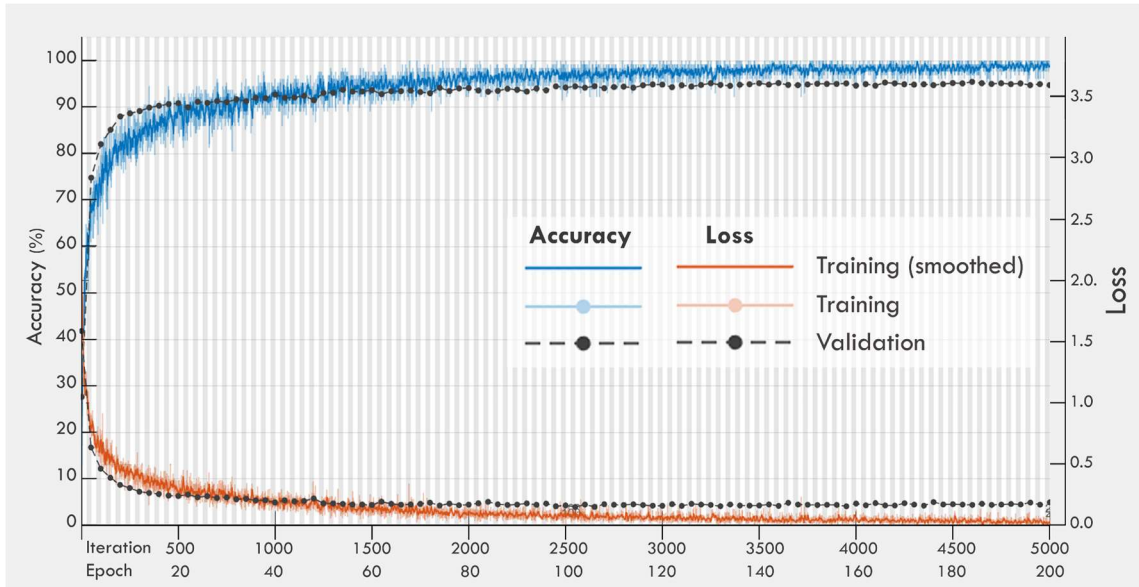
belong to the Khalal or Khalal-with-Rutab classes due to the high similarity between these and their adjacent classes. The Khalal and Khalal-with-Rutab classes (phases) represent the harvesting indicator of date bunches in the Khalal and Rutab stages, respectively. Hence, they cannot be merged or excluded. Instead, it is possible to increase the number of training examples of these two classes compared with the other classes, which will increase the ability of the CNN to learn their fine features and therefore improve the overall accuracy of the maturity classification model. This was done in the final experiment.

TABLE 15. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the five maturity classes

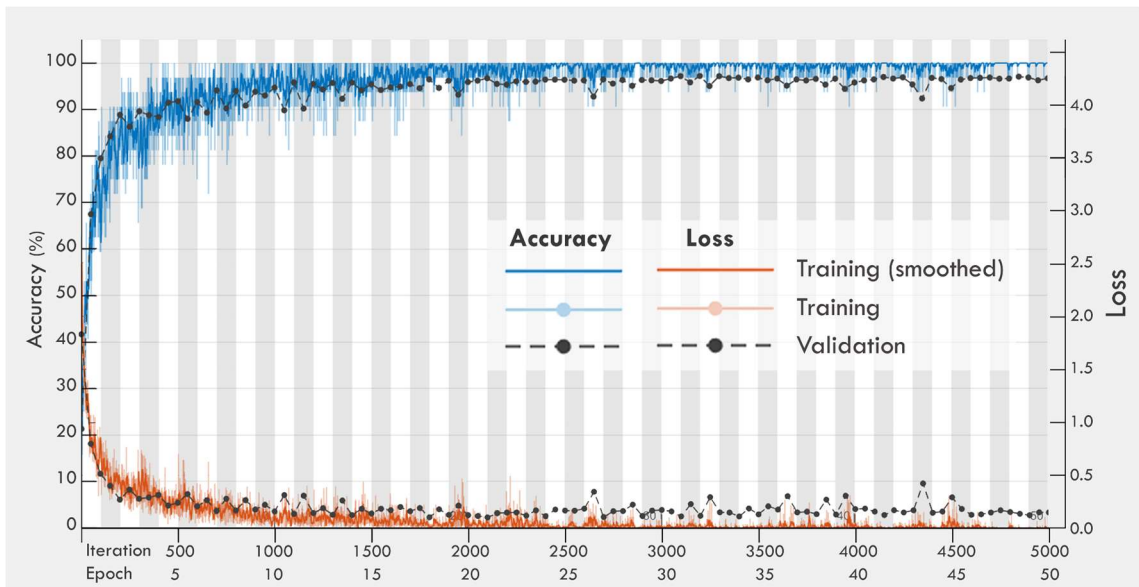
		(a) Predicted class based on AlexNet					(b) Predicted class based on VGG-16						
		Immature	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar	Immature	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar		
Actual class	Immature	1098	2	3	0	5	99.1	1101	2	0	1	4	99.4
	Khalal	9	736	111	6	1	85.3	4	812	47	0	0	94.1
	Khalal-with-Rutab	0	66	1056	62	15	88.1	0	76	1074	46	3	89.6
	Pre-Tamar	0	0	2	229	3	97.9	0	0	1	229	4	97.9
	Tamar	1	0	0	6	193	96.5	0	0	0	4	196	98.0
		99.1	91.5	90.1	75.6	88.9	93.4	99.6	91.2	95.7	81.8	94.7	95.8

3 - The third experiment used the same classes as in the second experiment but with twice the number of training examples for the Khalal and Khalal-with-Rutab classes and fewer testing images. The CNN models were trained for 5000 iterations on 3277 training images (922 for Khalal and Khalal-with-Rutab and 461 for the other classes). The test accuracies rose to 94.98% and 97.25% for AlexNet and VGG-16 models, respectively. The CNN models reached nearly their highest test accuracies after around 3000 iterations and stabilized in the subsequent iterations, as shown in **FIGURE 38**. The

confusion matrices in [TABLE 16](#) show that increasing the training samples of the Khalal and Khalal-with-Rutab classes enhanced the accuracy of the CNN models. However, these two classes still challenge the maturity classification model due to their low inter-class variation and wide intra-class variation in our dataset. [FIGURE 39](#) shows samples of some misclassified images. [TABLE 17](#) presents the results of the three experiments.



(a) AlexNet



(b) VGG-16

FIGURE 38. The behavior of the training and validation accuracies and losses of the date maturity classification models based on the five maturity classes with doubling the number of training examples of the Khalal and Khalal-with-Rutab classes.



FIGURE 39. Samples of misclassified images using the maturity classification CNN models on the five maturity classes

TABLE 16. The confusion matrices (as illustrated in TABLE 11) of the maturity classification models on the testing dataset based on the five maturity classes with doubling the number of training examples of the Khalal and Khalal-with-Rutab classes.

		(a) Predicted class based on AlexNet					(b) Predicted class based on VGG-16				
		Immature	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar	Immature	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar
Actual class	Immature	1095	4	3	0	6	1101	5	0	0	2
	Khalal	1	360	40	0	1	0	390	12	0	0
	Khalal-with-Rutab	0	34	685	19	0	0	43	690	5	0
	Pre-Tamar	1	0	1	224	8	0	0	1	229	4
	Tamar	0	0	0	4	196	0	0	0	3	197
		99.8	90.5	94.0	90.7	92.9	100.0	89.0	98.2	96.6	97.0
						95.0					97.3

TABLE 17. Results of the proposed maturity classification system on the testing dataset using the different settings of the maturity classes.

Exp. No.	Setting of maturity classes	Pre-trained model	Accuracy (%)	PPV (%)	TPR (%)	<i>f</i> score (%)	Classification rate (fps)	
							GPU*	CPU*
1	Seven maturity classes with an equal number of training samples	VGG-16	92.3	86.98	92.3	89.56	57.43	1.36
		AlexNet	90.1	82.51	90.11	86.14	83.75	15.66
2	Five maturity classes with an equal number of training samples	VGG-16	95.78	92.61	95.78	94.17	48.93	1.31
		AlexNet	93.36	89.05	93.36	91.16	93.17	16.47
3	Five maturity classes with an unequal number of training samples	VGG-16	97.25	96.17	97.25	96.71	48.17	1.34
		AlexNet	94.98	93.56	94.98	94.27	93.68	15.35

* GPU: Nvidia GeForce GTX 1060 6 GB, CPU: Intel Xeon E5-2600 with 28 GB RAM.

4.4.5 Harvesting Decision System

Dates are harvested and sold as ripe fruit in three stages: Khalal, Rutab, and Tamar. The choice to harvest in one or other stage depends on many factors such as climatic conditions, date fruit type, and market demand [7]. The proposed system determines the decision to harvest date bunches in two steps. First, the user enters the required harvesting stage for each date type, according to the climatic conditions, market demand, and so on. Next, the system will automatically recognizes the types of date bunches in the orchard and defines their maturity stages, then making the harvesting decision.

Date fruits can be harvested either by picking individual mature dates (selective harvesting) or by cutting the whole bunch when the most dates are mature (bunch-based harvesting). Selective harvesting can start from immature-2 phase and continue until Tamar phase. Dates at the Rutab stage usually harvested as individual dates; however, selective harvesting is highly labor intensive and slow. Harvesting by cutting the whole bunch is the method used in large orchards for commercial production. Therefore, in this study, we investigate the harvesting decision associated with bunch-based

harvesting. In bunch-based harvesting, date bunches labeled Pre-Khalal, Khalal, and Khalal-with-Rutab are harvested in the Khalal stage, whereas only date bunches labeled Khalal-with-Rutab and Tamar are harvested as Rutab or Tamar, respectively, as shown in **FIGURE 40**.

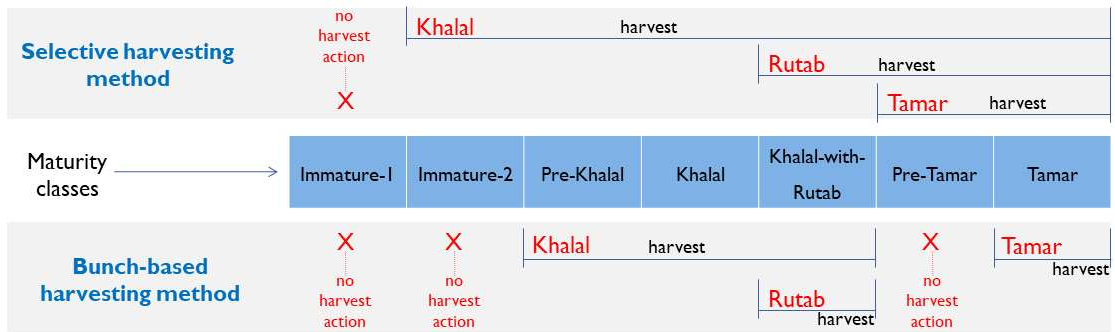


FIGURE 40. The harvesting decision of dates depending on the harvesting methods during the development period.

According to the date fruit characteristics, Barhi and Khalas are consumed in all maturity stages, whereas Meneifi, Naboot Saif, and Sullaj are consumed in Rutab or Tamar [47], as shown in **FIGURE 41**. In this experiment, we test the harvesting decision model for the five date types in the most frequently harvesting stage selected for each type. We assume that the required harvesting stage of Barhi is Khalal, of Meneifi and Sullaj is Rutab, and of Khalas and Naboot Saif is Tamar.

Date type	Maturity stage		
	Khalal	Rutab	Tamar
Barhi	√	√	√
Meneifi		√	√
Sullaj		√	√
Khalas	√	√	√
Naboot Saif		√	√

FIGURE 41. Possible harvesting stages of the five date types in the dataset.

The fine-tuned VGG-16 models of date type and maturity classification in the previous experiments were used to determine the harvesting decision in this experiment. The harvesting decision system was tested using a dataset containing 2682 images of the five date types. Each date type in the dataset was categorized into two classes (harvest and not harvest), as shown in **FIGURE 42**. The performance of the harvesting decision system on the testing dataset is illustrated by the confusion matrix in **TABLE 18**. This matrix consists of four large cells that report the number of true positives, false positives, false negatives, and true negatives of the decision to harvest date bunch images. Each large cell consists of five small cells related to the five date types. The rows represent the required harvesting decision for each date type and the columns represent the predicted harvesting decision. The overall framework performance is shown in **TABLE 19**. This table reports the speed and accuracy measures of the date type and maturity classification models based on the VGG-16 network and the performance of the harvesting decision system based on the entered decision factors and two classification results.

		Maturity phase					
		Immature	Pre-Khalal	Khalal	Khalal-with-Rutab	Pre-Tamar	Tamar
Date type	Barhi	not harvest		harvest			
	Meneifi	not harvest			harvest		
	Sullaj	not harvest			harvest		
	Khalas	not harvest				harvest	
	Naboot Saif	not harvest				harvest	

FIGURE 42. Labeling each date type in the testing dataset according to the harvesting decision. Decisions are based on the assumption that the required harvesting stage of Barhi is Khalal, Meneifi and Sullaj is Rutab, and Khalas and Naboot Saif is Tamar.

TABLE 18. The confusion matrix of the harvesting decision model on the testing dataset.

		Predicted harvesting decision by our framework											
		Harvest					Not harvest						
		Barhi as Khalal	Meneifi as Rutab	Sullaj as Rutab	Khalas as Tamar	Naboot Saif as Tamar	Barhi before Khalal	Meneifi before Rutab	Sullaj before Rutab	Khalas before Tamar	NabootSaif before Tamar		
Required harvesting decision	Harvest	Barhi as Khalal*	239	1	0	0	0	0	0	0	0	0	99.6
		Meneifi as Rutab*	0	73	0	0	0	0	1	0	0	0	98.6
		Sullaj as Rutab*	1	1	430	0	0	0	0	29	0	0	93.3
		Khalas as Tamar*	0	0	0	27	0	0	0	0	0	0	100
		Naboot Saif as Tamar*	0	0	0	0	63	0	0	0	0	1	98.4
	Not harvest	Barhi before Khalal	0	0	0	0	0	318	1	0	0	0	99.7
		Meneifi before Rutab	0	3	0	0	0	1	287	0	1	0	98.3
		Sullaj before Rutab	0	0	2	0	0	0	0	293	0	0	99.3
		Khalas before Tamar	0	0	0	1	0	0	1	0	501	0	99.6
		NabootSaif before Tamar	0	0	0	0	2	2	0	0	0	403	99.0
		99.6	93.6	99.5	96.4	96.9	99.1	99.0	91.0	99.8	99.8	98.6	

* The required harvesting stage for each date type, which is entered manually depending on many factors such as climatic conditions, date fruit characteristics, and market demand.

TABLE 19. Performance of the proposed classification systems (speed and accuracy measures).

Method	Accuracy (%)	PPV (%)	TPR (%)	fscore (%)	Average classification time (msec./image)	
					GPU*	CPU*
Date type classification system	99.01	98.82	99.01	98.92	20.6	740.2
Date maturity classification system	97.25	96.17	97.25	96.71	20.7	745.9
Harvesting decision system	98.59	97.46	98.59	98.02	35.9	1298

* GPU: Nvidia GeForce GTX 1060 6 GB, CPU: Intel Xeon E5-2600 with 28 GB RAM.

4.4.6 Comparison with Other Classification Methods

We compared the performance of the proposed classification method with that of the methods used in other studies of date fruit type and maturity classification, as shown in TABLE 20. The author in [27] used the SVM to classify date types using LBP and WLD

local texture descriptors combined with size and shape features. He classified date images of four date types and achieved an accuracy of 98.1%. In [31], the authors used 15 features including color, shape, size, and texture descriptors to classify dates according to their types using neural networks. The study reported an accuracy of 98.6%; however, only 140 images were used to classify and test seven date classes. The authors in [32] classified date images based on their types using a dataset containing 5000 images of 10 date types. They used an RGB color histogram, GLCM, and four shape features including area and eccentricity, combined with the GMM. They reported an accuracy of 97.5% with a classification time of 0.029 seconds. All these studies used single date images with a uniform background. Hence, previously used methods are unsuitable for real-life applications in natural environments where the variety in the data is enormous (e.g. in date orchards).

TABLE 20. Performance comparison between the methods.

(a) Date fruit type classification

Method	Dataset type	Number of testing images	Accuracy (%)	Average classification time (msec./image)	
				GPU	CPU
Proposed with VGG-16 fine-tuning	Date fruit bunches in an orchard environment	3542	99.01	20.6	740.2
Proposed with AlexNet fine-tuning			96.51	11.9	56.6
[27]	Single dates with uniform background	10-fold validation on 800	98.1	-	-
[31]		10-fold validation on 140	98.6	-	-
[32]		1000	97.5	29	-
[33]	Single and multiple dates after the production	1000	99.2	-	-

(b) Date fruit maturity classification

Proposed with VGG-16 fine-tuning	Date fruit bunches in an orchard environment	2682	97.25	20.7	745.9
Proposed with AlexNet fine-tuning			94.98	10.7	65.1
[36]	Single dates with a uniform background	353	97.51	-	-
[37]		100	88.33	340	-

Few studies of maturity classification have been performed, and they all have the same limitations discussed above. One such study [36] used a color distribution analysis and back projection to classify one date type into four maturity classes. The authors achieved a classification accuracy of 97.5% on single date images. In another study [37], a taxonomy classification method with RGB color and texture features including contrast, entropy, and homogeneity was used to classify a date type into four maturity classes. The authors reported an overall accuracy of 88.33% with a 0.34-second classification time.

In the most recent approach [33], the researchers proposed a date type classification based on a pre-trained CNN model and reported an accuracy of 99.2%. They used images of single and multiple dates for four date types. However, the studied date types had a wide visual inter-class variation and their images were all taken after the production stage. Hence, the classification task was much simpler than the classification in this work.

All previous approaches have used single or multiple date images taken after the harvesting or production stages. By contrast, this work deals with date bunch images in different pre-maturity and maturity stages with a large degree of variation in an orchard environment, which makes them challenging to classify.

TABLE 21 compares the performance of the proposed approach with that of other recent approaches using the same dataset that we created. The methods described in [27], [31], and [33] were carried out and evaluated using the same environment and machine used in the previous experiments. To conduct the methods of [27] and [31], color and texture features were used and size and shape features were ignored because they cannot identify the maturity stage of dates or their types as bunches. A k-nearest neighbor

classifier was chosen to implement [31] and LBP was chosen as the texture descriptor in [27]. **FIGURE 43** shows the average accuracy obtained by the proposed approach in the three classification problems and by the approaches in [27], [31], and [33]. **FIGURE 44** presents the average classification time of each method using a machine with and without a GPU. The results show that the proposed approach achieved excellent classification accuracies, which outperformed those of the other approaches and had a low classification time using a GPU.

TABLE 21. Performance comparison between the methods using the same dataset adopted in this work.

Method	Accuracy (%)			Average classification time of harvesting decision system (msec/image)	
	Date type classification	Date maturity classification ⁽¹⁾	Harvesting decision	GPU ⁽²⁾	CPU ⁽²⁾
Proposed with VGG-16	99.01	97.25	98.59	35.9	1298
Proposed with AlexNet	96.51	94.98	95.51	22.3	115.8
[27]	81.22	80.38	80.45	125	129.2
[31]	58.82	63.03	62.01	37.6	37.9
[33]	95.92	93.40	95.28	17.5	103

⁽¹⁾ Using the same training and testing datasets described in the maturity classification model subsection - the third experiment.

⁽²⁾ GPU: Nvidia GeForce GTX 1060 6 GB, CPU: Intel Xeon E5-2600 with 28 GB RAM.

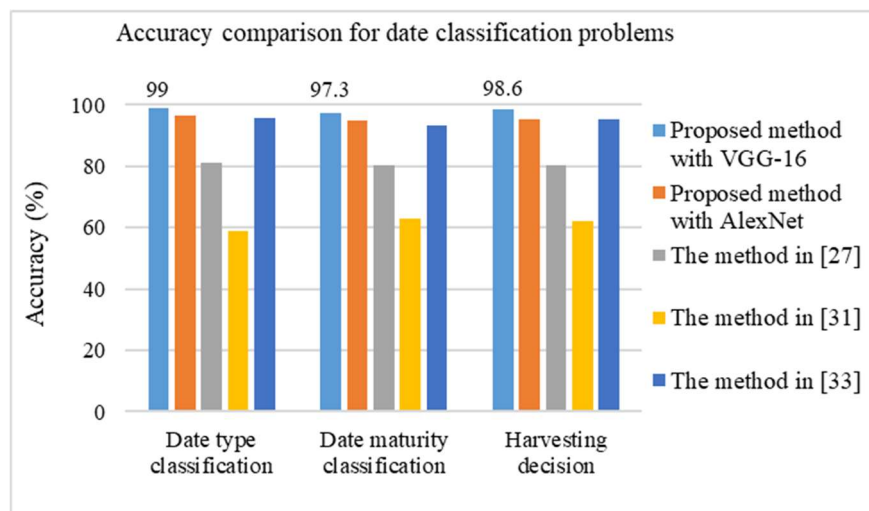


FIGURE 43. Comparison between the accuracy obtained by the proposed approach for date classification problems and the approaches in [27], [31], and [33].

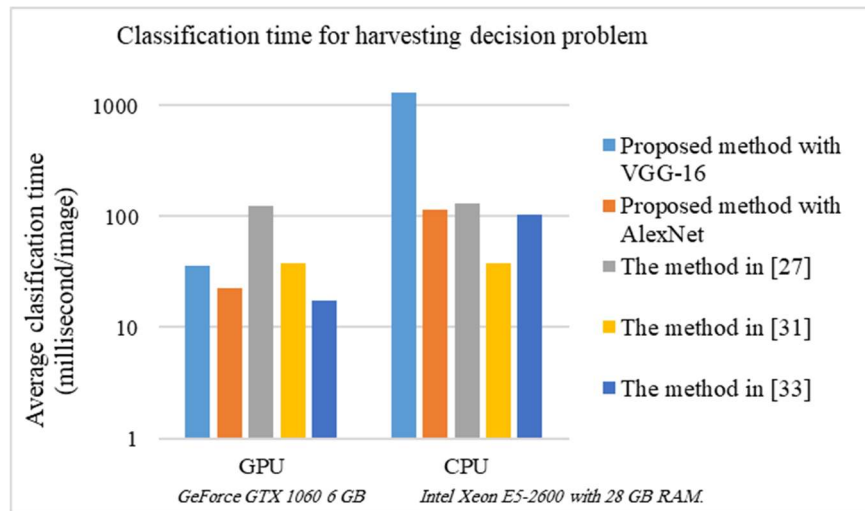


FIGURE 44. Average classification time of the date harvesting decision system based on the proposed method and the methods in [27], [31], and [33], using a machine with a GPU and CPU.

4.5 CONCLUSION

A real-time machine vision framework for date fruit harvesting robots in an orchard environment was proposed based on deep learning. The framework consisted of three classification systems used to classify date fruit bunches according to their type, maturity, and harvesting decision. Transfer learning with fine-tuning was used in the classification tasks. Two pre-trained CNN models were investigated, AlexNet and VGG-16. To build a robust machine vision system, we used a rich image dataset of five date types for all maturity stages. The dataset was designed with a large degree of variation that represents the challenges in natural environments and date fruit orchards. The proposed approach achieved outstanding classification accuracies on this challenging dataset with a high classification rate. The results showed that a pre-trained CNN could achieve robust date fruit classification without the pre-processing of images to remove background noise or enhance illumination. The best accuracies were obtained by the fine-tuned VGG-16 model, which achieved 99.01%, 97.25%, and 98.59% accuracies with classification times of 20.6, 20.7, and 35.9 msec for the date fruit type, maturity, and harvesting decision classification models, respectively.

CHAPTER 5

DATE FRUIT SEGMENTATION IN AN ORCHARD

5.1 INTRODUCTION

In robotic harvesting for date fruit, the robot first will recognize the type and maturity of date bunches and determine the harvesting decision: ready to harvest or not. Then, the segmentation will take place to localize each date bunch and locate the position to cut the stalk. This is needed to move the arm of the robot to harvest the bunch. Fruit segmentation is also important to estimate the weight of date bunches and for yield prediction. In this chapter, we propose a robust system for automatically segmenting date fruits from trees and orchard background. In the proposed system, a color image of a date palm is broken into structurally meaningful regions of pixels (superpixels) using the simple linear iterative clustering algorithm. Then, color and texture features in the form of local binary pattern are extracted from each superpixel. Finally, the support vector machine is used to classify each region into fruit or background.

The models, codes, data, and test videos of the developed system is available in the following website:

<https://sites.google.com/view/daterobotic/fruit-detection-segmentation>

5.2 PROPOSED APPROACH

FIGURE 45 shows a block diagram of the proposed date fruit segmentation system. The input of the system is a color image of date bunches on a tree. Once the image is inputted to the system, some pre-processing is applied to enhance the edge information. Then, the image is clustered into regions of pixels (superpixels) using simple linear iterative clustering algorithm (SLIC) [60]. The superpixels in the color image are converted into several color components. For each superpixel, texture information, in the form of LBP, and color features are extracted from the color components separately and fused to form the feature vector of the superpixel. Finally, the feature vector is used to classify each superpixel using SVM into two classes: fruit or background. The steps in details are discussed in the following.

5.2.1 Pre-processing

In the pre-processing step, first, the image is scaled to 1200*800 to lower processing time. The gradient magnitude is extracted from the image and normalized (each pixel gradient value is divided by the highest gradient value in the image). The normalized image is then multiplied by the original image. Finally, histogram equalization is applied to the final image. This pre-processing enhances the superpixels to adhere well to edge information of the original image.

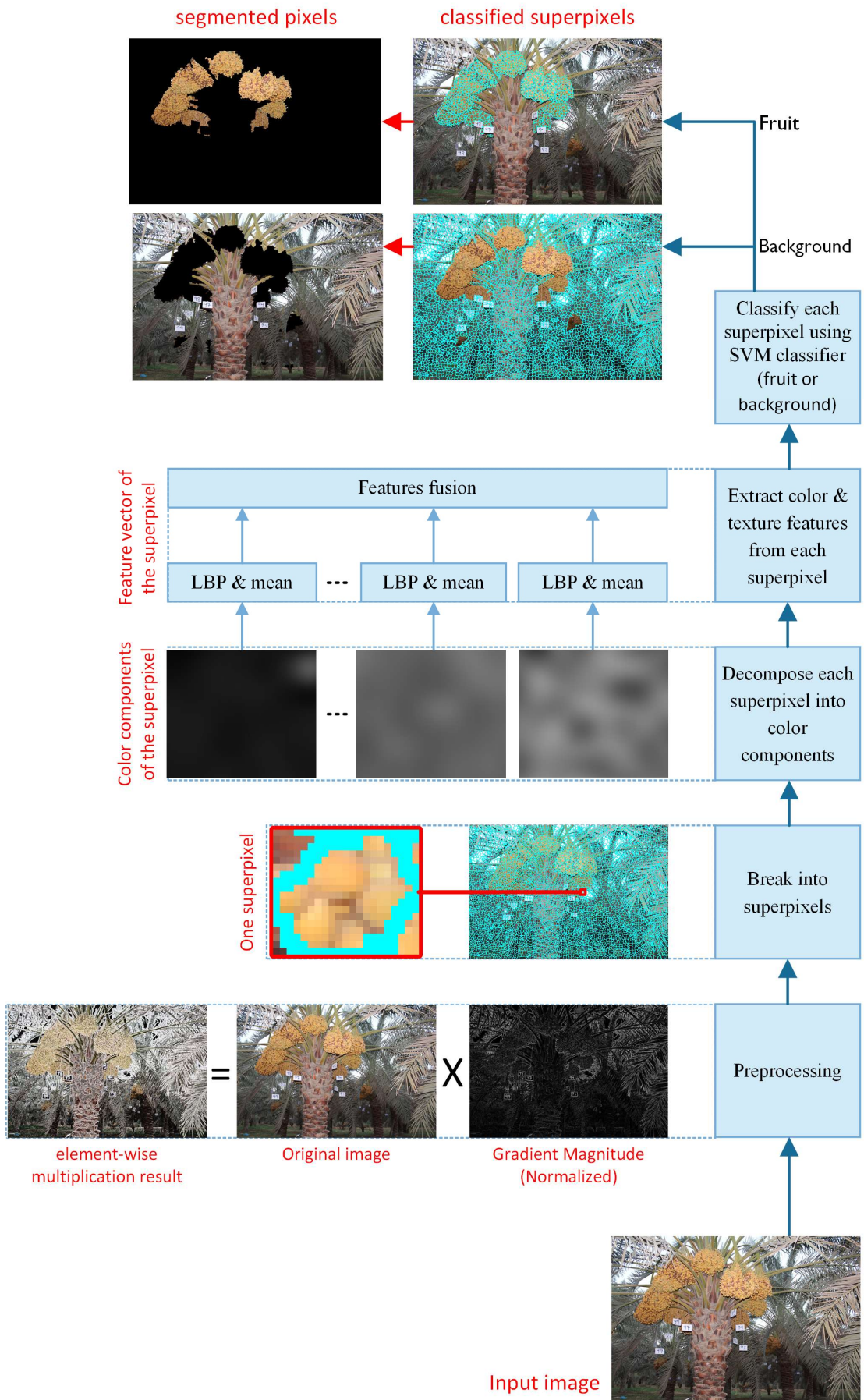


FIGURE 45. Block diagram of the date fruit segmentation system.

5.2.2 Superpixels Clustering

Superpixels clustering is a powerful technique to perform segmentation and classification. Superpixels divide an image into small local regions according to color and spatial information, where the boundaries of each region adhere to the information of the original image. We employ the SLIC [60] approach to generate superpixels, which adapts k-means algorithm to perform clustering. Standard k-mean searches the entire image while SLIC searches a limited region, as shown in **FIGURE 46**. SLIC clusters pixels into superpixel regions based on their color, represented by the components of CIELAB color space $[l, a, b]$, and their position in the image $[x, y]$.

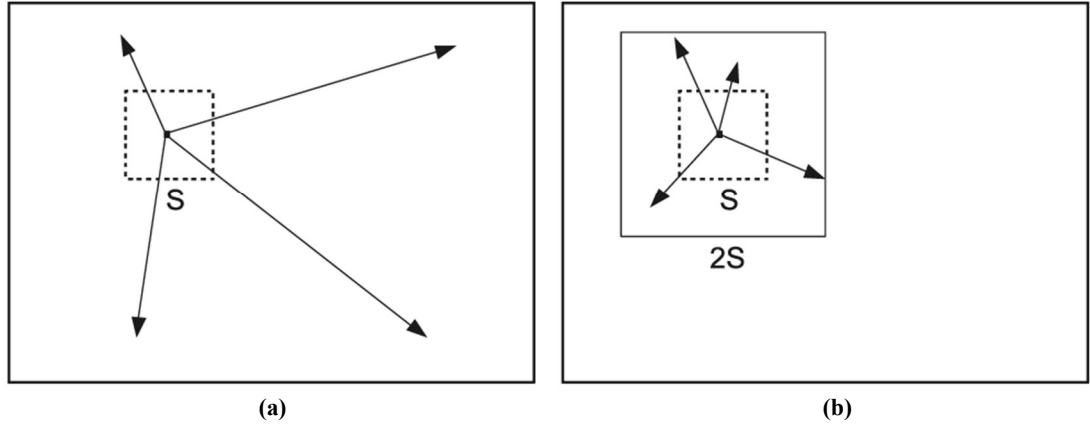


FIGURE 46. Standard k-mean searches the entire image (a), while SLIC searches a limited region (b).

For an image with N pixels and K superpixels, the average size of each superpixel is equal to N/K pixels, and each superpixel has a center at every grid interval $S = \sqrt{N/K}$.

If $C_k = [l_k, a_k, b_k, x_k, y_k]$ are superpixel centers, and $P_i = [l_i, a_i, b_i, x_i, y_i]$ are the parameters of a pixel (i) that is located within a $2S \times 2S$ region around the superpixel center, the distance of each pixel from the nearest cluster center is calculated as in Eq. (5).

$$\begin{aligned}
 d_{lab} &= \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \\
 d_{xy} &= \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \\
 D_s &= d_{lab} + \frac{m}{S} d_{xy}
 \end{aligned} \tag{5}$$

Where D_s is the sum of the color and spatial Euclidean distances normalized by the grid interval S . The parameter m is used to adjust the compactness of the superpixels. A larger value of m makes superpixels more compact and more regularly shaped (i.e. square). A smaller value makes superpixels adhere well to image boundaries, but have irregular shape and size. For details, the reader can refer to [60], [61].

SLIC takes as input two parameters, the number of superpixels (K) and their compactness (m). Then, according to the image size, the SLIC generates superpixels of average size equal to N/K . In our experiments in this chapter, we set the compactness value to $m = 10$, which provides a good balance between spatial closeness and color similarity.

Increasing the number of superpixels in an image reduces the probability of getting superpixels containing pixels that belong to both fruit and background. The number of superpixels can be increased by either increasing image resolution or decreasing the size of superpixels (number of pixels). However, increasing the number of superpixels rises the required computation time to perform superpixel clustering as well as the time to perform feature extraction and classification.

Based on many experiments we performed, we found that the optimal settings for superpixels are $K=6000$ regions, $S^2=160$ pixel for each region, and adaptively (automatically) refined compactness with average value $m = 10$. **FIGURE 47** shows that for a scene that has one date palm, around 6000 superpixels allow us to cluster the fruit and background better. We selected the average size of each superpixel to be 160 pixels which helped us extract the texture information of the fruit and background. Therefore, we downsampled the input image into 1200×800 which fit dividing the image into 6000 superpixels with 160-pixel size.

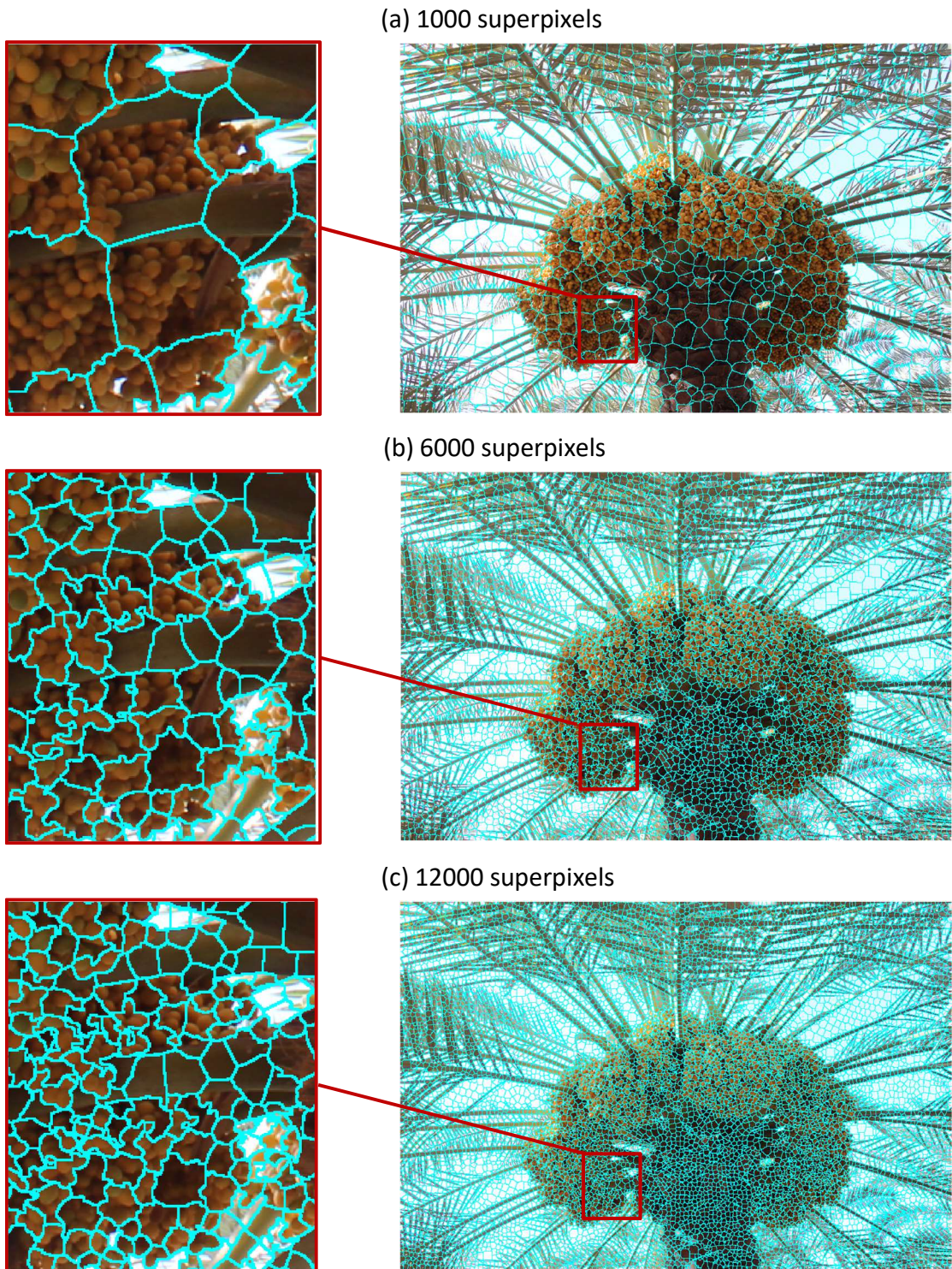


FIGURE 47. Date palm image clustered using the SLIC algorithm with a different number of superpixels.

5.2.3 Color Components

In this study, we investigate the color components of three color spaces, RGB {red (R), green (G), and blue (B)}, YCbCr {luminance (Y), chrominance (Cb and Cr)}, and HSV {hue (H), saturation (S), and value (V)}. It can be noted that a color image is passed

through three color filters (R, G, and B) inside the camera. YCbCr color space stores the color in terms of luminance and chrominance, where the human eyes are less sensitive to chrominance than luminance. Chrominance components can unveil some of the cues of differentiating the dates that can be impossible to detect by the naked human eyes. On the other hand, HSV attributes are the closest approximation to human interpretation of color. The logic behind using these color spaces is that the dates vary in colors and therefore contain important information in different color spaces.

Given a digital pixel represented in RGB format, 8 bits per sample, where 0 and 255 represents the black and white color, respectively, the YCbCr components can be obtained according to equations (6-8) [62].

$$Y = 16 + \frac{65.738 * R}{256} + \frac{129.057 * G}{256} + \frac{25.064 * B}{256} \quad (6)$$

$$Cb = 128 - \frac{37.945 * R}{256} - \frac{74.494 * G}{256} + \frac{112439 * B}{256} \quad (7)$$

$$Cr = 128 + \frac{112.439 * R}{256} - \frac{94.154 * G}{256} - \frac{18.285 * B}{256} \quad (8)$$

Given a digital pixel represented in RGB format, if $R' = R/255$, $G' = G/255$, $B' = B/255$, $C_{max} = \max(R', G', B')$, $C_{min} = \min(R', G', B')$, and $\Delta = C_{max} - C_{min}$, the HSV components can be obtained according to equations (9-11) [62].

$$H = \begin{cases} 0^\circ & \Delta = 0 \\ 60^\circ \times \left(\frac{G' - B'}{\Delta} + \text{mod}6 \right), C_{max} = R' \\ 60^\circ \times \left(\frac{B' - R'}{\Delta} + 2 \right), C_{max} = G' \\ 60^\circ \times \left(\frac{R' - G'}{\Delta} + 4 \right), C_{max} = B' \end{cases} \quad (9)$$

$$S = \begin{cases} 0 & , C_{max} = 0 \\ \frac{\Delta}{C_{max}} & , C_{max} \neq 0 \end{cases} \quad (10)$$

$$V = C_{max} \quad (11)$$

5.2.4 Local Binary Pattern

The LBP is a powerful texture descriptor that labels the pixels of an image by decimal numbers called LBP codes [63]. LBP encodes the local structure around the center pixel by comparing the pixel value to its eight neighbor pixels in a 3 x 3 neighborhood. If the value of the center pixel is less than the value of its neighbor, it is encoded with 1, otherwise, it is encoded with 0. This gives eight binary values that are concatenated to form an 8-digit binary number. The binary number is converted to the corresponding decimal value which used as a label for the given center pixel. In the proposed method, we used the uniform pattern [64] to reduce the length of the feature vector and implement a simple rotation invariant descriptor. We used 8 pixels neighboring ($p=8$) with radius 1 ($R=1$) for the circular position.

5.2.5 Support Vector Machine

Support vector machines (SVM) is a supervised learning technique that analyzes data for regression and classification. SVM performs classification by constructing an N-dimensional hyperplane that optimally separates the classes of data into two categories. SVM defines optimality by maximizing the distance between the separating hyperplane to the nearest neighbor in each of the separated classes [65].

SVM maps, nonlinearly, the original input space into a higher-dimensional feature space. The generalization capabilities of the classifier are maximized by this mapping. Because SVM is a two-class classifier, the output given by this classifier for each image sample can be interpreted as the probability that the sample belongs to a specific class. Support vectors are created during the training phase, and these vectors are used to classify the samples during the testing phase. Given the training samples, $\{(x_i, lab_i)$:

$i=1,2,\dots, n\}$ where $lab_i \in [-1,+1]$ is the class label, and x_i is the feature vector of the i th training samples, the optimum boundary is defined as [65]:

$$f(x) = w \cdot x + b \quad (12)$$

Where w is the weight to be multiplied by the input neurons. We can obtain w and b by solving the following optimization equation, where ε is the error, and C is the penalty parameter ($C>0$):

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_i \varepsilon_i \quad (13)$$

5.3 RESULTS AND DISCUSSIONS

To train the proposed system and evaluate it, we labeled the superpixels in images of eight date palm into fruit and background. We first selected some parts of the images that belong to the date fruit or background, samples are shown in [FIGURE 48](#). We used Labeler app in Matlab 2017b to label the fruit and background regions, as shown in [FIGURE 49](#). A total of 91 background regions and 293 fruit regions were labeled from eight images of Barhi date in the Khalal and Rutab maturity stages. Then, using SLIC, these regions were divided into superpixels. For the labeled regions of the eight image, we extracted 15460 superpixels. Each superpixel was classified using the SVM into fruit or background. With 10-fold cross validation, we achieved 99.8 classification accuracy for the superpixels. However, using these superpixels for segmentation will result in low segmentation accuracy because these superpixels did not cover the different regions of the images and especially it did not include the boundary regions between the fruit and background. Therefore, we created a graphical interface (UI) to help segmenting the whole image into background and foreground manually, as shown in [FIGURE 50](#). The interface first performs initial segmentation using the previous model,

then we manually label all pixels of the image by accepting the correct labeling and correcting the wrong labeling. We did this labeling on 16 images (the initial eight plus eight others). Next, we used the superpixels of thirteen of these images for training and the superpixels of the remaining three images for testing the classification and segmentation system as explained below.

a) Date fruit



b) Background



FIGURE 48. Samples of date fruit and background regions.

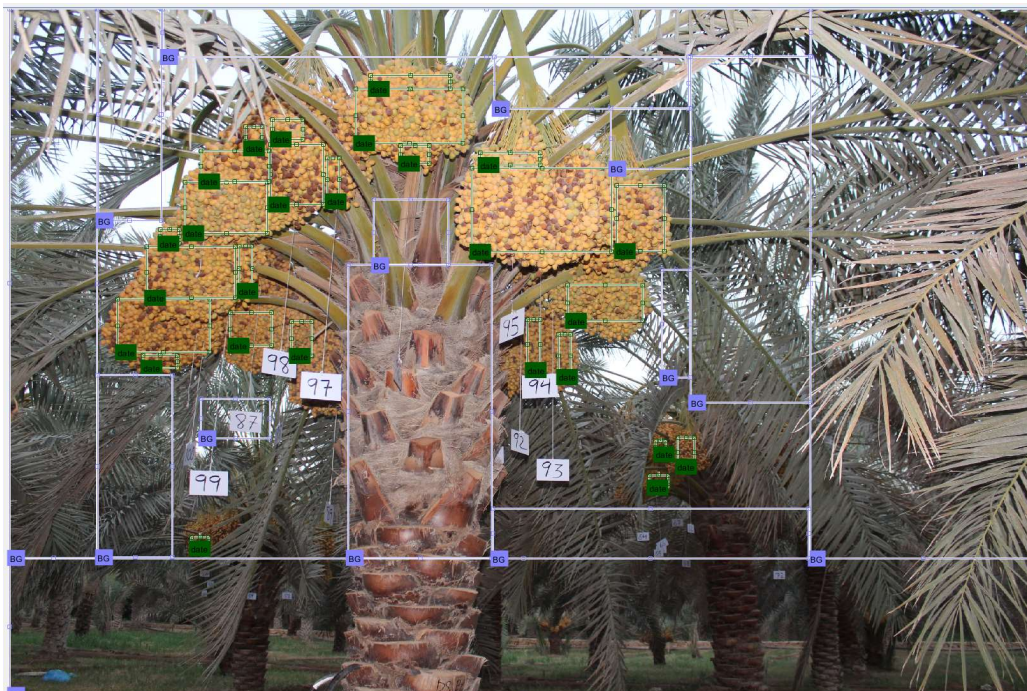


FIGURE 49. Label an image into fruit and background regions using image Labeler app in Matlab.

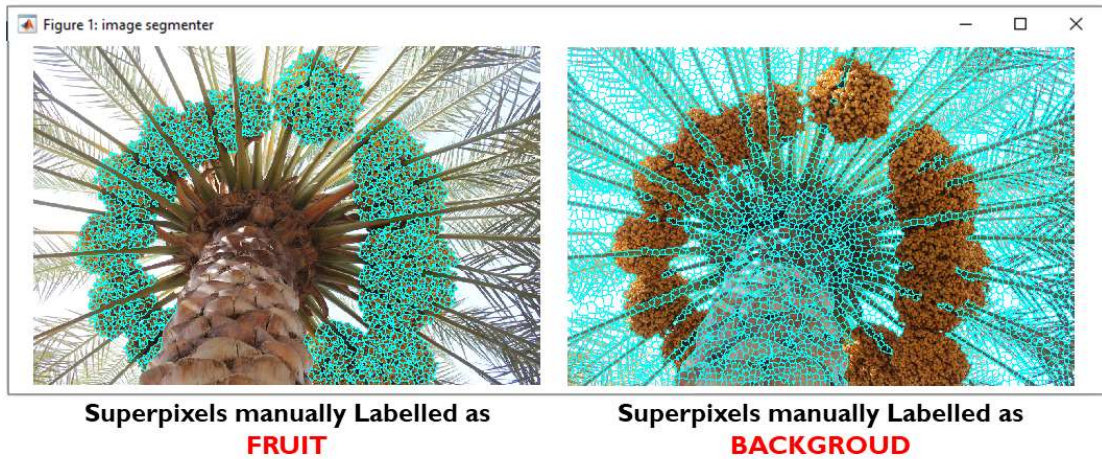


FIGURE 50. Graphical interface (UI) to help segmenting the images into fruit and background.

In this study, we investigated different color and texture features including mean, median, mode, entropy, gradient magnitude, and LBP with different color spaces and color components. We found that LBP combined with color mean features for each color component is sufficient to provide the best accuracies. **TABLE 22** presents, with 10-fold cross validation, the classification accuracy for the superpixels using color mean and LBP features for different color spaces and color components. The results show that the combination of the color components R, Cb, and H gives the best accuracies. **FIGURE 51** shows (R, Cb, H) color components of date fruit image.

TABLE 22. Cross-validation accuracy of the proposed date fruit segmentation system.

Color space/components	RGB	YCbCr	HSV	R, Cb, H
Accuracy (%)	97.3	98.2	97.6	98.4

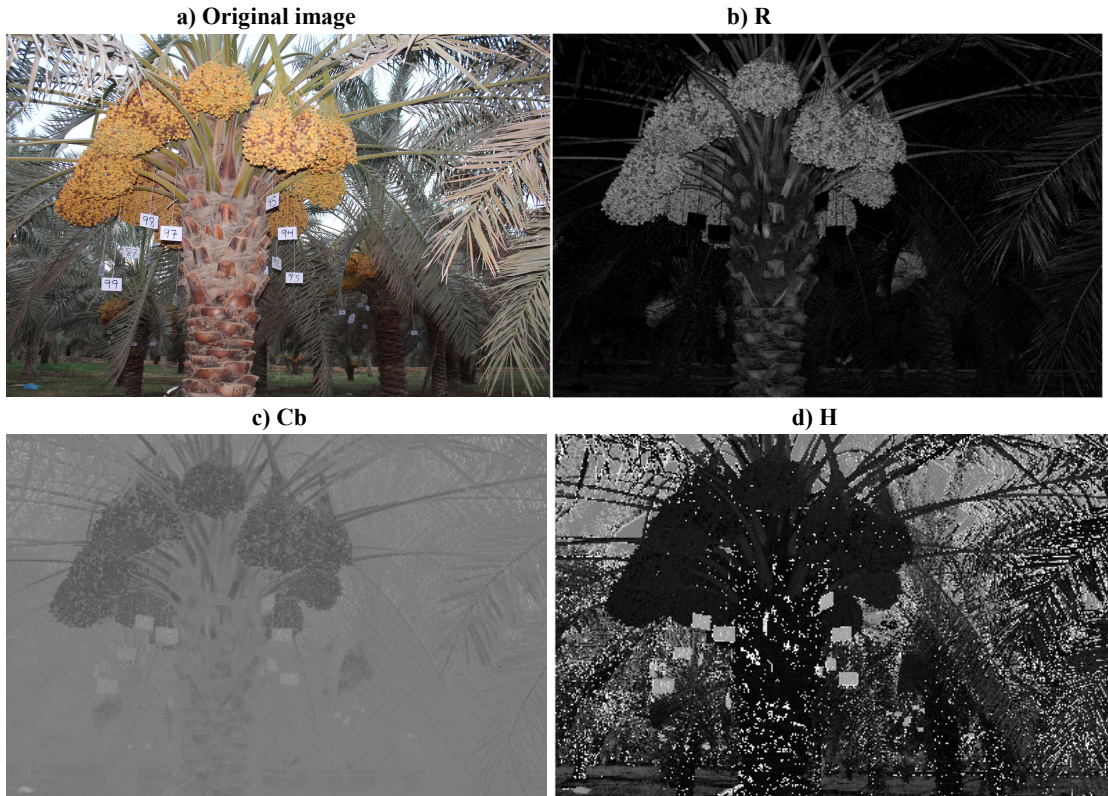


FIGURE 51. Samples of the color components of images that can discriminate the date bunches from the background. (a) Original image; (b) R image component; (c) Cb image component; (d) H image component.

We used the measures of recall Eq. (8), precision Eq. (9), and f-score Eq. (10), to evaluate the pixel-based accuracy of the proposed segmentation system. **TABLE 23** shows the segmentation performance using the pixels of testing images in the three maturity stages (Immature, Khalal, and Tamar). Samples result of the proposed date fruit segmentation system using the (R, Cb, H) color components are shown in **FIGURE 52**.

$$\text{Recall (true positive rate TPR)} = \frac{Region_{manual} \cap Region_{segmented}}{Region_{manual}} \times 100 \quad (8)$$

$$\begin{aligned} \text{Precision (positive predictive value PPV)} \\ = \frac{Region_{manual} \cap Region_{segmented}}{Region_{segmented}} \end{aligned} \quad (9)$$

$$fscore = \frac{1}{\alpha/TPR + \alpha/PPV} \quad (10)$$

where $\alpha = 0.5$ gives equal weights to precision and recall

TABLE 23. The performance of the proposed date fruit segmentation system.

	TPR (recall)				PPV (precision)				f-score			
Image code	M1.S4.B.74	B5.S2.B.27	N5.S5.B.3	all	M1.S4.B.74	B5.S2.B.27	N5.S5.B.3	all	M1.S4.B.74	B5.S2.B.27	N5.S5.B.3	all
Date type	Mencifi	Barhi	Naboot Saif		Mencifi	Barhi	Naboot Saif		Mencifi	Barhi	Naboot Saif	
Maturity stage	Khalal	Immature	Tamar		Khalal	Immature	Tamar		Khalal	Immature	Tamar	
RGB	0.937	0.807	0.804	0.886	0.897	0.894	0.902	0.897	0.916	0.848	0.844	0.889
YCbCr	0.951	0.868	0.864	0.918	0.884	0.867	0.885	0.882	0.916	0.867	0.873	0.898
HSV	0.953	0.868	0.837	0.913	0.888	0.869	0.904	0.889	0.919	0.869	0.865	0.899
R&Cb&H	0.951	0.887	0.852	0.918	0.887	0.856	0.894	0.884	0.917	0.870	0.871	0.899

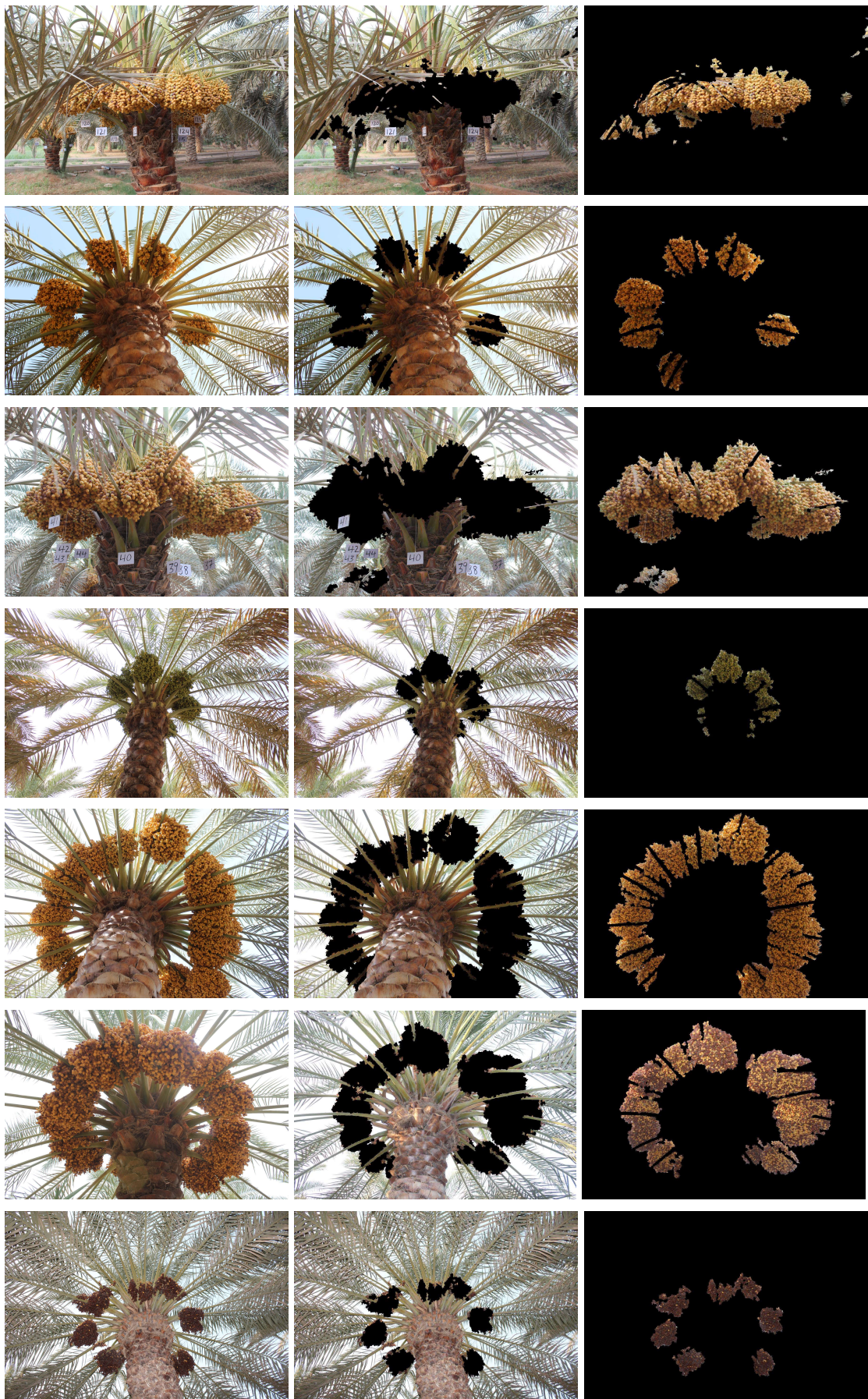


FIGURE 52. Samples segmented images using our proposed fruit segmentation method.

5.4 CONCLUSION

We proposed a robust method to segment date fruits from trees and orchard's background. We used SLIC to initially group pixels in an image into superpixels. Then, each superpixel was classified into fruit or background using LBP, as a feature descriptor, and SVM, as a classifier. The use of superpixels combined with LBP achieved a high-quality segmentation result.

CHAPTER 6

CONCLUSION AND FUTURE WORK

In this study, we proposed an efficient deep learning based classification framework for date fruit harvesting robots in an orchard environment. The framework consisted of three classification systems to classify date fruit images in real-time according to their type, maturity, and harvesting decision. In the classification systems, we utilized deep convolutional neural networks (CNN) due to their robust ability in automatic feature representation for challenging classification tasks as compared to the handcrafted feature-based approaches. We employed transfer learning with fine-tuning using two pre-trained CNN models namely: AlexNet and VGGNet. We achieved excellent classification accuracies using transfer learning with short training time using a relatively small dataset. The best accuracies were obtained by the fine-tuned VGG-16 model, while AlexNet achieved the highest classification speed. The results showed that a pre-trained CNN could achieve robust date fruit classification without pre-processing of images to remove background noise or enhance illumination. We trained the classification models using a challenging dataset that has a large degree of variations in date orchard environment including variations in angles, scales, illumination conditions, and with date bunches covered by bags. We accomplished excellent results on this challenging dataset.

In the case of date fruit segmentation, we proposed a robust method to segment date fruits from trees and orchard's background. We used simple linear iterative clustering to initially group pixels in an image into superpixels. Then, color and texture features in the form of local binary pattern are extracted from each superpixel. Finally, the

support vector machine is used to classify each region into fruit or background. We investigated different color and texture features including mean, median, mode, entropy, gradient magnitude, and LBP with different color spaces and color components. The results showed that LBP combined with color mean features for (R-B), Cb, and H color components achieve the best accuracies.

We also built a comprehensive date fruit dataset to address the requirements of many applications in the pre-harvesting and harvesting stages including automated harvesting, fruit detection and segmentation, classification, maturity analysis, and visual yield estimation. The dataset was provided with a large degree of variations to reflect the challenges in a natural environment and date fruit orchards. The-dataset has been fully labeled, coded, and released with their associated files to the research community in IEEE DataPort repository [1]

Finally, we would like to point out that the machine vision for date fruit harvesting robot is still a challenging issue. As for future work, currently, we are improving the dataset by including more videos. In the classification tasks, we are investigating more recent CNN models to minimize the usage of memory and lower computational complexity. In the case of date fruit segmentation, we are developing deep learning-based semantic segmentation system, which offer better accuracy and speed results.

The outputs of our current and future research are published in the following website <https://sites.google.com/view/daterobotic>

In the website, we are publishing the resources of our research in date fruits, including source codes, datasets, models, and real time test videos.

PUBLICATIONS

- H. Altaheri, M. Alsulaiman, M. Faisal, and G. Muhammed, “Date Fruit Dataset for Automated Harvesting and Visual Yield Estimation”, IEEE DataPort, v1, 2019. [Online]. Available: <http://dx.doi.org/10.21227/x46j-sk98>.
- H. Altaheri, M. Alsulaiman, and G. Muhammad, “Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning,” IEEE Access, vol. 7, no. 1, pp. 117115–117133, Aug. 2019.
- H. Altaheri, M. Alsulaiman, G. Muhammad, S. U. Amin, M. Bencherif, and M. Mekhtiche, “Date fruit dataset for intelligent harvesting,” Data Br., vol. 26, p. 104514, Oct. 2019.

REFERENCES

- [1] H. Altaheri, M. Alsulaiman, M. Faisal, and G. Muhammed, "Date Fruit Dataset for Automated Harvesting and Visual Yield Estimation," 2019. [Online]. Available: <http://dx.doi.org/10.21227/x46j-sk98>.
- [2] "Food and Agriculture commodities production FAOSTAT." .
- [3] A. A. Al-Janobi A., "Evaluation of field test of harvesting system for picking dates fruits based on robotic arm.," in *International Conference on Robotics and Associated High-technologies and Equipment for Agriculture*, 2012, pp. 183–188.
- [4] K. Kapach, E. Barnea, R. Mairon, and Y. Edan, "Computer vision for fruit harvesting robots–state of the art and challenges ahead," *Int. J. Comput. Vis. Robot.*, vol. 3, pp. 4–34, 2012.
- [5] A. Gongal, S. Amatya, M. Karkee, Q. Zhang, and K. Lewis, "Sensors and systems for fruit detection and localization: A review," *Comput. Electron. Agric.*, vol. 116, pp. 8–19, 2015.
- [6] J. M. Al-Khayri, S. M. Jain, and D. V. Johnson, Eds., *Date Palm Genetic Resources and Utilization*, 1st ed. Dordrecht: Springer Netherlands, 2015.
- [7] Z. Abdelouahhab and E. Arias-Jimenez, *Date palm cultivation.*, 1st ed. Food and Agriculture Organization (FAO), 2002.
- [8] A. A. Kader and A. M. Hussein, "Harvesting and postharvest handling of dates," *ICARDA, Aleppo, Syria*, vol. 4, p. 15, 2009.
- [9] P. Li, S. Lee, and H.-Y. Hsu, "Review on fruit harvesting method for potential use of automatic fruit harvesting systems," *Procedia Eng.*, vol. 23, pp. 351–366, 2011.
- [10] M. Karkee and Q. Zhang, "Mechanization and automation technologies in specialty crop production," *Resour. Mag.*, vol. 19, no. 5, pp. 16–17, 2012.
- [11] Y. Edan, D. Rogozin, T. Flash, and G. E. Miles, "Robotic melon harvesting," *IEEE Trans. Robot. Autom.*, vol. 16, no. 6, pp. 831–835, 2000.
- [12] A. Silwal, A. Gongal, and M. Karkee, "Apple identification in field environment with over the row machine vision system," *Agric. Eng. Int. CIGR J.*, vol. 16, no. 4, pp. 66–75, 2014.
- [13] R. Linker, O. Cohen, and A. Naor, "Determination of the number of green apples in RGB images recorded in orchards," *Comput. Electron. Agric.*, vol. 81, pp. 45–57, 2012.
- [14] R. Zhou, L. Damerow, Y. Sun, and M. M. Blanke, "Using colour features of cv. 'Gala' apple fruits in an orchard in image processing to predict yield," *Precis. Agric.*, vol. 13, no. 5, pp. 568–580, Oct. 2012.
- [15] L. Qiang, C. Jianrong, L. Bin, D. Lie, and Z. Yajing, "Identification of fruit and

- branch in natural scenes for citrus harvesting robot using machine vision and support vector machine,” *Int. J. Agric. Biol. Eng.*, vol. 7, no. 2, pp. 115–121, 2014.
- [16] S. Liu and M. Whitty, “Automatic grape bunch detection in vineyards with an SVM classifier,” *J. Appl. Log.*, vol. 13, no. 4, pp. 643–653, Dec. 2015.
- [17] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [18] D. M. Bulanon, T. F. Burks, and V. Alchanatis, “Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection,” *Biosyst. Eng.*, vol. 101, no. 2, pp. 161–171, 2008.
- [19] D. Stajanko, M. Lakota, and M. Hočevár, “Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging,” *Comput. Electron. Agric.*, vol. 42, no. 1, pp. 31–42, Jan. 2004.
- [20] O. Safren, V. Alchanatis, V. Ostrovsky, and O. Levi, “Detection of Green Apples in Hyperspectral Images of Apple-Tree Foliage Using Machine Vision,” *Trans. ASABE*, vol. 50, no. 6, pp. 2303–2313, 2007.
- [21] H. Okamoto and W. S. Lee, “Green citrus detection using hyperspectral imaging,” *Comput. Electron. Agric.*, vol. 66, no. 2, pp. 201–208, May 2009.
- [22] I. Sa *et al.*, “DeepFruits: A Fruit Detection System Using Deep Neural Networks,” *Sensors*, vol. 16, no. 8, p. 1222, Aug. 2016.
- [23] S. Bargoti and J. Underwood, “Deep fruit detection in orchards,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 3626–3633.
- [24] J. Rakun, D. Stajanko, and D. Zazula, “Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry,” *Comput. Electron. Agric.*, vol. 76, no. 1, pp. 80–88, Mar. 2011.
- [25] F. Kurtulmus, W. S. Lee, and A. Vardar, “Green citrus detection using ‘eigenfruit’, color and circular Gabor texture features under natural outdoor conditions,” *Comput. Electron. Agric.*, vol. 78, no. 2, pp. 140–149, Sep. 2011.
- [26] S. Chaivivatrakul and M. N. Dailey, “Texture-based fruit detection,” *Precis. Agric.*, vol. 15, no. 6, pp. 662–683, Dec. 2014.
- [27] G. Muhammad, “Date fruits classification using texture descriptors and shape-size features,” *Eng. Appl. Artif. Intell.*, vol. 37, pp. 361–367, 2015.
- [28] A. A. Aljanobi, A. S. Al-hamed, and A. S. Al-Suhaibani, “A setup of mobile robotic unit for fruit harvesting,” in *International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD)*, 2010, pp. 105–108.
- [29] A. I. Hobani, A. Thottam, and K. Ahmed, “Development of a Neural Network Classifier for Date Fruit Varieties Using Some Physical Attributes,” *Res. Bull. Agric. Res. Center, King Saud Univ.*, no. 126, pp. 5–18, 2003.

- [30] M. Fadel, "Date Fruits Classification Using Probabilistic Neural Networks," *Agric. Eng. Int. CIGR J.*, vol. IX, no. 1989, 2007.
- [31] A. Haidar, H. Dong, and N. Mavridis, "Image-based date fruit classification," in *International Congress on Ultra Modern Telecommunications and Control Systems*, 2012, pp. 357–363.
- [32] A. Oussama and M. L. Kherfi, "A new method for automatic date fruit classification," *Int. J. Comput. Vis. Robot.*, vol. 7, no. 6, pp. 692–711, 2017.
- [33] M. S. Hossain, G. Muhammad, and S. U. Amin, "Improving consumer satisfaction in smart cities using edge computing and caching: A case study of date fruits classification," *Futur. Gener. Comput. Syst.*, vol. 88, pp. 333–341, 2018.
- [34] A. Kamal-Eldin and S. Ghnimi, "Classification of date fruit (*Phoenix dactylifera*, L.) based on chemometric analysis with multivariate approach," *J. Food Meas. Charact.*, vol. 12, no. 2, pp. 1020–1027, Jun. 2018.
- [35] Z. Schmilovitch, A. Hoffman, H. Egozi, R. Ben-Zvi, Z. Bernstein, and V. Alchanatis, "Maturity determination of fresh dates by near infrared spectrometry," *J. Sci. Food Agric.*, vol. 79, no. 1, pp. 86–90, Jan. 1999.
- [36] D. Zhang, D. J. Lee, B. J. Tippetts, and K. D. Lillywhite, "Date maturity and quality evaluation using color distribution analysis and back projection," *J. Food Eng.*, vol. 131, pp. 161–169, 2014.
- [37] R. Pourdarbani, H. R. Ghassemzadeh, H. Seyedarabi, F. Z. Nahandi, and M. M. Vahed, "Study on an automatic sorting system for Date fruits," *J. Saudi Soc. Agric. Sci.*, vol. 14, no. 1, pp. 83–90, 2015.
- [38] A. Roukhe, B. Abdenabi, and N. EL Barbri, "Sorting dates fruit bunches based on their maturity using camera sensor system," *J. Theor. Appl. Inf. Technol.*, vol. 56, no. 3, 2013.
- [39] K. M. Ismail and K. A. Al-Gaadi, "Development of an electronic sensor for Date sorting based on moisture content," *Misr J. Ag. Eng.*, vol. 26, no. 4, pp. 1923–1932, 2009.
- [40] A. Manickavasagan, N. K. Al-Mezeini, and H. N. Al-Shekaili, "RGB color imaging technique for grading of dates," *Sci. Hortic. (Amsterdam)*, vol. 175, pp. 87–94, 2014.
- [41] A. Al-Janobi, "Application of co-occurrence matrix method in grading date fruits.," *ASAE Annual International Meeting*. Orlando, Florida, USA, 1998.
- [42] A. A. Al-Janobi, "Date inspection by color machine vision," *J. King Saud Univ.*, vol. 12, no. I, pp. 69–79, 2000.
- [43] D. Wulfsohn, Y. Sarig, and R. V. Algazi, "Defect sorting of dry dates by image analysis.," *Can. Agric. Eng.*, vol. 35, no. 2, pp. 133–139, 1993.
- [44] Y. Al Ohali, "Computer vision based date fruit grading system: Design and implementation," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 23, no. 1, pp. 29–

36, 2011.

- [45] A. Nasiri, A. Taheri-Garavand, and Y.-D. Zhang, “Image-based deep learning automated sorting of date fruit,” *Postharvest Biol. Technol.*, vol. 153, pp. 133–141, Jul. 2019.
- [46] N. Alavi, “Quality determination of Mozafati dates using Mamdani fuzzy inference system,” *J. Saudi Soc. Agric. Sci.*, vol. 12, no. 2, pp. 137–142, 2013.
- [47] Min. Agr. Saudi, *The famous date varieties in the Kingdom of Saudi Arabia*, 2nd ed. Riyadh: Ministry of Agriculture, Saudi Arabia, 2011.
- [48] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” *Adv. Neural Inf. Process. Syst.*, pp. 1097–1105, 2012.
- [49] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *ICLR*, 2015.
- [50] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [51] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, Jan. 2015.
- [52] L. Deng, Jia and Dong, Wei and Socher, Richard and Li, Li-Jia and Li, Kai and Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.
- [53] J. Dean *et al.*, “Large scale distributed deep networks,” in *Neural Information Processing Systems*, 2012, pp. 1223–1231.
- [54] J. Donahue *et al.*, “Decaf: A deep convolutional activation feature for generic visual recognition,” in *International conference on machine learning*, 2014, pp. 647–655.
- [55] B. Chu, V. Madhavan, O. Beijbom, J. Hoffman, and T. Darrell, “Best practices for fine-tuning visual classifiers to new domains,” in *European conference on computer vision*, 2016, pp. 435–442.
- [56] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN features off-the-shelf: an astounding baseline for recognition,” in *IEEE conference on computer vision and pattern recognition*, 2014, pp. 806–813.
- [57] A. Salvador, M. Zeppelzauer, D. Manchon-Vizuete, A. Calafell, and X. Giro-i-Nieto, “Cultural event recognition with visual convnets and temporal models,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 36–44.
- [58] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” in *Advances in neural information processing systems*, 2014, pp. 3320–3328.
- [59] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov,

- “Dropout: a simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [60] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “SLIC superpixels compared to state-of-the-art superpixel methods,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [61] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “Slic superpixels,” *Ec. Polytech. Fédéral Laussanne (EPFL), Tech. Rep.*, vol. 149300, pp. 155–162, 2010.
- [62] C. G. Rafael, E. W. Richard, and L. E. Steven, *Digital image processing using matlab*. Tata McGraw-Hill, 2011.
- [63] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 12, pp. 2037–2041, 2006.
- [64] O. Barkan, J. Weill, L. Wolf, and H. Aronowitz, “Fast high dimensional vector multiplication face recognition,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1960–1967.
- [65] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.