HIGH-LEVEL DATA FUSION FOR *'INSIDIOUS FRUIT ROT'* *(IFR)* DETECTION OF *HARUMANIS CULTIVAR*

by

Nurul Syahirah Binti Khalid
1640612295

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Electrical Engineering Technology
UNIVERSITI MALAYSIA PERLIS

2020

ACKNOWLEDGMENT

My utmost gratitude goes to the Creator Ya Wakil Ya Hakim Ya Wahhab - for all the experiences, lessons and gifts in completing my PhD journey. Million thanks and deepest appreciation to my beloved main supervisor, Associate Professor Dr. Shazmin Aniza binti Abdul Shukor ,who have provided me with endless support, guidance, fruitful discussion advice throughout my study. Her mentorship, skills, expertise, dedication, passion and curiosity for researches have inspired me. Many thanks to my co-supervisor , Dr. Fathinul Syahir bin Ahmad Sa'ad for the support during my journey. I would also appreciate to Associate Professor. Ir. Dr. Abu Hassan Abdullah for his first insight and guidance throughout the early days which has helped me to build the foundation in my research work.

My sincere thanks to Postgraduate Society of the School of Mechatronic Engineering, for the discussion on statistical analysis and machine learning, which has helped me to bring the thesis to full completion. My deepest appreciation also to Dr. Maz Jamilah binti Masnan for consistent, valuable guidance and help throughout this study. This study would not have been possible without the financial support and opportunity from the Ministry of Education Malaysia as well as UniMAP.

Special thanks to the Department of Standards Malaysia and Federal Agricultural Marketing Authority (FAMA) Selayang and Perlis for supplying the batches of *Harumanis cv.* samples especially to Encik Rahim, Encik Mazalan, Encik Khairi, Encik Shukri and all other FAMA staff for helping me either direct or indirectly. Thanks to Puan Daleyla and Encik Hafiz from Malaysia Agriculture Research and Development Institute (MARDI) Sintok, Kedah for the knowledge about disorder in *Harumanis cv.*

I am forever indebted to my beloved parents, Khalid bin Ahmad Zahari and Jamaliah bt. Md. Basir for their continuous encouragement and du'a. My humble thanks to all my family members for the endless moral support. Finally, my deepest appreciation and thanks to my beloved husband Ismail Ishaq Ibrahim for his sacrifices, understanding, motivation, du'a and never-ending loves. I hope that this tiny masterpiece would instigate more significance researches for the goodness of mankind. May Allah accept this work as good-deed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

cv.	Cultivar
RM	Ringgit Malaysia
kg	Kilogram
IFR	Insidious Fruit Rot
FAMA	Department of Standards Malaysia and Federal Agricultural Marketing Authority
DOA	Department of Agriculture
MARDI	Malaysia Agriculture Research and Development Institute
NIR	Near-Infrared Spectroscopy
JDL	Joint Directors of Laboratories
FAO	Food and Agriculture Organization of the United Nations
NMR	Nuclear Magnetic Resonance
CT	Computerized Tomography
MRI	Magnetic Resonance Imaging
E-tongue	Electronic tongue
E-nose	Electronic nose
SEE	Standard Error of Estimation
2D	Two Dimension
3D	Three Dimension
DM	Dry Matter
SSC	Soluble Solid Content
VOC	Volatile Organic Compound
LAAS	Laboratoire d'Analyse et d'Architecture des Syst`emes
U.S.	United State
LLDF	Low-Level Data Fusion
ILDF	Intermediate/Mid- Level Data Fusion
HLDF	High-Level Data Fusion
VIS-NIR	Visible Spectroscopy
D-S	Dempster-Shafer
MVT	Majority Voting
HS	Healthy Sample
SG	Specific Gravity
PLS	Partial Least Square

PCR	Principal Component Regression
CCD	Charged Coupled Device
PVC	Polymerizing Vinyl Chloride
MANOVA	Multivariate Analysis of Variance
RGB	Red-Green-Blue
CSSM	Cross Sectional Scanning Method
SSM	Skin Scanning Method
SNV	Standard Normal Variate Transformation
MSC	Multiplicative Scatter Correction
S-G	Savitzky-Golay
SNR	Signal-to-Noise-Ratio
PCA	Principal Component Analysis
MLR	Multiple Linear Regression
LVs	Latent Variables
PC	Principal Component
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error for Cross Validation
RMSEP	Root Mean Square Error of Prediction
SVM	Support Vector Machine
KNN	K-Nearest Neighbour
RBF	Radial Basis Function
SS1	Single-sensor 1 (Image processing technique)
SS3	Single-sensor 2 (NIR spectroscopy)
SS3	Single-sensor 3(E-nose)
Min	Minimum
Max	Maximum
SG	Specific Gravity

LIST OF SYMBOLS

%	Percent
m	Mass
V	Volume
ρ	Density
nm	Nanometer
mL	Mililiter
$\pm$	Plus minus sign
$^{\circ}C$	Degree Celsius
g	Gram
SG	Specific Gravity
ρ_{mango}	Density Mango
ρ_{water}	Density Water
T	Threshold Value
r	Radius
h	Height
V	Volt
W	Watt
mm	Milimeter
R	Pearson
R	Reflectance
K	Kernel Function
g/cm^3	Gram per Cubic Centimeter
V_{sm}	Volume of Sample
V_{am}	Volume of Water after Sample is Submerged
V_{bm}	Volume of Water before Sample is Submerged
π	Pi
R^2	Multiple Correlation
X_{bl}	Baseline Manipulation Data
X_m	Sample Data
X_o	Reference Data

Gabungan Data Tahap-Tinggi Bagi Mengesan “Pereputan Buah Dalaman”(IFR) Bagi Kultivar Harumanis

ABSTRAK

Harumanis cv. dikenali sebagai raja mangga di mana ianya dapat menghasilkan pengeluaran yang banyak setiap tahun tetapi kualitinya perlu dinilai sebelum dijual atau dieksport. Pendekatan sedia ada bagi menilai kualiti buah ini bergantung semata-mata pada pemerhatian manual berdasarkan pengalaman peribadi penilai, di mana hanya parameter luaran sahaja yang diambil kira dan ini boleh menghasilkan ralat disebabkan ketidakjituan oleh pemeriksa manusia. Tambahan pula, cara ini tidak dapat menilai kualiti ciri dalaman *Harumanis cv.* memandangkan *Harumanis cv.* cenderung diserang kecacatan fisiologi yang hanya boleh dikesan melalui pemusnahan. Salah satu kecacatan fisiologi utama adalah kerosakan dalaman yang dikenali sebagai *“Pereputan Buah Dalaman” (IFR)* yang memberi kesan terhadap seluruh pengeluaran. Maka, kaedah tidak musnah dalam penggredan buah dengan mengambil kira kewujudan *IFR* adalah perlu. Kamera *Charged Coupled Device* (CCD), Spektroskopi Infra-Merah Jarak Dekat (NIR) dan Hidung elektronik (E-hidung) adalah teknik analitikal yang digunakan secara berasingan dalam memeriksa kualiti makanan bagi memeriksa graviti khusus dan bau. Dalam kerja ini, gabungan ketiga-tiga jenis penderia ini akan digunakan untuk meningkatkan kepercayaan dalam mengklasifikasikan keadaan *Harumanis cv.* jika dibandingkan dengan bergantung hanya pada sesuatu teknik penderia. Untuk mendapatkan ketepatan penilaian yang tinggi, himpunan data aras-tinggi dicadangkan bagi menggabungkan pelbagai perwakilan menjadi satu dengan menggunakan teknik gabungan *Dempster-Shafer* (D-S) dan *Majority Voting* (MVT). Dua jenis klasifikasi dasar; *Support Vector Machine (SVM)* dan *K-Nearest Neighbours (KNN)* digunakan untuk menganalisa prestasi sesuatu penderia. Eksperimen dijalankan untuk mengesahkan pendekatan yang dicadangkan. Empat peringkat ujian termasuk; pertama, gabungan graviti khusus daripada pemprosesan imej dan spektroskopi NIR; seterusnya, gabungan graviti khusus daripada spektroskopi NIR dan bau daripada E-hidung; ketiga, gabungan aroma daripada E-hidung dan graviti khusus daripada pemprosesan imej; dan akhirnya, gabungan graviti khusus dengan daripada pemprosesan imej, graviti khusus daripada spektroskopi NIR dan bau daripada E-hidung. Keputusan eksperimen menunjukkan prestasi klasifikasi teknik yang dicadangkan memberi prestasi yang lebih baik dibandingkan dengan satu penderia secara am, di mana teknik gabungan yang terbaik adalah D-S bagi graviti khusus daripada pemprosesan imej dan graviti khusus daripada spektroskopi NIR dengan ketepatan 96.87%. Sementara itu, teknik terbaik untuk gabungan kesemua penderia solo menggunakan gabungan D-S (95.80% ketepatan) dibandingkan teknik MVT (87.92% ketepatan). Ini membuktikan *IFR* boleh dikesan secara teknik tidak musnah dengan menggabungkan penderia yang sesuai dan teknik ini diharap dapat memberi faedah dalam membantu petani dan pakar pertanian dalam penggredan dan penilaian dalaman *Harumanis cv.* pada masa hadapan.

High-Level Data Fusion for 'Insidious Fruit Rot'(IFR) Detection of Harumanis cv.

ABSTRACT

Harumanis cv. is known as the king of mangoes where every year, large amount of this mangoes is produced and its quality assessments need to be evaluated before being sold or exported. Existing approach to evaluate the quality of the fruit is solely depending on manual observation based on personal experiences, where only external parameters are considered and this could lead to some errors due to inconsistencies made by human inspection. Furthermore, this approach has problems in evaluating the internal characteristics quality of *Harumanis cv.*, as *Harumanis cv.* are prone to physiological disorder that can only be detected destructively. One of the mainly physiological disorder attack is the internal breakdown known as *'Insidious Fruit Rot' (IFR)* which has affected the overall production of the fruit. Therefore, there is a need to have a non-destructive method in grading the fruit according to its internal quality based on the presence of *IFR*. Charged Coupled Device (CCD) camera, Near-Infrared (NIR) Spectroscopy and Electronic nose (E-nose) are the analytical techniques which have been used separately in the food quality appraisal to measure specific gravity and aroma of fruits. In this work, fusion of these three sensors will be used as a mean to increase the reliability of classification of *Harumanis cv.* condition internally as compared to depending only to a single analytical technique. To get higher accuracy for quality appraisal, high-level data fusion is proposed to fuse these multiple representations using Dempster-Shafer (D-S) and Majority Voting (MVT) fusion techniques. Two types of based classifier; Support Vector Machine (SVM) and K-Nearest Neighbours (KNN) were used to analyze the performance of the single-sensor used. Experimental procedures are then performed to validate the proposed approach. Four testing stages were included; firstly, specific gravity fusion from image processing and NIR spectroscopy; next, specific gravity and aroma fusion from NIR spectroscopy and E-nose; thirdly, fusion of aroma from E-nose and specific gravity from image processing; and lastly, fusion of specific gravity from image processing and NIR spectroscopy with aroma from E-nose. The experimental results show that the proposed method has better performance classification compared with using single-sensor technique in general, where the best fusion method is by applying D-S for specific gravity from image processing technique and specific gravity from NIR spectroscopy with 96.87% accuracy for overall cases. Meanwhile, the best technique in fusing all single-sensors is by applying D-S fusion (95.80% accuracy) compared to MVT method (87.92% accuracy). This proves that *IFR* can be detected non-destructively by fusing related techniques and this method is hope to be beneficial in helping farmers and agricultural experts in grading and evaluating *Harumanis cv.* internally in the future.

CHAPTER 1 : INTRODUCTION

1.1 Research Background

Mango is categorized as one of the most ancient cultivated fruits crops and world's most popular tropical fruits with a rising production trend every year (Ganiron, 2014). Mango or its scientific name *Mangifera indica L.* belongs to dicotyledonous family members of *Anacardiaceae* and taxonomically belongs to the genus *Mangifera*. This tropical fruit has its origin from Indo-Burma region and has been cultivated in India more than 4000 years (State, 2016). Beginning in the 16th century, mangoes were disseminated from India to other tropical countries in Asia such as the Philippines, Indonesia, China and Thailand by sailors, traders and missionaries. In 1993, Galan (1993) reported that the production of mango worldwide has increased by approximately 50 percent between 1971 and 1993.

Mango has been an enjoyable aromatic and flavour fruit, with exclusive taste and contains rich vitamin A compared to other fruits (Nip, 1991). It is mostly eaten up raw as a dessert fruit and small bulk are also prepared into mango juice, jellies, jams, nectars and preserves. There are more than a thousand varieties of mango under cultivation but only a small number of cultivar *(cv.)* are grown on a commercial scale. For the time being, mango is an important commercial crop not only in India, but also in Indonesia, Thailand as well as in Malaysia. According to Vasanthaiah, Ravishankar, & Mukunda (2007), the highest concentration of mango species is reported in Malayan Peninsular, followed by the Sunda Islands and the Eastern Peninsula with many of the genus being common between of them. Figure 1.1 presents the spread of mango from the original region cultivation to Malaysia and around the world.

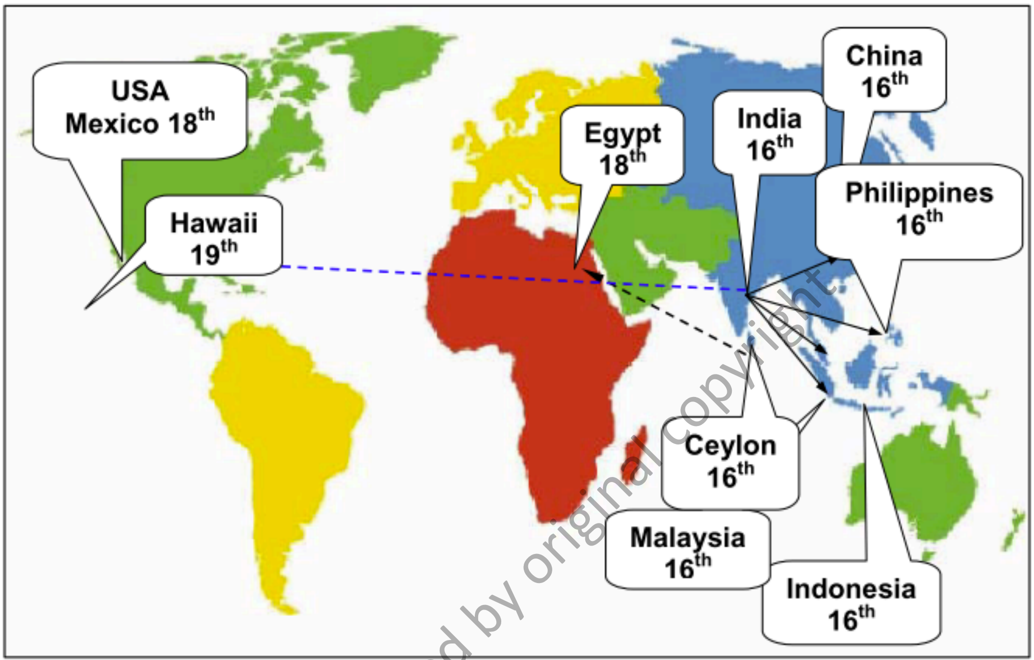

Figure 1.1 The spread of mango from India to another country (Harbinson & Kooten, 2005)

In Peninsular Malaysia, mango is grown in a limited, mixed property or orchard. According to "Malaysia Fruit Crops Statistics" report produced by Department of Agriculture (DOA) Putrajaya, Malaysia in 2017, the area of mango orchard in Peninsular Malaysia is 6048.29 hectarage. Today, there are over 300 mango cultivars with considerably in different size, shape, colour, flavour and fibre contents in Malaysia. A good mango is juicy, sweet and may also have its own attractiveness and variation.

One of the well-known mango cultivar that has its own unique characteristics and has been commercially exploited is *Harumanis cv.* (Faridah, Rosidah, & Jamaliah, 2010). *Harumanis cv.* is also known as the ''King of Mangoes'' and very well-liked in Malaysia due to its deliciousness, sweetness and aromatic fragrance. This cultivar is distinctive to Perlis and known in the national agenda as a specialty fruit from Perlis for the world with the total yield of RM41,986,500 with 1825.5 tons in 2017 (Malaysia Fruit Crop Statistics, 2017). As stated by Information Marketing Department of FAMA in 2019, this fruit sales have been reached up to an average of RM38.65 per kilogram in 2016 (Information Marketing Department of FAMA, 2019). Perlis' soil and climatic conditions makes it extremely suitable for *Harumanis cv*. It requires a significant dry weather period to trigger the flowering stage and the productive phase can be significantly affected by changes in weather. *Harumanis cv.* is a yearly fruit bearing tree type and reproductive phase of the *Harumanis cv.* trees usually begins in January and ends in early June. This cultivar is very suitable for the export market as it has favourable colour, sweetness and excellent eating quality with good aroma. Table 1.1 shows number of exported *Harumanis cv.* from Perlis to Japan in 2010 which is aimed to increase at 100 metric tons in 2020 (Farook et al., 2011) and this obviously revealed that *Harumanis cv.* has a high demand in market. As export of fruits is incredibly dependent on quality, the government has constructed a rigid quality control regulation for the export of this fruit since poor conditions of fruits may lead to the rejection and definitely causes market loss (Jha et al.,2012). As a consequence, quality control of *Harumanis cv.* plays important role in deciding the export value of the fruit to the world.

Table 1.1 The target and actual export of *Harumanis cv.* mango to Japan (Farook et al., 2011)

Year	Target	Export
2010	3.1 metric tons	500kg
2011	3.1 metric tons	2.4 metric tons
2020	100 metric tons	?

Considered as an exotic fruit and as a perishable product, a good quality *Harumanis cv.* is highly desired and yield a valuable price in the world market. However, they are prone to pests, disease and physiological disorder attack. According to the 2017 report by Malaysia Agriculture Research and Development Institute (MARDI), the expansion of this cultivar is reportedly hindered due to the incidence of internal tissue breakdown or commonly known as '*Insidious Fruit Rot' (IFR)* as shown in Figure 1.2. This physiological disorder was reported in Perlis where losses amounting to 80% have been recorded by growers, which resulted in poor fruit quality and thus reduced marketability. Thus, proper non-destructive approaches that able to detect the existence of this internal disorder are much appreciated, which normally requires information from more than one independent sensors. Hence, the possibility of using a multi-sensory data fusion approach in the same practice as the human brain combines the information resulting from multiple senses will be investigated in order to gain a more accurate results. Figure 1.3 shows *Harumanis cv.* sample in good sensory condition.

Figure 1.2 *Harumanis cv.* affected by *IFR*

Figure 1.3 Healthy *Harumanis cv.* sample