




Destructive and non-destructive measurement approaches and the application of AI models in precision agriculture: a review

Maidul Islam^{1,5} · Suraj Bijjahalli² · Thomas Fahey^{1,5} · Alessandro Gardi^{1,3,5}  · Roberto Sabatini^{1,3,5} · David W. Lamb^{4,5}

Accepted: 6 January 2024 / Published online: 27 February 2024
© The Author(s) 2024

Abstract

The estimation of pre-harvest fruit quality and maturity is essential for growers to determine the harvest timing, storage requirements and profitability of the crop yield. In-field fruit maturity indicators are highly variable and require high spatiotemporal resolution data, which can be obtained from contemporary precision agriculture systems. Such systems exploit various state-of-the-art sensors, increasingly relying on spectrometry and imaging techniques in association with advanced Artificial Intelligence (AI) and, in particular, Machine Learning (ML) algorithms. This article presents a critical review of precision agriculture techniques for fruit maturity estimation, with a focus on destructive and non-destructive measurement approaches, and the applications of ML in the domain. A critical analysis of the advantages and disadvantages of different techniques is conducted by surveying recent articles on non-destructive methods to discern trends in performance and applicability. Advanced data-fusion methods for combining information from multiple non-destructive sensors are increasingly being used to develop more accurate representations of fruit maturity for the entire field. This is achieved by incorporating AI algorithms, such as support vector machines, k-nearest neighbour, neural networks, and clustering. Based on an extensive survey of recently published research, the review also identifies the most effective fruit maturity indices, namely: sugar content, acidity and firmness. The review concludes by highlighting the outstanding technical challenges and identifies the most promising areas for future research. Hence, this research has the potential to provide a valuable resource for the growers, allowing them to familiarize themselves with contemporary Smart Agricultural methodologies currently in use. These practices can be gradually incorporated from their perspective, taking into account the availability of non-destructive techniques and the use of efficient fruit maturity indices.

Keywords Digital farming · Artificial intelligence · Unmanned aircraft system · Fruit maturity · Sensor fusion · Field robotics

Introduction

Over the last few decades, agriculture has progressively evolved from an experience-based and subjective process, to the present state, which is characterised by an increasing use of sensors and data-driven decision-making. According to the analysis performed by the Food and Agricultural Organization (FAO) in 2017, the world population will reach 9.6 billion by 2050 (Trilles et al., 2020). To feed this growing population, the global food production must be increased by 50%. On the other hand, the practice of agricultural intensification can have a substantial impact on the environment, resulting in the deterioration of soil quality caused by erosion from wind and water, as well as the pollution of air and water owing to excessive use of nutrients and agrochemicals. It may also result in a loss of biological and ecological diversity. To address the negative impacts of highly productive but intensive farming, it is crucial to transition towards more environmentally friendly and sustainable agricultural practices (Issad et al., 2019). Precision Agriculture (PA) has emerged as one of the most promising solutions, since it predicated the use of cutting-edge technologies such as proximal and remote sensing, automation and control systems, power and data management. This domain has extended its reach to encompass a broad spectrum of applications, including but not limited to crop disease monitoring, pesticide control, irrigation and water management, storage management, and weed and soil management (Jha et al., 2019). The main goal of PA is to optimize the available resources in order to achieve sustainable production with the lowest cost (Alreshidi, 2019; Issad et al., 2019). The International Society for Precision Agriculture (ISPA) initiated a year-long endeavor that involved the contributions of 46 experts in the field of precision agriculture, culminating in the formulation of the subsequent definition:

Precision Agriculture is a management strategy that gathers, processes and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability and sustainability of agricultural production. (Andujar, 2023).

In short, PA is, by its definition, data rich. Yet, much of practical PA today still relies upon human decision making. The role and influence of increasing amounts of available data in that decision making can be greatly enhanced by the implementation of Artificial Intelligence (AI) into the decision process (Alreshidi, 2019). AI algorithms allow a machine to emulate humans in solving complex problems for which analytical solutions are difficult or impossible to derive (Pathan et al., 2020). For instance, in agriculture, AI can be trained utilizing sensor data to estimate and interpret physiological or organoleptic changes in crops/fruits and at the same time, classify them according to their physiological properties. Hence, its predictive capabilities enable the prediction of crop yields, disease outbreaks, and optimal planting times, resulting in more informed decision-making. AI has nowadays achieved a large success in different applications (e.g. healthcare (Yu et al., 2018), business (Akerkar, 2019), transportation (Machin et al., 2018), education (Roll & Wylie, 2016), robotics and automation (Perez et al., 2018; Wisskirchen et al., 2017)) and has been demonstrated in agricultural robotics, optimization management, automation, knowledge-based systems, expert systems and decision support systems (Alreshidi, 2019).

Overview of fruit maturity estimation

The estimation fruit quality and maturity, and the subsequent fruit classification steps are considered essential parts of the harvest process. Accurate maturity estimation and classification methodologies mitigate the risk of economic loss due to untimely harvesting. Traditionally, destructive approaches have been used to identify the maturity of the fruit, but this is not a very efficient approach since the fruit is destroyed in the process. For example, in Iran, many fruits, such as hawthorn and mulberry, are harvested either manually by shaking the tree branches or mechanically using a shaker. However, this harvesting process can lead to variations in the ripeness levels of the fruits as some hawthorn fruits may be unripe, some ripe, and some overripe. One possible reason for this variation is the irregular flowering time of hawthorn trees, which causes the fruits to ripen at different times. The uneven ripeness levels can negatively impact the marketability of the harvested hawthorns, leading to a decline in economic value and an increase in product waste. In Iran, annually around 20 million tons of these agricultural products go to waste during the post-harvest stage. This amount of waste is equivalent to the food that 20 million people could consume in a year (Azadnia et al., 2023).

Moreover, the traditional process must be frequently repeated with many samples to get representative values for the entire orchard. In effect, destructive testing provides a very selective view of the broader field as there is high variability between fruit maturity from cultivar to cultivar and from plant to plant. One alternative is organoleptic fruit maturity estimation by an experienced grower with years of accumulated expert knowledge. However, expert knowledge is difficult to acquire, difficult to communicate (or translate) to others and hence scale, and therefore cannot be widely applied. Active and passive remote sensing technologies have advanced considerably to the point where reasonably inexpensive sensors can be easily deployed in fields. There is now growing use of AI to process remote sensing measurements and infer maturity levels of fruit to save time, reduce cost and avoid laborious inspection processes. As a result, non-destructive approaches emerge as a more promising PA technique in which sensor data and AI are combined. These sensors measure the fruit maturity indices externally and provide this information to AI models to estimate the fruit maturity and quality and classify accordingly. The bulk of AI algorithms used for fruit quality and maturity estimation belong to the Machine Learning (ML) category. ML methods are designed to ‘learn’ non-linear mappings between input measurements and output inferences that are not observable directly. Some commonly considered ML approaches for estimation are Multiple Linear Regression (MLR), Partial Least Square (PLS) and Principal Component Regression (PCR) while Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forests, K-nearest neighbours (KNN), K-means clustering and Neural Networks (NN) widely adopted ML approaches to sorting out fruits. “[Sensor fusion and artificial intelligence for sensor data processing](#)” section extensive describes the ML approaches that have been adopted for both the classification and estimation of fruit maturity.

Overview of machine learning applications in smart agriculture

ML is popular especially because it does not require extensive analytical modelling (Liaos et al., 2018) and it is rather simple to setup a computer to learn from structured data. In agriculture, the data can be supplied from local sensors data storage, cloud based data

sources that accumulates soil properties, temperature or humidity from multiple sensors independently, or any agricultural company or organization like Food and Agriculture Organization (FAO). To infer the fruit quality and maturity, a computer must receive a large volume of data from multiple sources. ML algorithms extract patterns from data based on user-defined labels during the training process. User-defined labels provide description of data, categorized in a dataset which are set by the user. For example, while developing a ML model for categorizing maturity, a user would supply a dataset of numeric values of multiple maturity indexes. Each level of maturity would be accompanied by a label indicating the specific type of maturity such as ‘immature’, ‘mature’, or ‘ripe’. Therefore, understanding features and metrics that can be extracted from data is crucial since this step is the foundation for the estimation of the fruit quality and maturity. In the case of fruit maturity level identification, ML algorithms are expected to identify features from the dataset, such as maturity indexes and determine the ranges of the maturity indexes according to user-defined labels, whereas metrics interpret the accuracy in determining the ranges. ML approaches in agriculture have been reviewed by several researchers. For instance, Liakos et al. (2018) classify and briefly discuss the ML approaches and show their applications to four major aspects of agriculture, namely livestock management, crop management, soil management and water management. Under crop management, they discuss briefly the application of ML in yield prediction, weed detection, disease detection, and crop quality and species detection. Ip et al. (2018) highlight the applications of ML in the area of crop protection that covers plant disease, weed and pest management.

Aim and structure of the article

In recent years, there has been considerable work on the use of ML for the detection of fruit maturity and quality, both in-situ and in sorting and packing sheds. In order to use ML approaches to estimate fruit maturity, suitable datasets are required that comes from destructive and non-destructive sensors. These sensors primarily measure some maturity indexes of the fruits. Therefore, finding the suitable maturity indexes of a particular type of fruit is challenging since each type of fruit exposes its physical and organoleptic changes in its unique way. Apart from that, some of these maturity indexes can only be measured considering either destructive or non-destructive approaches. Hence, an extensive study is required that can outline the important maturity indexes and their measuring approaches so that researchers in agriculture can take decisions suitably in their future works. Besides, another challenge may appear when it is a matter of considering suitable ML approaches in order to interpret different type of datasets where some ML approaches need to be introduced with features while others not. It is worth mentioning that sensor data type varies from sensor to sensor. To address this challenge, the strength and limitations of these ML approaches should be discussed and analyzed, particularly for this application of fruit maturity estimation. Additionally, the interpretability and data sensitivity for the ML algorithms have been addressed that play significant roles in terms of understanding how the maturity indexes are interpreted by the algorithms in this application. Thus, the ML users can consider the features during model development and utilize the algorithms correctly.

Considering a large number of fruit maturity indexes, discussed in literature, this study outlines the common and effective maturity indexes to offer readers a clear understanding of them. Apart from that, this study significantly contributed to developing a framework about the applicability of different destructive and non-destructive methods in measuring different fruit maturity. Hence, the destructive and non-destructive approaches have been

classified and comprehensively described that can help a reader to consider suitable fruit maturity testing methods. In addition, this study seeks to review the work on fruit maturity estimation using ML approaches, including a comprehensive critical review of maturity indexes is presented to inform the potential future design of novel ML models. Moreover, this review contributes by analysing commonly used ML approaches, particularly in fruit maturity level estimation and their corresponding strengths and limitations. Finally, the article shows the trend of currently considered non-destructive approaches (i.e., spectral imagery and spectroscopy) in estimating fruit maturity and suitable ML approaches that can interpret the sensors data suitably.

This article is organized as follows. “[Fruit maturity indexes and testing methods](#)” section reviews commonly considered maturity indexes and their derivation from in-field measurements. Therefore, it also offers an idea of widely applied destructive and non-destructive approaches to measure the maturity indexes. “[Sensor fusion and artificial intelligence for sensor data processing](#)” section focuses on a comprehensive review of AI approaches with sensor fusion in fruit maturity, beginning with AI fundamentals, followed by a comparative analysis of their advantages and limitations in the context of estimating and/or predicting fruit maturity. “[Challenges and future research and development trends](#)” section highlights several challenges pertaining to ML algorithms, maturity indexes and sensors are identified. Additionally, several promising areas of future work are identified. This section also demonstrates the trends in remote sensors and ML algorithms to give a conclusive idea about their directions and progress.

Fruit maturity indexes and testing methods

Maturation is a stage of fruit development that refers to the achievement of either physiological or horticultural maturity. Physiological maturity defines the absolute biological stage of fruit development whereby fruit can carry on the process of ontogeny (i.e. further development) once detached from the host plant (Kader, 1997). Noted that non-climacteric fruits (e.g. grapes, cherries, citrus) do not continue to ripen after picking whereas climacteric fruits (apples, pears, peaches) continue to ripen. Horticultural maturity, on the other hand, is the stage of fruit physiological development associated with meeting specifications as defined by processors or consumers (Brovelli & Cisneros-Zevallos, 2007). This section addresses commonly considered maturity indexes and their measuring approaches, including their classifications.

Fruit maturity indexes

In this review, maturity indexes have been classified into two categories based upon observability; namely internal and external observables. Figure 1 illustrates the classification of fruit maturity indexes, described in detail below and in the subsequent sections of the article. Internal observables include the fruit texture and composition metrics such as moisture, dryness, firmness, sugar, acidity, starch, flesh colour and oil content. Some other internal observables such as specific gravity and heat capacity are also considered in a few works to define the fruit maturity but they have been omitted from this review since they are not broadly considered for a number of fruits.

On the other hand, external observables such as fruit size, shape, weight, skin colour, chlorophyll content and number of days after full bloom have been discussed extensively

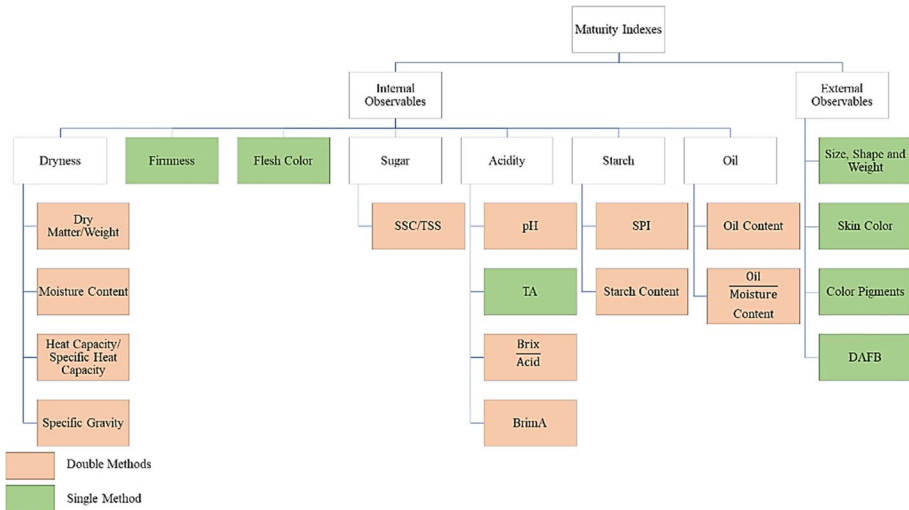


Fig. 1 Maturity Indexes classification based on the observables

since they are commonly considered in horticulture, particularly for horticultural maturity. Conversely, leaf changes, aroma other colour pigments (e.g., chlorophyll content, anthocyanin and carotenoid) have not been explored in full detail since their application is limited to certain fruits. The maturity indexes can be measured by destructive, non-destructive or both methods. Indeed, most of them can be measured by both methods, i.e. ‘double methods’, while some can only be measured by a single method. Table 1 summarises the main observables, maturity indexes and their measuring approaches.

Sugar concentration

In general, both in climacteric fruits (e.g., apple, pears, plums, avocados etc.) where starch starts to break down to sugars when the fruit starts to ripen, and in non-climacteric fruits (e.g., strawberries, grapes, cherry, citrus fruits etc.) where sugars are accumulated during the process of maturation (Prasad et al., 2018), sugar concentration is considered a key indicator of fruit maturity. However, instead of being climacteric fruits, summer fruits such as peach, nectarine, plums and apricots continue to ripen after harvest. Hence, off and on its own, sugar concentration cannot be used to estimate, in an absolute sense, the ‘maturity’. Instead, other maturity indicators such as firmness, size, colour, aroma etc., are also required to be taken into consideration at the same time in order to be confirmed about the maturity level of these fruits. Hence, multiple sensor data can be fused altogether using ML approaches. Noted that Brix is the technical term that is used for both total soluble solids (TSS) and SSC, albeit Brix only refers to the sugar contents of fruits (Magwaza & Opara, 2015). Degree Brix ($^{\circ}$ Brix) is the unit of measure that defines the gram of sucrose equivalents per 100 g of solution, i.e. juice. The soluble sugars include sucrose, glucose and fructose and are, in general, measured by destructive methods like gravimetric methods (using hydrometer) (Nor et al., 2014), refractometry (Harrill, 1998), High-Performance Liquid Chromatography (HPLC) (Richmond et al., 1981) and potentiometric chemical sensors (also known as electronic tongues) (Beullens et al., 2008). Since these methods are time-consuming,

Table 1 Maturity indexes and associated measuring approaches

Observables	Derived maturity indices	Relevant equation(s)	Destructive approaches/instruments	Non-destructive approaches/instruments
Sugar Contents	TSS, SSC	N/A	Hydrometer (Nor et al., 2014), Refractometer (Harrill, 1998), HPLC (Richmond et al., 1981), Electric tongue (Beullens et al., 2008)	Hyperspectral imaging (Sugiyama & Tsuta, 2010), vis-NIRS (Jie et al., 2014), NMR (Zhang & McCarthy, 2013; Zion et al., 1995), ATR-FTIR (Bureau et al., 2009), dielectric spectroscopy (Castro-Giráldez et al., 2010)
Acidity	pH, TA, Brix/Acid, BrimA,	$\text{BrimA} = \text{SSC} - k \text{ (TA)}$ $\text{TA}_{\text{wt/vol}} (\%) = \frac{N \cdot V_1 \cdot E_{\text{wt}}}{V_2 \cdot 100} \cdot 100$ $\text{TA}_{\text{wt/wt}} (\%) = \frac{N \cdot V_1 \cdot E_{\text{wt}}}{W \cdot 100} \cdot 100$	pH meter/instrument (Chassagne-Berces et al., 2010, acid-base titration (Kafkas et al., 2007), HPLC test (Fawole & Opara, 2013; Moing et al., 1998), voltammetry (Kotani et al., 2019) and pH meter test (Lebrun et al., 2008)	NIR Hyperspectral Imaging (Li et al., 2018a, 2018b)
Firmness	N/A	N/A	Magness–Taylor test (Jantra et al., 2018), electronic pressure tester (Lehman-Salada, 1996) and Effegi firmness tester (Harker et al., 1996)	Mechanical resonance (Goliáš et al., 2003), vibrational excitation (Peleg et al., 1990), reflectance spectrometry (Nicolai et al., 2008; Paz et al., 2008), impact force response/velocity (Wang et al., 2009) and chlorophyll fluorescence (Bron et al., 2005)
Starch content	SPI or Starch Index	N/A	Iodine potassium-iodide (I ₂ -KI) solution (Blankenship et al., 1993; Brookfield et al., 1997; Zude et al., 2006)	NIR spectroscopy (Peirs et al., 2003)

Table 1 (continued)

Observables	Derived maturity indices	Relevant equation(s)	Destructive approaches/instruments	Non-destructive approaches/instruments
Solid content	DM, TSS, Moisture Content	$DM (\%) = \frac{\text{dry sample mass}}{\text{fresh sample mass}} \cdot 100$ $\text{Moisture Content } (\%) = \frac{\text{fresh sample} - \text{dry sample}}{\text{fresh sample}} \cdot 100$	Oven drying (Doerflinger et al., 2015a; Palmer et al., 2010)	NIR (McGlone et al., 2003; Travers et al., 2014), SWNIR (Subedi & Walsh, 2011), NMR (Chen et al., 1993), Ultrasonic attenuation (Mizrach et al., 1999) for DM, NIR (Blakey & Van Rooyen, 2011), dielectric method (Gun & Chen, 2010) for Moisture Content
	Oil contents	$\text{Oil Content} = \text{dry matter } (\%) \cdot \frac{\text{oil weight}}{\text{dry pulp weight}}$	Soxhlet extractor (Magwaza & Tesfay, 2015; Stefanoudaki et al., 1999), refractometer (Lavee & Wodner, 2004) HPLC (Lichenthaler, 1987)	NMR (Magwaza & Tesfay, 2015), vis-NIRS (Ncama et al., 2018), laser-light backscattering (Ali et al., 2020) DA meter (Cocetta et al., 2017; Zhang et al., 2020), vis-NIR spectrometer (Infante et al., 2011; Zhang et al., 2017), Raman spectroscopy (Trebolazabala et al., 2017)
Chloro-phyll contents	Index of Absorbance	$I_{AD} = A_{670} - A_{720}$		

manual, laborious and destructive, a few non-destructive methods have emerged as alternatives (Jie et al., 2014). Several spectroscopic methods such as hyperspectral imaging (Sugiyama & Tsuta, 2010), visible to near-infrared (vis/NIR) spectroscopy (Jie et al., 2014), nuclear magnetic resonance (NMR) (Zhang & McCarthy, 2013; Zion et al., 1995), and Fourier transform NIR spectroscopy with attenuated total reflection (ATR-FTIR) (Bureau et al., 2009) are frequently used non-destructive methods to determine the sugar contents of fresh produce. Dielectric spectroscopy is another non-destructive method that is also found in some literature (Castro-Giráldez et al., 2010).

Acid contents

Changes in acidity are significant for several fruits during their maturation and ripening. In both climacteric (Pech et al., 2008) and most of the non-climacteric fruits (Batista-Silva et al., 2018), acidity is an observable indicator of maturation. Unlike sweetness, acidity declines gradually as the fruit reaches maturity although the rate of acidity changes is affected by the cultivars and season. The acidity of fruits are primarily measured considering titratable acidity (TA) and pH, and both of them provide a unique insight into fruit quality (Tyl & Sadler, 2017). TA can be calculated using either Eqs. (1) or (2) below. Equation (1) calculates the percentage of an acid with respect to weight, and Eq. (2) measures it with respect to volume:

$$TA_{wt/wt}(\%) = \frac{N \cdot V \cdot Eqwt}{W \cdot 100} \cdot 100 \quad (1)$$

where N, V, Eq.wt and W define the normality of titrant (usually NaOH), the volume of titrant, equivalent weight of the acid and mass of the fruit.

$$TA_{wt/vol}(\%) = \frac{N \cdot V_1 \cdot Eqwt}{V_2 \cdot 100} \cdot 100 \quad (2)$$

where V_1 and V_2 symbolize titrant's volume and sample's volume, respectively.

Apart from TA, the Brix-Acid ratio is a desired metric, often in preference over Brix or acid alone is and in some instances is considered an objective measurement that reflects the consumer acceptability (Tyl & Sadler, 2017). The measurement of Brix includes the sugars, acids and other acid components and hence acid components is often considered at the numerator. As a result, Brix-Acid ratio may have a chance of lower reliability in terms of defining fruit maturity (Magwaza & Opara, 2015). For instance, Zhang et al. (2020) consider six (6) different cultivars of peach of fifty (50) samples to predict their maturity relying on multiple maturity indexes such as IAD, skin colour, firmness, extractable juice, TA, SSC, and Brix-Acid ratio. Among them, IAD and firmness prove their reliability in interpreting the maturity while SSC, TA and Brix-Acid ratio are unable show considerable reliability in terms of estimating peach maturity.

In order to measure the pH of fruits, a pH meter or pH instrument (Chassagne-Berces et al., 2010) is used. While TA is measured, in general, using acid–base titration (Kafkas et al., 2007), other instrumental approaches such as HPLC (Fawole & Opara, 2013; Moing et al., 1998), voltammetry (Kotani et al., 2019) and pH meters (Lebrun et al., 2008) also have drawn the attention of the researchers. Near-infrared (NIR) hyperspectral imaging can be considered as a new promising non-destructive technique to measure the pH efficiently (Li et al., 2018a, 2018b).

Firmness

The firmness of fruit flesh decreases with the ripening process, and hence overripe fruits are relatively soft (Jarimopas & Kitthawee, 2007). Therefore, it is considered as one of the important maturity indicators to classify fruit according to their maturity levels (Macrelli et al., 2013). Several types of penetrometers are widely used to measure fruit firmness including the Magness–Taylor tester (Jantra et al., 2018), electronic pressure tester (Lehman-Salada, 1996) and Effegi firmness tester (Harker et al., 1996). Zahed Fathizadeh et al. (2021) classify the non-destructive approaches to measure the apple firmness into three (3) categories based on their working principles and extensively discuss them as follows:

- (1) Acoustic, vibration and mechanical response (i.e., mechanical resonance, vibrational excitation, impact force response/velocity, ultrasonic methods and intelligent firmness detector);
- (2) Optical methods (i.e., time resolved and NIR spectrometry, chlorophyll fluorescence and laser excitation); and
- (3) Electrical resistance and capacity measurements.

Although this particular study focuses on the non-destructive tests of firmness of apples, these methods are applicable to other fruits as well.

Starch contents

During active growth, many fruits accumulate carbohydrates that are stored in starch form utilizing photo-assimilates. As they approach maturity in the trees, starch starts to be hydrolysed (Fuchs et al., 1980) and turns into soluble carbohydrates (Doerflinger et al., 2015b). Once the fruits are detached from the trees, the starch begins to transform into soluble sugars in order to contribute to the metabolic processes of fruit ripening (Doerflinger et al., 2015b). Thus the fruits lose starch gradually with the ripening process (Blankenship et al., 1993; Doerflinger et al., 2015b). Interestingly, the starch releases sugars like glucose, sucrose and fructose that are responsible for the respiration and enhancement of sweetness of the fruits (Doerflinger et al., 2015b). Therefore, for starchy fruits such as apple, mango, avocado, banana, kiwi etc., starch content is considered an effective maturity index. Starch pattern index (SPI) (Doerflinger et al., 2015a), SSC (Subedi & Walsh, 2011), and dry matter content (DMC) (Palmer et al., 2010) are some maturity indexes that are significantly influenced by the starch content of the fruits. Starch pattern index (SPI) defines the starch degradation in fruits (Brookfield et al., 1997) by staining with an iodine potassium-iodide solution. It is a widely applied destructive approach (Blankenship et al., 1993; Brookfield et al., 1997; Zude et al., 2006). Peris et al. has introduced NIR spectroscopy to determine the starch index in apples and achieved satisfactory results in determining the maturity stage of the fruit though it is a time-consuming approach (Peirs et al., 2003).

Dry matter contents

Dry matter is a measure of carbohydrate content and hence it is related to the starch content. It is also significantly related to moisture content. Dry matter increases with the fruit

growth and development and declines with fruit maturity. Conversely, moisture content increases when dry matter starts to decline at the very beginning of fruit ripening (Olareswaju et al., 2016). Accordingly, dry matter and moisture content are increasingly considered as two effective maturity indexes to determine the fruit maturity stage that will define the harvest time (Magwaza & Tesfay, 2015; Rajkumar et al., 2012). In general, a sample of the pulp of a starchy fruit is oven dried in order to measure the weight of dry matter or dry matter concentrations (Doerflinger et al., 2015a; Palmer et al., 2010) which is a destructive approach. McGlone et al. (2003) established a mathematical equation for DM with starch content as follows:

$$DM = SS + St + NS \quad (3)$$

where SS, St and NS are respectively the percentages (w/w) of water, soluble solids and starch and non-starch insoluble solids. In addition, Magwaza and Tesfay also introduced two equations to calculate DM and Moisture Content as follows (Magwaza & Tesfay, 2015):

$$DM (\%) = \frac{M_d}{M_f} \cdot 100 \quad (4)$$

$$\text{Moisture Content } (\%) = \frac{M_d - M_f}{M_f} \cdot 100 \quad (5)$$

where M_d and M_f define the mass of the dry sample and the mass of the fresh sample.

In general, non-destructive approaches such as NIR (McGlone et al., 2003; Mishra & Woltering, 2023; Travers et al., 2014), SWNIR (Subedi & Walsh, 2011), NMR (Chen et al., 1993), ultrasonic attenuation (Mizrach et al., 1999) for DM, NIR (Blakey & Van Rooyen, 2011) and dielectric method (Gun & Chen, 2010) for Moisture Content can be taken into consideration in pre-harvest to estimate fruit maturity.

Oil content

Oil content is considered one of the crucial factors along with dry matter and moisture content to estimate the maturity of fruits like avocado (Magwaza & Tesfay, 2015), palm and olive (Guzmán et al., 2015). In the case of palm, oil content does not increase after harvesting (Saeed et al., 2012), and hence, it is very crucial to decide the suitable palm in situ during harvesting. In general, oil content increases with the decrease of water (i.e. moisture contents) when fruits are in the trees. Since the maturity of palm is significantly correlated with its texture and colour, chlorophyll content is considered as one of the factors to estimate its maturity that can lead to estimate the oil content in palm fruits (Balasundram et al., 2006). Magwaza and Tesfay (Magwaza & Tesfay, 2015) have introduced an equation to measure the Oil Content as follows:

$$\text{Oil Content } (\%) = \text{drymatter}(\%) \cdot \frac{\text{oilweight}}{\text{drypulpweight}} \quad (6)$$

The Soxhlet extractor (Magwaza & Tesfay, 2015; Stefanoudaki et al., 1999), refractometer (Lavee & Wodner, 2004) and single-outlet piston press (Tantanawat et al., 2020) are examples of destructive methods used to measure oil content; the Soxhlet extraction method being the most popular approach. Alternatively, NMR (Magwaza & Tesfay, 2015), vis-NIRS (Ncama

et al., 2018) and laser-light backscattering (Ali et al., 2020) are examples of non-destructive methods.

Chlorophyll content

Chlorophyll content degradation in fruit can be considered one of the efficient maturity indicators (Zhang et al., 2020) and even can be better in indication than starch, firmness and SSC (Costamagna et al., 2012). The ripening process brings various physiological and biochemical changes in the fruits (Kasampalis et al., 2020), and hence, it is responsible for the degradation of chlorophyll in the fruits. For instance, chlorophyll shows strong correlation with firmness in apple (Song et al., 1997) and peach (Zhang et al., 2020) and with skin colour of mango (Jacobi et al., 1998) in their ripening process. To measure the chlorophyll content, index of absorbance difference (IAD) has been introduced in the literature that degrades with the degradation of chlorophyll content (DeLong et al., 2014). Fruit absorbance (A) can be expressed using Lambert's Beer Law as stated (Ziosi et al., 2008):

$$A = \log_{10} I^{-1} \quad (7)$$

where I is the interactance spectra.

Hence, index of absorbance can be measured using the formula as stated (DeLong et al., 2014; Ziosi et al., 2008):

$$I_{AD} = A_{670} - A_{720} \quad (8)$$

where A_{670} and A_{720} are the A values at 670 nm and 720 nm wavelengths.

Chlorophyll content can also be measured using both destructive and non-destructive methods. Chromatography is the only destructive method that has been used to measure the chlorophyll content (Lichtenthaler, 1987). Delta Absorbance (DA) meter (Cocetta et al., 2017; Zhang et al., 2020) and Vis–NIR spectrometer (Infante et al., 2011; Zhang et al., 2017) particularly maintain a similar working principle considering light absorbance and help to find out the chlorophyll content through IAD. Raman spectroscopy (Trebolazabala et al., 2017) and chlorophyll fluorescence (Kolb et al., 2006) are two other non-destructive approaches to measure the chlorophyll content in the fruits though they are not used for measuring I_{AD} .

Size, shape and weight

Reaching a certain size and weight can be considered a possible maturity index, but it certainly can be alone to declare the maturity of a fruit (Crisosto, 1994). In general, the size and weight of fruit according to its maturity level can vary from another because of its cultivar (Cheng & Breen, 1992), cultivation practice, climate conditions, region of cultivation etc. (Crisosto, 1994). In addition, fruit shoulder is remarkably highlighted as an indicator to assess the maturity of a fruit (Pereira et al., 2009; Sahu & Potdar, 2017). However, to rely upon these maturity indexes, additional maturity indexes are highly recommended to be coupled with them, such as skin colour (LaRue, 1989; Prabha & Kumar, 2015).

Skin colour

Skin colour is considered as one of the indicators of fruit maturity. Hence, fruits like bananas (Prabha & Kumar, 2015), tomatoes (Wan et al., 2018), dates (Lee et al., 2008),

plums (Kaur et al., 2018) etc., in general, expose their colours vividly during their maturity process. Noted that the colour change also varies from cultivar to cultivar, where some cultivars are very responsive with their maturity through skin colour while some other cultivars are not such responsive (Prasad et al., 2018). However, classifying fruits according to their maturity using camera have received tremendous interest since the computer vision system has significantly developed, become cheaper and available. Computer vision particularly works relying on the skin colour of the fruits. Hence, recently, skin colour has become a popular maturity index to segregate mature fruit from immature though relying upon skin colour completely to identify a mature fruit is not certainly a practical approach always. However, in some cases, like nectarines, peaches or apples, skin colour is highly responsive to the sunlight that leads to reach a wrong conclusion about the maturity of a fruit (Iglesias et al., 2008; Prasad et al., 2018). Therefore, particularly changes in ground colour (Mitchell et al., 1977) that is not much influenced by sunlight or flesh colour (Josan & Chohan, 1982) in fruit can be comparatively a better indicator than exposed surface colour (Prasad et al., 2018).

Days after full bloom

Days after Full Bloom (DAFB) is also considered as a maturity index for fruit though it is not as efficient as other indexes since it is highly variable depending on the weather. It can vary from 5 to 20 days in general (Washington State University, 2021). Therefore, DAFB helps to estimate only the season for fruit maturity of an orchard in general but is unable to provide the accurate information about the maturity of a fruit particularly.

Fruit maturity testing methods

The maturity indexes can be measured using either destructive or non-destructive methods or a combination of both. As aforementioned, in comparison with destructive methods, non-destructive methods, are more advantageous since they assist in estimating in situ maturity, measuring multiple maturity indexes simultaneously, and supporting real-time decisions (Li et al., 2018a, 2018b). Interestingly, during the process of ripening, several physical and chemical changes in the fruits can be noticed, such as fruit shape, size, colour, weight, ethylene production, firmness, texture, chlorophyll concentration, soluble solid contents, sugars, oils, acids, and respiration (Perkins-Veazie et al., 2000; Prasad et al., 2018). However, non-destructive methods are not able to measure and assess all the changes. Therefore, researchers particularly work on the maturity indices that can describe some physical and chemical changes during harvest time efficiently and hence, non-destructive methods can assist in estimating harvest time and ripeness. Reid (2002) outlines the maturity indices for specific fruits and commonly practiced methods (i.e. destructive and non-destructive) to determine the fruit maturity as delineated in Table 2. Figure 2 outlines the classification of available testing methods that are widely adopted to determine and estimate the fruit maturity and ripening based on some previous works.

This review work exploits the widely applied non-destructive and destructive approaches to identify fruits qualities and narrow down the focus on the possible non-destructive approaches which can be adopted in Smart Agriculture. Hence, the working principles of destructive and non-destructive approaches can be further studied from some previous literature. For example, Li et al. (2018a, 2018b) elaborate the colour detection, spectroscopy and spectral imaging essentially while El-Mesery et al. (2019) shows

Table 2 Maturity indices for specific fruits and methods of maturity determination (Reid, 2002; Vanoli & Buccheri, 2012)

Index	Examples	Method of determination	Category
Elapsed days from full bloom to harvest	Apples, pears	Record of day count	Non-destructive
Mean heat units during development	Apples	Computation from weather data	
Development of abscission layer	Apples, some melons, feijoas	Visual or force of separation	
Surface morphology and structure	Cuticle formation on grapes; Netting of some melons; gloss of some fruits (development of wax)	Visual	
Size	All fruits	Various measuring device, weight	
Specific gravity	Cherries, watermelons	Density gradient solution; flotation techniques, volume/weight	
Shape	Angularity of banana finger; Full cheeks of mangoes	Dimensions, ratio charts	
External colour	All fruits	Light reflectance, visual colour charts	
Internal colour and structure	Flesh colour of some fruits	Light transmittance, delayed light emission, Visual examination	Destructive
Total solids	Avocados, kiwifruit	Dry weight	
Firmness	Apples, pears, stone fruits	Firmness tester, deformation	
Starch content	Apples, pears	KI test, other chemical tests	
Sugar content	Apples, pears, stone fruits, grapes	Hand refractometer, chemical tests	
Acid content, sugar/acid ratio	Pomegranates, citrus, papaya, kiwi	Titration, chemical tests	
Juice content	Citrus fruits	Extraction	
Oil content	Avocados	Extraction, chemical tests	
Astringency (tannin content)	Persimmons, dates	Ferric chloride test	
Internal ethylene concentration	Apples, pears	Gas chromatography	

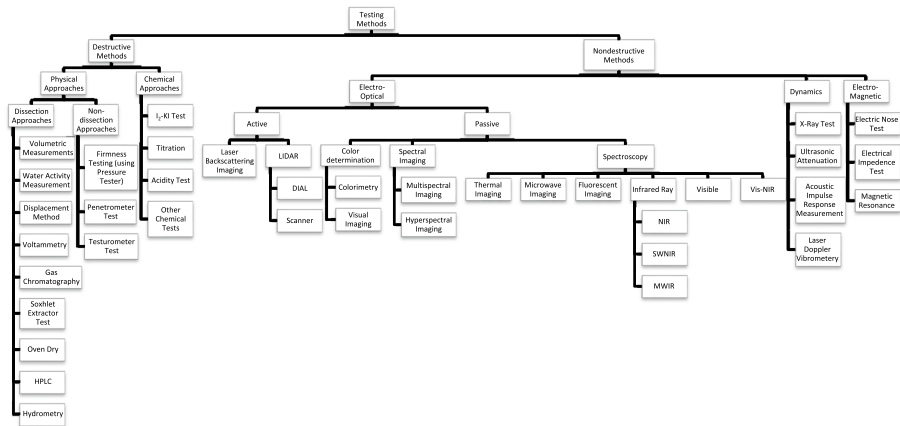


Fig. 2 Classification of fruit ripening and quality measurement testing methods

emphasis on spectroscopy, spectral imaging and some dynamics methods such as acoustic and ultrasonic methods. Lakshmi et al. (2017) briefly discuss the commonly utilised non-destructive methods and classify them according to their working principles. Nielsen (2017) addresses both the destructive and non-destructive approaches which are often practised in food industry.

Destructive methods

Destructive methods are defined as the approaches that measure the fruit maturity or ripening through destroying. These approaches include penetration, peeling fruit flesh/mesocarp, chemical applying reaction or exertion of external force on a fruit. Figure 3 illustrates some of the commonly adopted instruments for destructive testing of fruit maturity. Based on the actual measurement performed, this study classifies destructive methods as either physical or chemical. Physical approaches do not include chemical reactions during the process. Hence, they can be named as non-chemical approaches as well. Physical approaches can be further categorised as follows:

Dissection approaches These rely on some sort of dissection of fruits to extract a sample, which can then be used to measure a maturity index. For instance, volumetric measurement, water activity measurement, displacement method, voltammetry, gas chromatography, Soxhlet extractor test, oven-dry, HPLC, and hydrometry require a pulp primarily from the particular fruit to measure its maturity indexes.

Non-dissection approaches As opposed to dissection approaches, these do not require any sort of dissection of a fruit. Hence, pressure tester, penetrometer, testurometer and effigy tester follow non-dissection approaches since they take the measurement through penetration or creating pressure at the surface.

Chemical Approaches, on the other hand, usually follow a dissection process but focus on chemical characterisation of the sample as opposed to the physical quantities discussed above towards the fruit to measure its maturity indexes. For instance, starch index has traditionally been measured using I2-KI solution with a pulp of a fruit, and TA of a fruit is measured by titration.

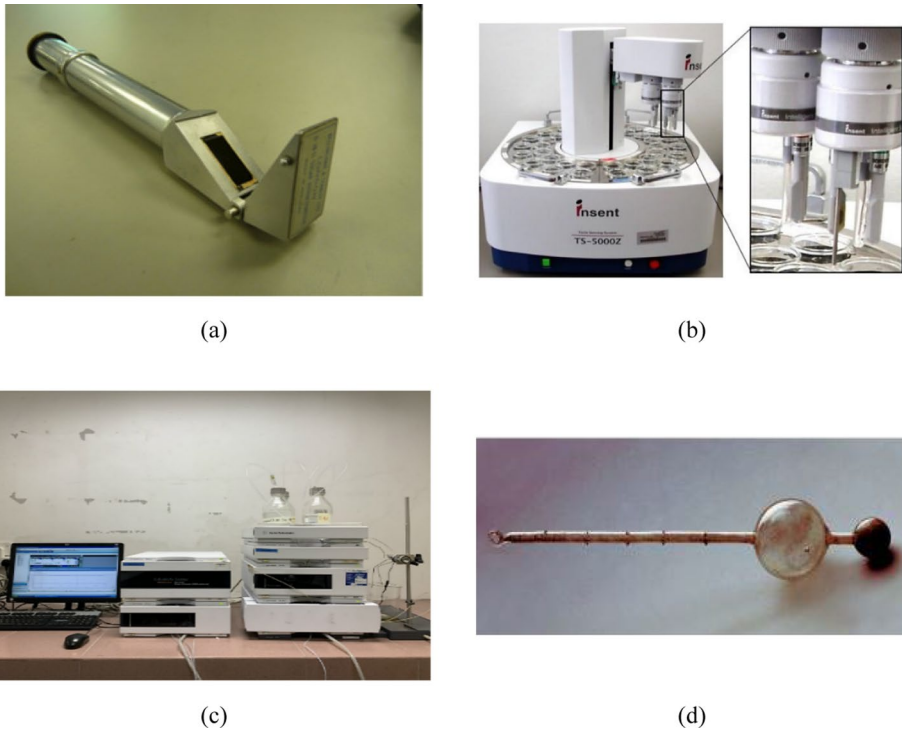


Fig. 3 Destructive SSC measuring instruments: **a** refractometer (Rawle, 2017); **b** electronic tongue (Podraz'ka et al., 2018); **c** HPLC (Monash University Malaysia, 2021); **d** hydrometer (Camuffo, 2019)

Non-destructive methods

Non-destructive methods can be broadly categorized according to the three following working principles: electro-optical, dynamics and electro-magnetic methods. Figures 4, 5, 6 show some of the commonly adopted non-destructive instruments, whereas Table 3 overviews established non-destructive methods.

Electro-optical methods have been defined as the methods that primarily use visual sensors with electronic signal acquisition and processing system. Therefore, it includes various active and passive systems such as: laser, LIDAR, colour detecting (i.e. visible imaging and colorimetry), spectroscopy (i.e. fluorescence, visible, infrared, microwave and thermal imaging), and spectral imaging (i.e. multispectral and hyperspectral imaging). In general, imaging spectroscopies are categorized under passive sensing system as they rely on controlled external energy sources for measuring the energy absorbance, reflectance, refraction or scattering. In contrary, laser and LIDAR are considered as active sensors since both of them generate signals to a fruit and receive back the signals from the same fruit. Thus, they measure the index of a fruit based on the transmitter and receiver signals.

LIDAR-based systems are active systems in that they measure the difference between the emitted pulse and the returned signal at the receiver (time delay, energy or power differential). As opposed to passive systems (hyperspectral) that measure the reflected radiation, active systems are less susceptible to atmospheric conditions, light variability, changes in



(a)

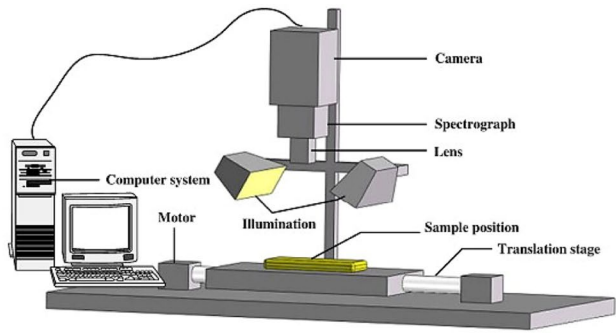


(b)

Fig. 4 Non-destructive instruments: **a** Vis–NIR Spectrometer (Trimble, 2019); **b** NMR Spectrometer (Bruker, 2021)



(a)



(b)

Fig. 5 **a** Hyperspectral camera (Middleton Spectral Vision, 2021) and **b** configuration for measuring fruit maturity (Cheng et al., 2014)

Fig. 6 Firmness tester (Facchini srl, 2021)

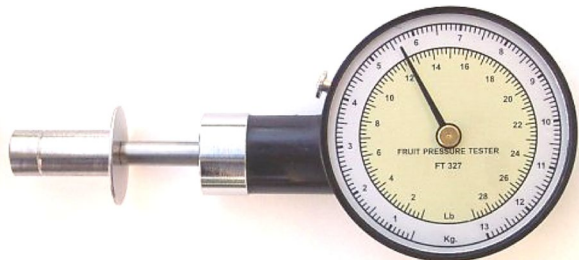


Table 3 Non-destructive methods and their measured maturity indexes

Non-destructive methods		Destructive methods		Fruits	References
Methods	Observables	Instruments	Observables		
Visible imaging	Colour (RGB)	Colorimeter; Texturometer; Potentiometer; Oxalic acid solution	Peel colour; Pulp firmness; pH; Ascorbic acid solution	Papaya	Pereira et al. (2018)
Hyperspectral imaging	Texture colour, SSC and firmness	Penetrometer; Refractometer	Firmness; SSC	Banana	Pu et al. (2019)
Multispectral imaging	Firmness, total soluble solids (TSS) and pH	Penetrometer; Refractometer	Firmness; TSS and pH	Pomegranate	Khodabakhshian et al. (2017)
Colorimetry	Measuring pH and flavour	pH electrode; Texture analyser; Saccharometer	pH; Hardness; SSC	Apple	Kim et al. (2018)
Fluorescent imaging	Chlorophyll fluorescence			Apple	Huybrechts et al. (2002)
Laser backscattering imaging	Moisture content (MC), shrinkage, lightness, chroma and hue	Water activity meter; displacement method	MC; Shrinkage	Papaya	Udomkun et al. (2014)
Thermal imaging	Heat capacity/temperature	–	–	Mango	Sumriddechajorn and Intaravanne (2013)
Microwave imaging	Seeds	–	–	Lemon and grape	Ghavami et al. (2019)
Magnetic resonance imaging	Sugar content	–	–	Watermelon	Miki et al. (1996), Srivastava et al. (2018)
Infrared thermography	Heat capacity/Heat diffusion	Pressure tester	Firmness	Apple and tomato	Offermann et al. (1998)
Near infrared (VIS–NIR) spectrometer	SSC and Dry Matter (DM)	Refractometer; weight loss	SSC; DM	Mango	dos Santos Neto et al. (2017)
Ultrasonic technique	Water contents and firmness	Displacement method	Water content, volume, dehydration	Orange	Morrison and Abeyratne (2014)
Acoustic impulse response	Coconut meat and water content	–	–	Coconut	Caladcad et al. (2020)
Electric nose	Aroma	–	–	Banana	Chen et al. (2018)
Electrical impedance spectroscopy	Days after floration	–	–	Mango	Neto et al. (2017)

Table 3 (continued)

Non-destructive methods		Destructive methods		Fruits	References
Methods	Observables	Instruments	Observables		
X-ray tomography	Arils, kernel, peel, juice content, air space	Displacement method	Volumetric measurements	Pomegranate	Arendse et al. (2016)

viewing angle, and the canopy structure (for airborne measurements). Relating the tree profile and volume LIDAR to fruit yield and maturity, has been successfully applied to peaches (Pascual et al., 2009), but is heavily related to other measurements such as irrigation frequency. These types of applications, underutilize the data gathered by the LIDAR sensor. In fruit maturity LIDAR spectroscopy has also been used to measure oil palm fruit maturity specifically measuring the relative reflectance of fruit bunches on trees (Zulkifli et al., 2018). LIDAR-based systems in agriculture have largely been confined to 3D scanning, point cloud techniques, despite the inherent versatility of laser systems. Molecular-line absorption is an interesting area of laser application for atmospheric sounding, determining atmospheric constituents, but relatively obscure in agriculture applications. Previous work (Gardi et al., 2016, 2017; Pham et al., 2019) has identified LIDAR-based systems for volatile organic compounds (VOC) and CO₂ estimation. Knowledge of these atmospheric constituents has clear causal implications for the photosynthetic and fruit maturity processes (Fahey et al., 2021). NASA's Active Sensing of CO₂ Emissions over Nights, Days and Seasons (ASCENDS) mission, employs monostatic Differential Absorption LIDAR (DIAL) to measure oceanic CO₂ flux, taking advantage of the ability to take day and night measurements across different conditions that allows the small laser beam to pass through gaps in clouds (Abshire et al., 2010). The push towards LIDAR-based or fusion systems in agriculture, highlights the practical difficulties of VIS-based systems in in-field applications.

Dynamic methods include X-ray computed tomography, ultrasonic attenuation, acoustic impulse response measurement and laser doppler vibrometry. In these methods, primarily, the atoms or molecules of a fruit are stimulated by a detector and based on the response from the detector, the maturity indexes of the fruit can be measured.

Electro-Magnetic methods establish relations between a detector and a fruit through the principles of electricity and magnetism and find out the maturity status of the fruit. Hence, it incorporates several non-destructive approaches such as Electric Nose Test, Electrical Impedance Test and Nuclear Magnetic Resonance.

It is worthy to mention that single non-destructive sensor is not enough efficient to offer a fruit maturity index with a higher accuracy. Thus, destructive methods are adopted in order to achieve the certainty in maturity indexes of fruits. Table 3 describes the commonly used non-destructive methods on different fruits in order to measure their maturity indexes that are supported by some destructive methods.

Sensor fusion and artificial intelligence for sensor data processing

With the advancement of sensor technology, multi-sensor fusion and AI classification/inference methods have achieved a fast-growing research area since it has the ability to overcome the limitations of individual sensors, reduce the uncertainty and error in data and cooperatively provide information that may not be achieved by a single sensor (Durrant-Whyte, 1990; Mitchell, 2007). This section overviews the applications of sensor fusion and AI to the determination of fruit maturity.

Sensor fusion in fruit maturity estimation

Dull (1986) highlighted the need to consider multiple information sources to determine the fruit quality and Benady (1994) applied the concept of sensor fusion on an intelligent

robot to determine the ripeness of melon through visual inspection and gas sensing. Steinmetz et al. (1999a, 1999b) demonstrated that multi-sensor fusion can significantly reduce the error in classifying fruit on the basis of maturity compared to using just a single sensor. With increasing levels of accuracy realised through sensor fusion, researchers are considering an ever-expanding range of different sensors to assess fruit maturity. For example, Steinmetz et al. (1999a, 1999b) considered vision coupled with spectrophotometry to predict a fruit maturity index, namely, the sugar content of apples. Mendoza et al. (2012) used a combination of an acoustic sensor, a bio-yield (firmness) tester and a visible and shortwave near-infrared (vis-SWNIR) spectrometer to measure apple firmness and SSC respectively. Zakaria et al. (2012) used the so-called ‘electronic nose’ along with an acoustic sensor to discriminate and classify the maturity and ripeness level of mangoes through measuring the aroma and firmness, achieving 84.4% classification accuracy.

Artificial intelligence in fruit maturity estimation

AI, a human intelligence-mimicking technology, is progressively covering almost all sorts of engineering areas nowadays, particularly that are data and logic-oriented though its initialization takes place under computer science. Its rigorous and multi-faceted applicability makes it pervasive in versatile research areas that would otherwise be computationally infeasible. Its domain of applications is so vast that cannot be highlighted only through naming, and hence, based on the characteristics of different applications, this study classifies it into ML, Genetic Algorithms, Fuzzy Logics and Expert Systems as shown in Fig. 7. This study mainly explains the applications of ML in fruit classification based on their maturity level, and thus, only the branches of AI in regards with only fruit maturity have been considered. Since agricultural growth is necessarily proportional to the growing population, smart agriculture has become a concern to researchers. Apart from that, traditional methods that have been practised are not sufficient to keep pace with increasing demands and hence, it expedites the development of smart agriculture. The necessity of smart agriculture then introduces AI that contributes to several agricultural areas such as crop diseases, pesticide control, irrigation and water management, storage management and crop,

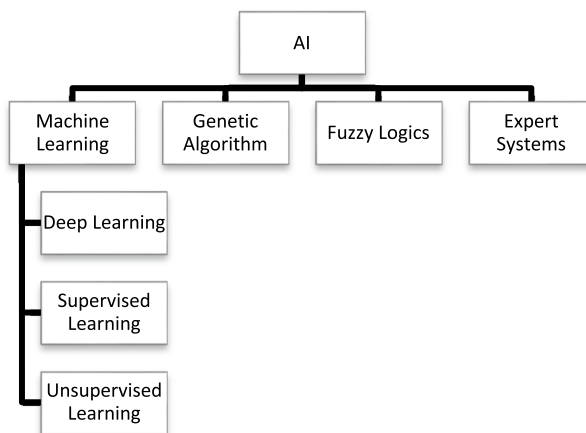


Fig. 7 Branches of Artificial Intelligence

weed and soil management (Jha et al., 2019). As a result, several areas of applications of AI have been found in agriculture such as expert systems (Murase, 2000; Rani et al., 2011), decision support or making (Attonaty et al., 1999; Thomopoulos et al., 2015), prediction (Chlingaryan et al., 2018), optimization (Talaviya et al., 2020), multi-agent systems (Bahri et al., 2020; Skobelev et al., 2018), machine vision (Patrício & Rieder, 2018), natural language processing (Mostaco et al., 2018) and ML (Liakos et al., 2018). This study particularly focuses on the employment of AI for the in-situ fruit maturity and ripeness and hence, ML is the area of interest of this study since it deals with pattern recognition.

Machine learning

The interest in ML is resurging as high computation with high volumes of data is no more a challenging task and its progress in last two decades is significantly prominent. The purpose of developing ML is to train a model that can update and improve itself through experience automatically. These ML algorithms train models by extracting characteristic patterns from data that are applicable to a given industrial scenario (Diez-Olivan et al., 2019). Therefore, ML mainly deals with computer science and statistics concurrently, and its core comes from data science and AI (Jordan & Mitchell, 2015). In the realm of AI, ML emerges to define models based on some algorithms that finally are able to identify, classify or predict data efficiently and it can be performed through suitable training processes.

Definitions

Machine Learning (ML) methods are a data-driven subset of AI wherein typically large datasets are used to train algorithms to ‘learn’ highly non-linear functions automatically or semi-automatically (Salcedo-Sanz et al., 2020) that are difficult to model analytically. Despite the large number of ML algorithms, each with numerous variants that have been developed, the following fundamental definitions are commonly employed in the ML community to most algorithms:

Model The output of an ML training process is a model which represents the relationship or mapping between the user-provided input/observables e.g. fruit colour, pH value, measured volume and the output of interest e.g. maturity level, quality index. Under the ML paradigm, the model is not derived analytically from first principles i.e., a physics-based model, but rather, is derived from data through an automated training algorithm.

Model parameters This refers to iteratively adjusted parameters by the training algorithm according to a user-defined cost criterion. When the model training is complete, the parameters are ‘frozen’, and the model can be deployed in the field for its chosen application. For example, within the Artificial Neural Network (ANN) class of ML algorithms, the model is parameterized through weights and biases of fundamental computing elements called neurons.

Training The procedure by which model parameters are adjusted according to a cost criterion. As an example, backpropagation is commonly used in ANN to adjust weights and biases according to a cost function formulated from the difference between the model output and the target output. Training time varies widely from hours to several weeks depending on the number of model parameters, the size of the available dataset and the ‘learning rate’ or step-size to adjust the parameters for each iteration.

Dataset This refers to the user-provided input data that is used to adjust and finalize the model parameters prior to its deployment in the field. In the ML context, it is essential to

make a distinction between the training dataset and the testing dataset. As the name suggests, the training dataset is used to adjust the model parameters during the training phase. Subsequently, an independent dataset is used to test the model to inform further parameter adjustments. The split of datasets into training and testing datasets is an important factor in ensuring that the ML model can generalize to unseen data during its operational deployment. Depending on the type of algorithm, the input training data can be completely 'raw' i.e., little to no pre-processing or preparation. This is the case for Deep Learning algorithms where features are automatically extracted from the data during the training process. For standard ML algorithms, Input datasets are typically analyzed beforehand to inform the development of hand-crafted features which are then input to the model.

Workflow

The objective of ML is to develop a model that can make an inference of the system or process being observed. The model essentially removes the need of deriving an analytical relationship that may be difficult or impossible depending on the non-linearity of the process. To estimate fruit maturity or identify a mature fruit, a ML algorithm receives the maturity indexes as the pre-processed data. Then the data are segregated into testing data and training data. In general, a large portion of the initial dataset is required to support the ML training, and the rest of the data are kept as testing data where the allotted data for testing and training depends on the features available in the data. It is important to note that the training data deserves consideration and the need for quality control by documenting performance with further independent sets is critical in practical adoption.

Then these training data are used to develop a model considering either a supervised or unsupervised learning approach. In unsupervised learning, the features are not labelled, and consequently, the algorithm identifies data clusters autonomously, without user guidance. Conversely, supervised learning algorithms are provided with labelled data by the user. Continuous data is trained through regression approaches, while discrete, categorized data is trained using classification methods, particularly designed for organizing data into distinct groups. Finally, this model is tested using the testing data, and the accuracy is measured. Depending on the application, Root Mean Square Error (RMSE) or confusion matrices are typically used as measures of model performance. Hence, if the model fails to interpret data with a pre-defined accuracy, the model is required to be re-developed. For more detail explanations about the workflow of ML algorithms, some previous works can be referred as (Bonaccorso, 2018; Janiesch et al., 2021; Mohammed et al., 2016). Figure 8 depicts the complete process of ML algorithms through a block diagram.

Classifications

ML algorithms can be classified into supervised and unsupervised learning based on their learning approaches, primarily though some works add semi-supervised learning and reinforcement learning (Kumar et al., 2019). As this work particularly focuses on the ML approaches in fruit maturity estimation, semi-supervised and reinforcement learnings are not widely practised. Hence, this sub-section focuses on supervised and unsupervised learning and their application to fruit maturity classification and estimation. Since this review work only explores the applications of different ML algorithms in fruit maturity estimation and their maturity-based classification, the working principles of some of the commonly considered ML algorithms have been discussed briefly. For further explanation,

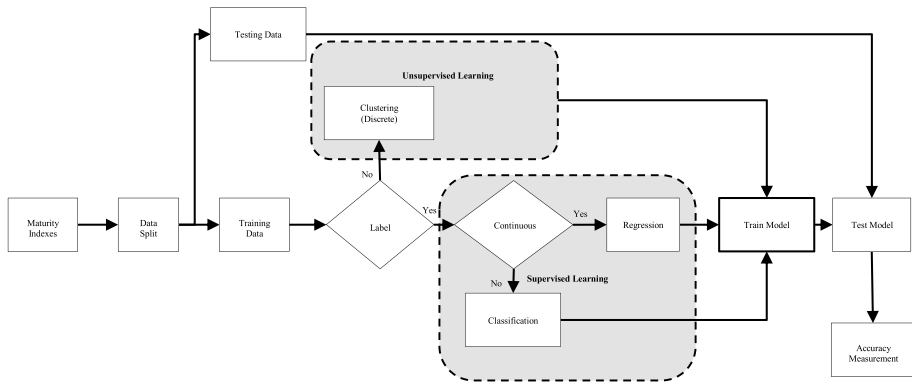


Fig. 8 Block diagram of ML algorithms' workflow

several works can be referred such as (Bonaccorso, 2018; Dhall et al., 2020; Mohammed et al., 2016; Sarker, 2021). Figure 9 shows the classification of ML algorithms that are commonly applied in Smart Agriculture.

Supervised Learning approaches can be commonly listed based on the algorithm as Naïve Bayes, SVM, Decision Tree, Random Forests and ANN (Fig. 10).

Supervised learning

Supervised learning necessitates labelling the data before the training process. During the training process, the training examples are input into the model. The model output predictions are compared against the label (true) value. The difference between the model prediction and the label is used to form a cost function. The model parameters are then iteratively optimized to minimize the cost function using optimization algorithms, i.e. Gradient Descent, AdaDelta, Adaptive Moment Estimation (Adam), RMSProp, Adagrad, Momentum, Nesterov Accelerated Gradient etc. Gradient Descent (GD) and its variants Batch GD and Stochastic Gradient Descent (SGD) are the most commonly employed optimization routines used in training. During the training process, patterns that characterize the underlying characteristics of the input–output relationship are extracted to inform the model prediction. In the fruit maturity context, ML models most commonly perform one of two tasks viz., classification or regression. In classification, observables or features extracted from sensor measurements or tests are input to the model, which infers and outputs the state of a particular crop as belonging to one of the multiple possible discrete categories, e.g., ripe, moderate ripeness, unripe. In regression, temporal features are extracted from the data to predict the future state of the observed crop. Both these applications are covered in greater detail in the following sections.

Classification based on algorithm

Supervised Learning approaches can be commonly listed based on the algorithm as Naïve Bayes, SVM, Decision Tree, Random Forests and Artificial Neural Network. Hence, this section explains the aforementioned supervised learning approaches since they are widely used in fruit maturity applications to either categorize fruits or estimate their maturity.

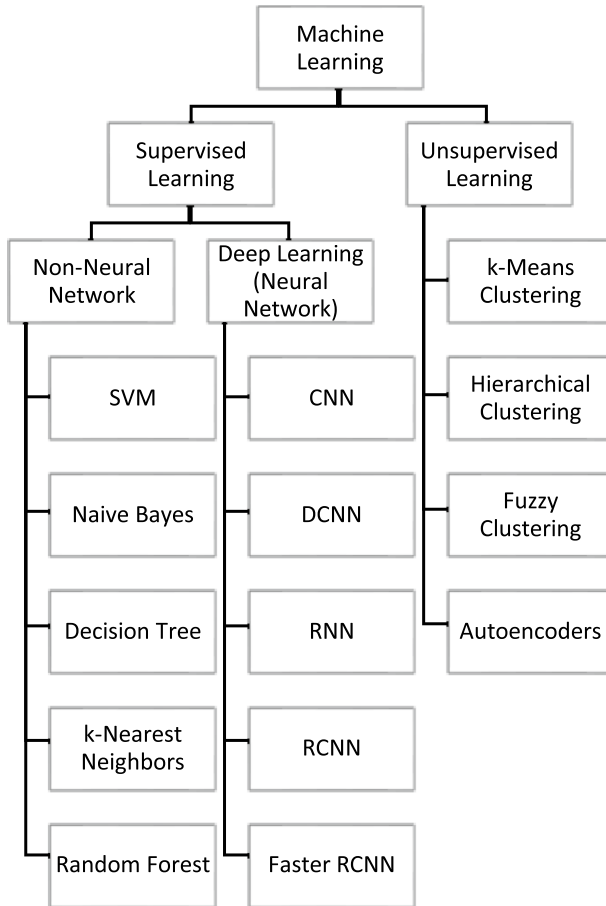


Fig. 9 Taxonomy of ML in Smart Agriculture (especially in Fruit Maturity application)

SVM Support Vector Machines (SVM) are powerful supervised machine-learning tools used mainly for classification, though less frequently for regression. They create a ‘decision boundary’ to classify new data into two categories. The standard SVM aims to find the widest ‘street’ between these categories, but initially had ‘hard margins,’ being highly sensitive to outliers. Soft-margin SVMs were later developed to allow for some margin violations. Since real-world datasets are often not linearly separable, SVMs use ‘kernel’ functions to transform data into a high-dimensional feature space. This facilitates the application of linear SVM in the transformed space to find the separation boundary. Common kernels include polynomial and Radial Basis Functions (RBF). In the context of this specific application, it is frequently observed that datasets exhibit nonlinearity when multiple features, such as fruit maturity indexes, are plotted against each other. Consequently, it becomes necessary to consider the utilization of a hyperplane that can effectively address the inherent high-dimensionality, as exemplified in references (De-la-Torre et al., 2019; Zhang & Wu, 2012).

Decision Trees Decision trees are a non-parametric supervised learning approach that solves classification and regression problems (Reyes, 2020). They use input attributes

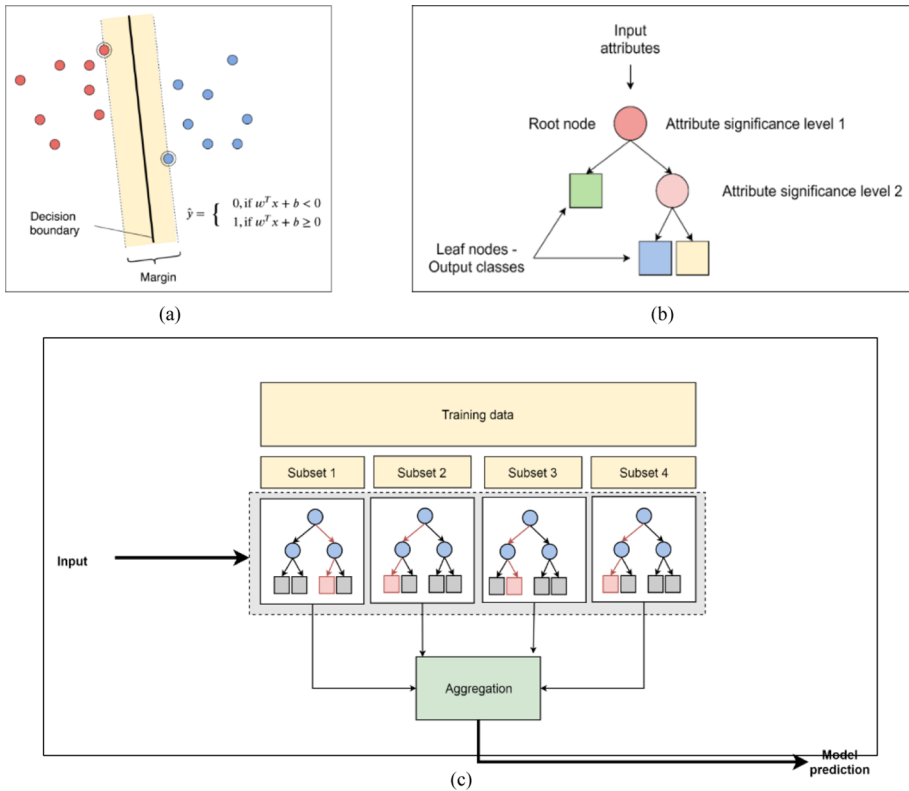


Fig. 10 a SVM, b Decision Tree and c Random Forest (Bijjahalli et al., 2020)

to make decisions, which can be output classes or numerical values. The tree structure involves arranging attributes based on data values, with the root at the top and nodes branching from it. Decisions are generated at the endpoints where no further splitting occurs (Alloghani et al., 2020). Hunt's algorithm drives the selection of sequences for each node, determining the order of decisions. The Gini impurity is introduced to achieve accurate classification based on these attributes, minimizing the Gini index to 0 (Plapinger, 2017). Decision Tree method holds significant appeal in the context of classifying fruits based on their maturity levels, primarily owing to its simplicity and interpretability (Mim et al., 2018). The literature supports its higher predictive accuracy for fruit classification, as evidenced in (Fadchar & Cruz, 2020; Mim et al., 2018; Suresha et al., 2012).

Random Forest Random Forest is a machine learning approach that leverages an ensemble of decision trees to predict outcomes. It trains multiple decision trees in parallel and determines the final decision based on the majority consensus of these trees. Random Forest employs bootstrapping to train each decision tree on different subsets of the training data (Misra & Li, 2019), ensuring diversity and reducing variance. This approach, collectively known as bagging, makes the algorithm more robust than a single decision tree when dealing with training samples and noisy datasets. Random Forest is particularly advantageous for fruit classification and maturity estimation tasks where multiple maturity indexes and large sample sizes pose challenges for algorithm

training. In such cases, Random Forest demonstrates promising performance, offering higher accuracy in both classification and regression applications. It is capable of outperforming Support Vector Machines (SVM) in certain scenarios, despite SVM's efficiency with large datasets (Kaur & Gupta, 2017). Random Forest's strength lies in its ensemble of decision trees, transforming weak learners into strong ones, ultimately resulting in higher prediction accuracy (Elhariri et al., 2014).

Naïve Bayes Naïve Bayes is a probabilistic classifier based on Bayes' theorem (Mushtaq & Mellouk, 2017). It assumes that each input feature independently and equally contributes to the target class. For each class value, it calculates conditional probabilities for each feature, then applies the product rule to derive a joint conditional probability for the features (Thornton, 2021). Bayes' rule is used to find the conditional probability for the class variable. It calculates conditional probabilities of each feature for each class, either assuming a multinomial distribution for discrete features or a Gaussian distribution for continuous features. The classifier then approximates the probability of a sample with known feature values belonging to a specific class. It selects the class with the highest combined conditional probabilities for all features and that class is the final prediction. It's important to remove correlated features, as they can be overemphasized (Misra & Li, 2019). Naïve Bayes is favoured in fruit classification due to its low training data requirements and faster computation compared to more complex algorithms like SVM, ANN, or Random Forest (Ibba et al., 2021). The Gaussian kernel variant of Naïve Bayes is particularly effective in classifying fruits with nonlinear sample data. However, Naïve Bayes is unsuitable for regression tasks and assumes feature independence, which can be advantageous when dealing with truly independent maturity indexes (Kusuma & Putra, 2018).

KNN K-Nearest Neighbour (K-NN) is a versatile non-parametric machine learning approach suitable for handling large, multidimensional datasets in both classification and regression tasks. K-NN stores training dataset examples and classifies new instances based on their similarity to stored examples. The choice of k , the number of neighbours to consider during classification, impacts the method's performance (Lindon et al., 2011). A suitable “ k ” is determined through distance measures like Euclidean, Hamming, Manhattan, Minkowski, or Chebychev (Isa et al., 2017). K-NN excels in small, noisy datasets (Liu & Zhang, 2012), making it apt for tasks like fruit classification. Weighted K-NN, which assigns higher weight to closer neighbors, enhances accuracy (Fan et al., 2019). However, K-NN demands more memory, longer training times, and slower runtime as dataset size increases (Amra & Maghari, 2017; Harrison, 2018).

ANN Artificial Neural Network (ANN) is a biologically inspired computer algorithm designed to mimic human brain information processing. It operates through interconnected neuron layers, including an input layer for data collection, a hidden layer for processing, and an output layer for delivering results (Chang et al., 2008). Neurons are assigned weights and biases, employing nonlinear activation functions like sigmoid, ReLU, or hyperbolic tangent (Zhang et al., 2018). The inner workings of hidden layer neurons remain largely opaque, making ANN a black-box model (Jo, 2021). It updates weights through techniques like back-propagation, addressing model uncertainties. However, ANN's performance relies on extensive training data, and its complexity hinders interpretability as layers increase. While ANN is effective with large, nonlinear datasets, the lack of interpretability poses challenges as layer count rises (Mijwel, 2018).

Classification based on data type

Based on the types of data and their application, supervised learning takes the following two forms in agriculture: (1) classification; and (2) regression approaches.

Classification approaches In horticulture, ML is also introduced to interpret multiple sensors data and fuse them through proper training in order to classify fruits according to their maturity and predict the harvest time. As a result, several supervised ML algorithms have been introduced to classify the maturity levels of different fruits, such as Support Vector Machine (SVM) (Caladcad et al., 2020), Naïve Bayes (Kusuma & Putra, 2018), Decision-tree (Wajid et al., 2018) and Random forest (Harel et al., 2020). Table 4 depicts the consideration of different ML approaches in classifying the fruits based on their maturity in recent years. Caladcad et al. (2020) use SVM as a classifier to classify coconuts into three (3) groups and interestingly, they use the acoustic sensor data that are the measurements of coconut meat and water volume. Finally, they achieve in classifying fruits according to their maturity level with an accuracy of 80%. Kusuma and Putra (2018) classify tomatoes into raw, ripe and rot categories considering Naïve Bayes approach. They develop a grayscale formula using RGB colour images and finally are able to reach 76% accuracy in classifying tomatoes (Kusuma & Putra, 2018). Similarly, Wajid et al. (2018), Jhawar (2016) and Harel et al. (2020) adopt decision tree and linear regression to classify orange, and Harel et al. use a random forest to classify sweet pepper and finally, achieve 93.13%, 97.98% and 95.82% accuracy in classifying them respectively. These three studies use RGB colour images to identify fruits' maturity levels. On the other hand, Saeed et al. (2012) use vegetation indexes to classify the palm fruits and manage to reach an accuracy of 82.2% where K-nearest neighbours have been used as the classifiers. It shall be noted that the accuracy in each study is measured based on the testing data performance in comparison with training data.

Regression approaches In situ fruit classification is not sufficient alone if the maturity cannot be predicted as it is still a laborious and time-consuming job. Therefore, researchers introduce another dimension of ML that can predict the fruit maturity indexes and determine the optimal harvest time. Hence, a regression model is developed at the beginning to test a group of data that has been collected and at this stage, it is known as a pre-process model. Then the best interpreting model can be considered as a predicting model, and then the model is yielded to be evaluated based on accuracy. Several regression models are taken into consideration such as Multiple Linear Regression (MLR) (Jha et al., 2007), Partial Least Square (PLS) (Li et al., 2018a, 2018b), and Principal Component Regression (PCR) (Mahesh et al., 2015). MLR is a regression method that determines potential explanatory variables which are important to predict the response variable (Hemmings & Hopkins, 2006). PLS, an extension of MLR, predicts the response of a variable from the nature of a large set of predictor variables through reducing the predictor variables to a smaller set of predictor variables and finally performs least square regression (Xia, 2020). Besides, PCR reduces a large set of explanatory predictor variables into a set of smaller principal components and at the same time, offer a regression model (Artigue & Smith, 2019). In the meantime, for scattering correction, Standard Normal Variate (SNV) (Marques et al., 2016; Torres et al., 2015) and Multiple Scattering Correction (MSC) (He et al., 2005; Saranwong et al., 2004), for describing scattering profile, Modified Lorentzian Distribution (MLD) (Peng & Lu, 2008) and for data smoothing, Savitzky–Golay (SG) filter are some commonly used approaches that lead to achieving suitable linear regression (Gorry, 1990). Table 5

Table 4 Algorithms used in classifying fruit maturity

Algorithms (classifiers)	Fruits	Data type	Features	No. of classes	Accuracy (%)	References
SVM	Coconut	Wave amplitude	Coconut meat and coconut water volume	3	80	Caladcat et al. (2020)
Naïve Bayes	Tomato	Histogram features ^a	Texture features	3	76	Kusuma and Putra (2018)
Decision Tree	Orange	RGB color and gray values	Texture features	3	93.13	Wajid et al. (2018)
Random forest	Sweet pepper	RGB color space, HSV color space and morphological features ^b	Texture features	4	95.82	Harel et al. (2020)
K-nearest neighbours	Oil Palm	Light reflectance of four different wavelengths	Vegetation indices ^c	3	82.2	Saeed et al. (2012)
Linear regression	Orange	RGB color value	Texture features	4	97.98	Jhawar (2016)
K-means clustering	Tomato	RGB, HIS and L* ^a *b color spaces	Texture features	6	80–90	Pacheco and López (2019)
Faster R-CNN	Coconut	RGB	Size, shape, illuminations and brightness	2	89.4	Parvathi and Selvi (2021)
ANN	Mulberry	R, B (from RGB) and extracted color image	Colour, geometric and texture features	3	99.13 ^d 98.26 ^e	Azarmdel et al. (2020)
CNN	Banana, Mango, Calamansi	RGB and grayscale image	Texture features	3	97 (Banana) 94 (Calamansi) 85 (Mango)	Aylon et al. (2019)
DCNN	Strawberry	RGB Image (of ROI)	Texture features	2	88.03 (mature) 77.21 (immature)	Habaragamuwa et al. (2018)

^ai.e. mean intensity, energy, entropy, contrast, smoothness and skewness

^bi.e. no. of pixels, eccentricity, extent and solidity

^cNormalized/renormalized difference vegetation index, simple ratio index, greenness index, red edge shift and plant senescence reflectance index

^dCorrelation-based feature selection subset

^eConsistency feature selection subset

describes the commonly considered regression models to offer a comprehensive idea about the earlier works in fruit maturity prediction. The table shows commonly used regression approaches such as ANN, PLS, MLR, and LDA-Competitive Learning Neural Network (CLNN). This table also describes the considered sensors for data acquisition and observables from the acquired data in those studies. Pre-process algorithms have also been highlighted that leads to predict the maturity indexes further. It is important to underline that these maturity indexes assist to predict when the fruits should be harvested since these maturity indexes gives the measurement of their gradual changes through ripening process.

Classification based on AI approaches

Supervised learning algorithms can be alternatively/further classified into two based on the neural network approaches: neural networks and non-neural networks.

Neural Networks Several approaches appear in horticulture to classify and predict fruit maturity. For instance, ANN (Amiryousefi et al., 2012), CNN (Ayllon et al., 2019), faster Region-based Convolutional Neural Network (Faster R-CNN) (Zhu et al., 2020), Deep Convolutional Neural Network (D-CNN) (Habaragamuwa et al., 2018) and Recurrent Neural Network (RNN) (Yossy et al., 2017), have been adopted to classify the maturity stages of banana, mango and calamansi, strawberry and blueberry respectively. Deep Learning (DL) approaches, including CNN, RNN and their derivatives, are more computationally demanding than conventional ANN (largely due to the higher number of hidden layers). However, DL is more capable than ANN in dealing with more complex non-linear problems (Raghavan et al., 2016). One of the most successful areas of applications of deep learning is pattern recognition (Bai et al., 2021). In spite of its notable advantages, DL has several drawbacks; for example it requires more model parameters during training and as a result, it is prone to overfitting due to over-parameterization (Thompson et al., 2020). Also, deep ANN have much higher computational costs (Pandian, 2021; Thompson et al., 2020) and, as a result, one or more GPU become necessary. Each additional GPU increase significantly the amount of power absorbed by the PC. Hence, DL require higher power consumption comparing to other and especially non-neural network methods (Thompson et al., 2020). Table 5 discerns the applications of these algorithms in several research works with their accuracy to provide a better understanding of the algorithms.

Non-Neural Networks These alternative approaches do not adopt the conventional ANN multi-layer structure to interpret data, and thus, they include SVM, Decision Tree, Random Forest, Naïve Bayes, k-means neighbour, linear regression and logistic regression.

Notably, none of the approaches is perfect unequivocally in classifying fruit maturity level and predicting the fruit harvesting time, and at the same time, they have their own merits and demerits. Therefore, a proper trade-off between the advantages and drawbacks of each method can help to reach higher accuracy in classifying fruits according to their maturity level and predicting the suitable time for harvesting.

Unsupervised learning

Unsupervised learning is a type of ML that learns pattern of data without human intervention. Therefore, unsupervised learning is applied where data are not labelled, and the system does not have any prior knowledge about the data (Debener et al., 2023; Károlyi et al., 2018). Clustering is the most popular approach among the unsupervised learning

Table 5 Algorithms to predict fruit maturity indexes

Algorithms	Sensor	Sensor observations	Pre-process algorithms	Prediction indexes	Fruits	References
ANN	Spectrometry (UV–Vis NIR)	Chlorophyll, TSS, Sugar content, pH	PCR, PLR, MLR	Harvest time	Citrus fruits	Srivastava et al. (2020)
PLS	Spectrometry (NIR spectral)	Moisture content, Dry matter (DM), Oil content	Multiplicative scatter correction (MSC), Savitzky–Golay first and Second derivative, Standard Normal Variate (SNV), Mean Normalization, and Orthogonal Signal Correction Methods	Moisture content and Dry matter	Avocado	Olarewaju et al. (2016)
	Acoustic firmness sensing, Bioyield firmness test, VIS–NIR spectroscopy, and online spectral scattering	Firmness and SSC	PLS with principal component analysis (PCA)	SSC	Apple	Mendoza et al. (2012)
	Spectrophotometers at the VIS–NIR and SWIR spectral range, Relaxation and Ultrasonic tests, and colour measurement	Internal content, colour, texture, spectral measurement, firmness, coefficient of elasticity, TSS, DM, AA, OP, total chlorophylls, carotenoids, texture	PLS, PCR, Kernel and SVM	Harvest time ^a	Bell pepper	Ignat et al. (2015)
	Acoustic impulse resonance frequency sensor and miniaturized VIS/NIR spectrometer	SSC, firmness, chlorophyll, relative water content	PLS	SSC and Firmness	Apple	Zude et al. (2006)
PLS, MLR	Spectrometry (NIR)	TSS, DM, Titratable Acidity (TA), Arithmetic and Geometric Mean Diameter, Aspect ratio, Sphericity	Baseline Corrections, Smoothing, Second Order Derivatives and Multiplicative Scatter Corrections (MSC)	$I_m = \frac{TSS \times DM}{TA}$	Mango	Jha et al. (2014)
MLR	Colorimeter	Skin colour values (L, a, b)	PLS, PCR, MLR	$I_m = \frac{TSS}{8} \times 100^c$	Mango	Jha et al. (2007)
LDA-CLNN	Electric nose and acoustic sensor	Aroma and Volatile gases	PCA and LDA	Ripeness level	Mango	Zakaria et al. (2012)

^aHarvesting time is measured considering a New combined quality index (NCQI)
^bMaturity Index that is correlated with TSS, DM and TA where η is a constant specific for each cultivar and region of its production
^cMaturity Index that is correlated with colour values

approaches and commonly used in a wide range of applications such as feature extraction, pattern recognition, image segmentation, vector quantization and data mining (Du, 2010). It divides the input data into several groups where the system is not aware of the groups beforehand. Surprisingly, it has become popular in fruit maturity because of its well-known applications like feature extraction and pattern recognition. For example, Tu et al. (2018) apply k-means clustering to classify the maturity stages of passion fruits and achieve 91.52% of accuracy. Moradi et al. (2012) consider fuzzy c-means to classify pomegranates based on their maturity level through magnetic resonance (MR) images and able to reach an accuracy of 85.93%. Gong et al. (Gong et al., 2014) use hierarchical clustering to classify apples into 5 groups based on their maturity level. Besides, a few works on autoencoders are available in the area of fruit maturity classification. Varga et al. (2021) adopts autoencoder to enhance the performance of Deep Neural Network in terms of classifying kiwis according to their level of maturity.

Comparison of methods

ML algorithms are widely being used in different applications, and hence, it is important to highlight on what basis they are chosen in different fields. Some of the algorithms are well suited for classification, while some others are highly effective for regression. For example, Naïve Bayes is commonly used for classification while it is not used for regression. SVM is considered as an efficient classifier when dealing with high-dimensional data. Moreover, KNN is popular for its efficiency in dealing with outliers. Maturity estimation of fruit and their classification based on maturity, highly deals with outliers, predictors dependence and sometimes high-dimensionality. Therefore, choosing a suitable algorithm is highly challenging for this application. Hence, Table 6 outlines the algorithms' features, advantages, limitations and suitable application areas to offer a clear understanding about them (Ma et al., 2019; Mao & Wang, 2012; Sen et al., 2020; Shrivastava et al., 2020; Tu, 1996).

Challenges and future research and development trends

ML-based methods promise several benefits to smart agriculture. Broad application areas that benefit from this line of research include fruit maturity prediction, fruit quality classification, crop disease detection, plant health monitoring, weed detection, water level in soil, etc. In each of these domains, ML methods complement conventional algorithms to provide greater adaptability to evolving environmental conditions. Alternatively, they can also be used to replace traditional analytic algorithms entirely. These methods support the ability to generalize from training patterns or data. Despite these promised benefits, there are several limitations and pitfalls to be carefully addressed. Hence, this study categorizes the challenges and limitations from three different perspectives, namely algorithms, maturity indexes and instruments.

Challenges

The following subsections discuss in detail some of the relevant challenges faced when applying advanced sensor fusion and AI techniques to agriculture and particularly fruit maturity estimation.

Table 6 Comparative discussion among the ML algorithms

Algorithm	Actions/features	Advantage	Disadvantage	Application
Decision tree	Takes decision at every stages based on predefined conditions	Simple and easy to interpret For pre-processing, it requires less data	Can form a complex tree structure if the number of input features is high	Regression and classification
Random forest	Primarily developed by many Decision Trees	Applicable with both numerical and categorical data One of the best for classification Can deal with the missing values with an accuracy Does not overfit a model Can deal with a large volume of data with higher dimensionality	Poor in stability i.e. sensitive to noise in the input features Not efficient as regression model	Regression and Classification
Support vector machine	Each decision tree offers a classification and random forest picks the highest voted prediction from the collected classifications n number of features can be grouped in n dimensional space by hyper-plane	Efficient in dealing with high-dimensional data Memory efficient since it uses subset training points Does not require huge volume of data Fast in prediction Efficient in diagnostic problems	Limited access in model activities Loss of interpretability as the size of the ensemble increases Training time can be higher some-times Performance degrades in presence of the overlapping of target classes Inefficient in estimation	Regression and classification
Naïve bayes	Adopts Bayesian theorem			Classification
K-nearest neighbors	Applies probabilities in every case and defines new cases based on highest probability outcome Stores all data and predicts the class of a new instance concentrating on the similarity at the nearest neighbours	Implementation is simple Efficient in noisy dataset Can work effectively with large dataset	Highly computational that leads to slow response Challenging to find a suitable value of k	Regression and classification

Table 6 (continued)

Algorithm	Actions/features	Advantage	Disadvantage	Application
ANN and derivatives	Artificial neurons form layers that receives information from outside	Can deal with a large volume of data Does not need to be defined the features Requires less statistical training to develop a model	Black box nature	Regression and classification
	These layers are connected through mathematical relations	Can detect nonlinear relationships between dependent and independent variables	Loss of interpretability as the number of input layers increase	
	Based on the weights in the connections, neurons feed the input data from one layer to another and hence, finally primary input data are construed into information	Has good adaptability and anti-noise capacity	Prone to overfitting as the number of model parameters increase	

Complexity

The complexity of a ML model depends on the complexity of the task that it is designed to perform. Complexity in this instance refers to the non-linearity of the underlying process/system, and consequently, the number of model parameters that are required to ‘fit’ the data.

For example, if strong maturity indexes such as starch index and firmness are available, a relatively simple model can be trained to classify the maturity of the fruit. If these indexes are not available and classification must be made on the basis of features such as colour or shape (as in the case of imagery), then a more complex model is required. A good example of this is Convolutional Neural Networks, where highly complex models are trained using hundreds of thousands of training samples. There is always a trade-off between model complexity and interpretability.

Data sensitivity

The prediction result from an ML model significantly depends on the size and variety of datasets. Interestingly, a lack of diversity in training dataset can make a model highly efficient in that particular dataset, but would make the algorithm perform poorly when a new dataset is used to estimate fruit maturity and ripening. In addition, a model developed with a small dataset can show the same problem in interpreting unseen data. In order to explain it comprehensively, the bias-variance trade-off model is illustrated in Fig. 11. A bias reflects the difference between the correct prediction value and average prediction value from the

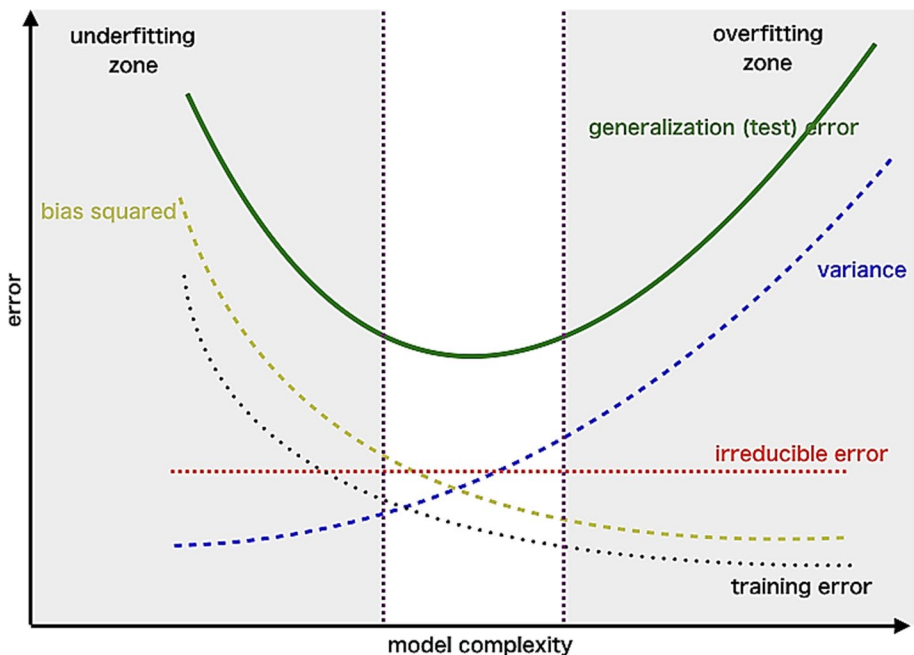


Fig. 11 Bias-variance trade-off (Papachristoudis, 2019)

designed model. An excessive consideration on bias leads the model to be oversimplified and generate high error between the testing and training data. The variance, on the other hand, describes the variability in the prediction of a model. High variance models perform well when it is introduced with unseen training data. However, at the same time, it leads to high error when testing data is applied and is responsible in making the model more complex. Therefore, to develop a well-performed model, a proper trade-off between bias and variance should be achieved and hence, generalization error should be minimized. Both the bias and variance provide useful insights on the nature of a ML model. For instance, when a model of multiple variance leads high error in terms of prediction, it may give an insight that consideration of some variances can be misleading in terms of model's performance. Therefore, PCA has been proposed to determine the most significant variances to explain a model. Similarly, bias can also help to interpret a model where two variables maintain a relationship. For instance, Root Mean Square Error of Prediction (RMSEP) and R-squared (R^2) are commonly considered to explain the relationships between two variables and that is possible with the help of the presence of bias of a model. Hence, both bias and variance can help to design a suitable model for prediction.

Interpretability

Conventional approaches assist in predicting the maturity of fruits based on some key maturity indexes in general. However, some ML approaches, like Neural Network-based ML algorithms (in particular Deep Learning), cannot natively provide interpretations or explanations of their outputs, i.e. how they arrived at a particular prediction. As such, the user may not be aware of the most significant maturity indexes that the model utilized to predict the maturity level of a given fruit sample. For example, IAD and firmness shows strong correlation in peaches and thus, it can interpret the fruit ripeness efficiently. However, SSC or SSC/TA ratio are unable to deduce any considerable correlation which can interpret the peach maturity or ripeness (Gasic et al., 2013). There is a possibility of a black-box modelled ML algorithm to choose either SSC or SSC/TA ratio or both altogether for a group of samples and able to interpret maturity or ripeness of peach. This interpretation can be local and not be applicable to other group of peaches. Being the ML algorithm a black-box model, a user is unable to identify in which features the algorithm are adding weight to culminate the result of identifying or estimating peach maturity or ripeness. While the body of knowledge in explainable AI now entails several promising techniques, further research in ML explanation methodologies focusing on smart agriculture is required to integrate this capability where some works like feature attribution methods have already been started adopting in the sector (Montavon et al., 2018; Paudel et al., 2023).

Incorrect/insignificant maturity index selection

It is very well known in horticulture that different fruits have their own ways of exposing their maturity level. Therefore, a maturity index that is significant to measure the maturity level of a fruit is not necessarily applicable to others. For example, the skin colour of a peach or apple can be deceiving in estimating their maturity since they are highly responsive to sunlight. In that case, the measurement of multiple maturity indexes can be a promising solution. Moreover, the maturity level commonly varies from cultivar to cultivar of the same type of fruits. Therefore, the same standard, considered for a fruit's maturity index to estimate its maturity level, cannot be considered to estimate the maturity level of

the same type of fruits with different cultivars; even though they are from the same orchard and plants have grown up under same environments.

In some cases, the measurements from multiple maturity indexes can lead to a wrong conclusion. For instance, considering starch index and oil content together is a wrong choice to measure the maturity level of a peach or nectarine. In contrast, these maturity indexes are highly suitable for measuring the maturity of an avocado. Therefore, choosing appropriate maturity indexes is highly significant in terms of explaining a fruit's maturity level.

External factors

Solar radiation/irradiance, water irrigation, fruits' positions in the canopy and nutrition in plants have impacts at certain extents on the fruits though they are ideally not significant. However, it is worth noting that multiple external factors altogether can lead to a wrong estimation of the fruits in an orchard in terms of their maturity. Apart from that, it is widely known that weather, season and climate are some impactful factors that remarkably are responsible in defining the fruit maturity of an orchard.

Sensor limitations

Instrumental error is a commonly known challenge in all real-life applications. Therefore, several calibrations can be a way to mitigate the problem. Apart from that, the recorded performance of a sensor at ideal conditions may help determine if the sensor has fault inside by comparison.

Trends in non-destructive approaches and ML algorithms

In order to comprehend the pace and directions of ML applications in smart agriculture, a survey on previous studies, particularly on both sensors and ML algorithms, can aid in getting the complete picture. Out of fifty-three (53) relevant articles published in the last two decades, thirty-eight (38) articles have considered colour determination approaches that include colourimetry and visible imaging Fig. 12.

In addition, five (5) articles introduce spectroscopy approaches to classify fruit based on their maturity and it includes fluorescence, vis-NIR, NIR and MIR spectroscopy. Among these six (6) articles, three (3) articles consider Vis-NIR, and the other three (3) articles consider fluorescence, NIR and MIR spectroscopy. Spectral imaging approaches have been considered in six (6) articles that entail multispectral and hyperspectral imaging approaches, where four (4) of them are hyperspectral, and the rest two (2) are multispectral imaging approaches. Acoustic impulse and electrical impulse are minimally considered approaches among the articles that classify fruit based on their maturity level. It is essential to highlight that computer vision algorithms, specifically in visible imaging and colourimetry, have increased drastically in the last decade and hence, this technology has widely been considered to classify fruits. Figure 13 illustrates the timeline of ML algorithms that have been using in fruit maturity classification based on the surveyed articles. Between 2011 and 2020, it is apparent that ML model usage has sharply increased in fruit maturity classification because of the advancement of two key enablers—namely, advances in GPU efficiencies and the increasing availability of datasets. Among ML methods, SVM, and CNN have been two widely applied algorithms over the past five (5) years.

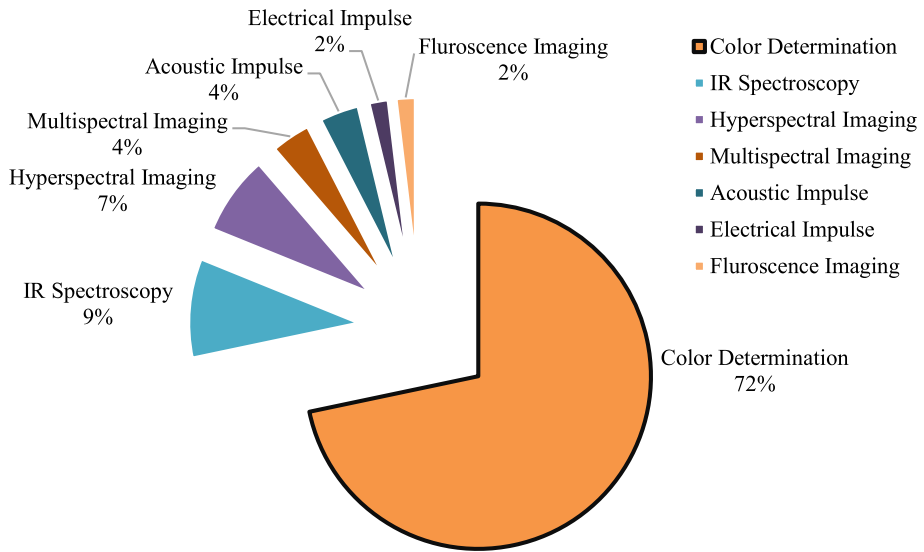


Fig. 12 Pie Chart of non-destructive approaches in classifying fruit maturity/ripening stages

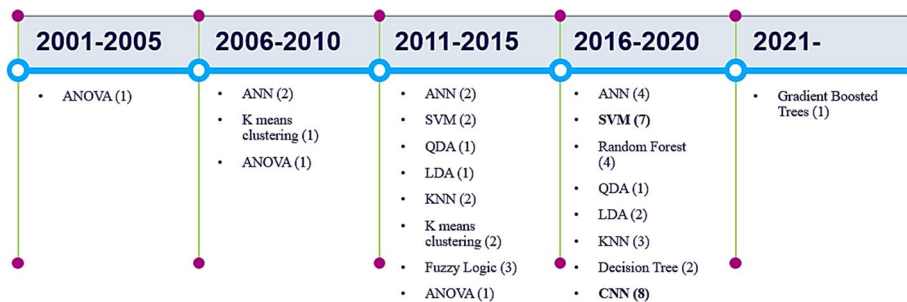


Fig. 13 Timeline of ML algorithms in fruit maturity

Apart from that, in the realm of fruit classification based on maturity, several applications of ANN can be found along with SVM and CNN in literature, as summarized in Fig. 14. Both ANN and CNN are types of deep learning mostly used in image processing applications requiring high computation. Thus, when both approaches are combined, it will cover 31% of the works which represent that majority of current works in this application are being performed by deep learning. As previously illustrated by Fig. 13, most works have focused on colour determination in the last two decades. As a result, deep learning is achieving greater interest in recent years for image processing applications due to the advancement of camera technology and availability. In addition, SVM is increasingly popularity for statistical data analysis of non-destructive sensors data. Noted that Decision tree offers overfitting with high dimensional data (Kotu & Deshpande, 2019) while QDA offers singularity to the covariance of the matrix and not fit for the variable of high dimensionality (Jiang et al., 2018). The pros and cons of the ML algorithms further have been discussed in “Instrument technology” section.

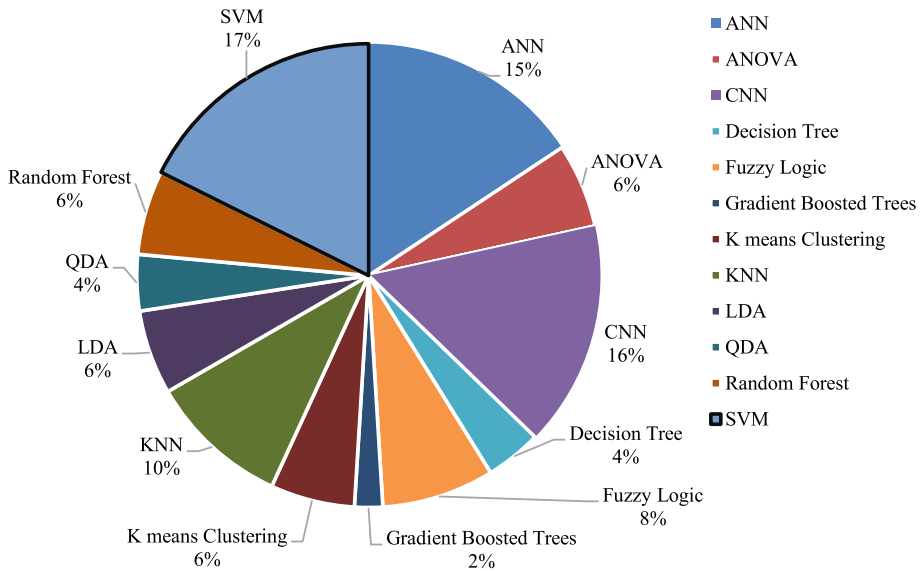


Fig. 14 ML approaches in fruit maturity

External factors

Solar radiation/irradiance, water irrigation, fruits' positions in the canopy, nutrition in plants, environment temperature and soil quality are some external factors that have impacts on the fruit quality and maturity mentioned previously. These factors have been studied before but not extensively in relation to fruit maturity. Apart from that, the impact of season and weather on fruit maturity are well-known facts in horticulture. Cirilli et al. highlight environmental factors such as irrigation, pruning and canopy management, fertilization, temperature, photoperiod, solar radiation, soil patterns, precipitation, climate, crop load etc. significantly impacts on the physiological and metabolic process of peaches, and hence, SSC and sugar contents in peaches vary (Cirilli et al., 2016). Therefore, the impacts of environmental factors offer scopes to make a relationship with fruit maturity and ripening. Henceforth, it may widen the opportunity to understand fruits' phenotype that can lead to estimate fruits' maturity more efficiently.

Instrument technology

In-situ techniques are increasingly being surpassed in accuracy and application in agriculture by remote electro-optical sensors. Emerging techniques such as refinement of spectral imaging LIDAR, and thermography in remote sensing will shape the future of crop yield estimation, fruit maturity, product quality and stress detection. Spectroscopy and spectral imaging have remarkable prospects in estimating fruit maturity as they deal with hundreds of spectral bands and provide a high-levels of understanding about an object (Kang et al., 2020). As an illustration, Wendel et al. (2018) have introduced an unmanned ground vehicle equipped with a hyperspectral camera designed to measure dry matter and estimate the maturity of mangoes within an orchard. In the course of spectral image analysis, ML algorithm, specifically a Convolutional Neural Network (CNN), were employed. In addition,

sugar contents, acidity, solid contents, oil contents and chlorophyll contents can be measured using spectroscopy that has been outlined in Table 1. Since spectral imaging can also include fruit colour, size and shape to estimate fruits' maturity level at the same time, several works combine both spatial and spectral features of an image (Fauvel et al., 2012; Ghassemian, 2016; Mirzapour & Ghassemian, 2015).

Since ML works particularly with the features, a suitable ML approach is a requirement to achieve higher accuracy. According to the survey, as shown in Fig. 14, SVM is a widely applied ML approach in fruit maturity and its competitors are ANN and CNN since most of the vision systems users choose them. Noted that neural network-based ML algorithms behave as 'black-boxes' which are challenging to interpret by human users. This is a drawback since the interpretability of the maturity indexes is important. The black-box structure gives stress on some features during model development that are in general unknown. Therefore, in this application, the stressed features must be known since the significance of the maturity indexes vary from fruit to fruit that has been discussed in previous sections. Thus, it may lead to the wrong predictions about the maturity levels of the fruits. Therefore, neural network-based ML algorithms are in general not suitable for this application until the link between weights and the functions can be revealed. In addition, these methods are computationally efficient and ideal for image processing applications. As discussed previously, to support image processing applications, additional GPUs are usually needed especially when dealing with a large volume of images. In such case, additional power supply and cooling system are frequently required.

Besides, non-neural network-based algorithms can be chosen based on their features such as prediction time, training time, required tuning, memory usage, high-dimensional data handling capacity etc. Since training and prediction time, tuning requirements and memory usage are not significantly considered because of the availability of highly computational computers, high-dimensional data handling capacity should receive the highest priority now after interpretability. From this aspect, SVM, Naïve Bayes and KNN can be suitable for this application. Particularly, SVM has achieved the reputation of classifying hyperspectral images in both spectral and spatial aspects (Chandra & Bedi, 2021; Kaul & Raina, 2022; Li et al., 2011). KNN is easy to implement but shows poor performance in prediction. Besides, Naïve Bayes considers all the features as independent and equally significant, and it is practically not considerable specifically in this application where fruits' maturity (Jahromi & Taheri, 2017).

Table 7 shows commonly used ML approaches in terms of their features that gives an understanding of choosing a suitable ML approach for this application.

Conclusions

The evolution and increasing capabilities of sensors and AI algorithms have supported their growing role in precision agriculture, particularly in the analysis of fruit maturity and ripening. Fruit maturity estimation methods include both destructive and non-destructive approaches with a variety of techniques developed over the years for a wide range of practical applications. This article categorized these methods with the aim of exploring their potential and evolving applicability. The techniques can be very-cultivar specific but recent developments have introduced progressively higher levels of flexibility and applicability. The classification of fruit maturity is typically performed with reference to suitable indices, including sugar contents, acidity, firmness and chlorophyll content. This article examined

Table 7 Outline of the available ML algorithms for fruit maturity estimation and their comparison

ML methods	High dimensionality	Maturity indexes interpretability	Practicality	Remarks
SVM	√	√	√	Easy to interpret Strong maturity indexes can be monitored and interpreted during model development Prediction and training periods are comparatively slow Requires medium memory usage
Naïve bayes	√	√	×	Able to hand high dimensionality of data Fast in prediction and training Requires medium memory usage Can be considered for image analysis
Decision trees	×	√	√	Not fit for all real-life applications since it considers all features are independent and equal Not efficient in handling high dimensional data Fast in prediction and training Requires small memory usage Prone to overfitting
K-nearest neighbours	√	√	√	Able to handle large data volumes but poor accuracy Easy to implement and interpret Moderate prediction speed and minimal training speed Requires minimum memory usage
Random forest	√	×	√	Lack of interpretability Able to handle high-dimensional data Slow in training and prediction

Table 7 (continued)

ML methods	High dimensionality	Maturity indexes interpretability	Practicality	Remarks
CNN	√	×	√	Lack of interpretability Able to handle a large volume of data Over parameterization Higher power consumption
RCNN	√	×	√	Suitable for image analysis Lack of interpretability Able to handle a large volume of data Over parameterization Higher power consumption
RNN	√	×	√	Suitable for image analysis Lack of interpretability Able to handle a large volume of data Over parameterization Higher power consumption
Faster RCNN	√	×	√	Suitable for image analysis Lack of interpretability Able to handle large volume of data Over parameterization Higher power consumption Suitable for image analysis

the mainstream ML approaches adopted in fruit maturity estimation and compared their performance in different practical applications. Neural network-based ML algorithms have received more attention than other AI techniques due to their inherent simplicity and suitability for image processing applications. Additionally, as opposed to non-neural network methods, these approaches do not require manual feature extraction from the raw data, which helps to streamline the estimation process. However, neural network algorithms usually have a black-box structure that prevents an immediate human-readable interpretation of their workings. These approaches are widely adopted for image-based applications like colorimetry or digital imaging application but not fit in other applications where multiple sensor streams are monitored and interfered sometimes. In that case, SVM and KNN can be preferable since they have the ability to deal with large datasets and are inherently more human-interpretable. SVM have been widely used for dealing with spectral and spatial data and offers satisfactory results whereas KNN is poor in accuracy sometimes.

Ongoing research is focusing on using non-destructive spectroscopy and spectral imaging to analyze sugar, acidity, firmness and chlorophyll contents, particularly from remote sensors. This is a major trend in the smart agriculture field, which has observed a move away from destructive testing, despite them being traditionally more accurate. Remote sensing, conversely, can offer a faster and wider-area overview of the entire yield, which is also practical for other post-harvest activities and sorting. Specifically, there is a considerable body of research addressing spectroscopy and spectral imaging techniques to increase both their accuracy and widen their applicability. For instance, measuring vegetation indexes to assess the maturity of individual fruits is a promising area of application of multispectral and hyperspectral imaging. Since different spectral bands have specific capabilities to characterize the reflectance properties of different surface features and considering the rapid technological advancements in optics, vegetation indexes and multi/hyper-spectral analysis can deliver more information about the maturity of fruits. Therefore, further research is recommended to explore their efficacy in this important application domain. Further advancements in efficient and effective ML algorithms tailored for the PA context are also recommended to best match the characteristics of fruit maturity-related sensor data and of the natural maturation process.

Acknowledgements This project was sponsored by the Food Agility Cooperative Research Centre (CRC) under Project No. FA042. The Food Agility CRC is funded by the Australian Government CRC Program, which supports medium to long-term industry-led research collaborations to improve the competitiveness, productivity and sustainability of Australian industries.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data availability This review article does not feature data of scientific significance for further research, so no data is made available. The readers are recommended to refer directly to the numerous works cited in this article to obtain the related data.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abshire, J. B., Riris, H., Allan, G. R., Weaver, C. J., Mao, J., Sun, X., Hasselbrack, W. E., Yu, A., Amediek, A., & Choi, Y. (2010). A lidar approach to measure CO₂ concentrations from space for the ASCENDS Mission. In *Lidar technologies, techniques, and measurements for atmospheric remote sensing VI*.
- Akerkar, R. (2019). *Artificial intelligence for business*. Springer.
- Ali, M. M., Hashim, N., & Hamid, A. S. A. (2020). Combination of laser-light backscattering imaging and computer vision for rapid determination of oil palm fresh fruit bunches maturity. *Computers and Electronics in Agriculture*, *169*, 105235.
- Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., & Aljaaf, A. J. (2020). A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised and Unsupervised Learning for Data Science*, 3–21.
- Alreshidi, E. (2019). Smart sustainable agriculture (SSA) solution underpinned by internet of things (IoT) and artificial intelligence (AI). arXiv preprint [arXiv:1906.03106](https://arxiv.org/abs/1906.03106)
- Amiryousefi, M. R., Zarei, M., Azizi, M., & Mohebbi, M. (2012). Modelling some physical characteristics of pomegranate (*Punica granatum* L.) fruit during ripening using artificial neural network. *Journal of Agricultural Science and Technology*.
- Amra, I. A. A., & Maghari, A. Y. (2017). Students performance prediction using KNN and Naïve Bayesian. In *2017 8th international conference on information technology (ICIT)*.
- Andujar, D. (2023). Back to the future: What is trending on precision agriculture? *Agronomy*, *13*(8), 2069.
- Arendse, E., Fawole, O. A., Magwaza, L. S., & Opara, U. L. (2016). Non-destructive characterization and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. *Journal of Food Engineering*, *186*, 42–49.
- Artigue, H., & Smith, G. (2019). The principal problem with principal components regression. *Cogent Mathematics & Statistics*, *6*(1), 1622190.
- Attonaty, J.-M., Chatelin, M.-H., & Garcia, F. (1999). Interactive simulation modeling in farm decision-making. *Computers and Electronics in Agriculture*, *22*(2–3), 157–170.
- Ayllon, M. A., Cruz, M. J., Mendoza, J. J., & Tomas, M. C. (2019). Detection of overall fruit maturity of local fruits using convolutional neural networks through image processing. In *Proceedings of the 2nd international conference on computing and big data*.
- Azadnia, R., Fouladi, S., & Jahanbakhshi, A. (2023). Intelligent detection and waste control of hawthorn fruit based on ripening level using machine vision system and deep learning techniques. *Results in Engineering*, *17*, 100891.
- Azarmdel, H., Jahanbakhshi, A., Mohtasebi, S. S., & Muñoz, A. R. (2020). Evaluation of image processing technique as an expert system in mulberry fruit grading based on ripeness level using artificial neural networks (ANNs) and support vector machine (SVM). *Postharvest Biology and Technology*, *166*, 111201.
- Bahri, O., Mourhir, A., & Papageorgiou, E. I. (2020). Integrating fuzzy cognitive maps and multi-agent systems for sustainable agriculture. *Euro-Mediterranean Journal for Environmental Integration*, *5*(1), 1–10.
- Bai, X., Wang, X., Liu, X., Liu, Q., Song, J., Sebe, N., & Kim, B. (2021). Explainable deep learning for efficient and robust pattern recognition: A survey of recent developments. *Pattern Recognition*, *120*, 108102.
- Balasundram, S., Robert, P., & Mulla, D. (2006). Relationship between oil content and fruit surface color in oil palm (*Elaeis guineensis* Jacq.). *Journal of Plant Sciences*, *1*(3), 217–227.
- Batista-Silva, W., Nascimento, V. L., Medeiros, D. B., Nunes-Nesi, A., Ribeiro, D. M., Zsögön, A., & Araújo, W. L. (2018). Modifications in organic acid profiles during fruit development and ripening: Correlation or causation? *Frontiers in Plant Science*, *9*, 1689.
- Benady, M. (1994). *Sensory perception for an agricultural robot*. Purdue University.
- Beullens, K., Mészáros, P., Vermeir, S., Kirsanov, D., Legin, A., Buysens, S., Cap, N., Nicolai, B. M., & Lammertyn, J. (2008). Analysis of tomato taste using two types of electronic tongues. *Sensors and Actuators b: Chemical*, *131*(1), 10–17.

- Bijjahalli, S., Sabatini, R., & Gardi, A. (2020). Advances in intelligent and autonomous navigation systems for small UAS. *Progress in Aerospace Sciences*, *115*, 100617.
- Blakey, R., & Van Rooyen, Z. (2011). Non-destructive measurement of moisture content using handheld NIR. *South African Avocado Growers Association Yearbook*, *34*, 9–11.
- Blankenship, S. M., Ellsworth, D. D., & Powell, R. L. (1993). A ripening index for banana fruit based on starch content. *HortTechnology*, *3*(3), 338–339.
- Bonaccorso, G. (2018). *Machine Learning Algorithms: Popular algorithms for data science and machine learning*. Packt Publishing Ltd.
- Bron, I. U., Ribeiro, R. V., Azzolini, M., Machado, E. C., & Jacomino, A. P. (2005). Chlorophyll fluorescence emission and its relation to skin color and firmness during ripening of guava fruit. *Fruits*, *60*(1), 25–32.
- Brookfield, P., Murphy, P., Harker, R., & MacRae, E. (1997). Starch degradation and starch pattern indices; interpretation and relationship to maturity. *Postharvest Biology and Technology*, *11*(1), 23–30.
- Brovelli, E., & Cisneros-Zevallos, L. (2007). Horticultural maturity revisited: From peaches to Echinaeca. *Fresh Produce*, *1*(1), 1–3.
- Bruker. (2021). *NMR spectroscopy approved for wine analysis*. Wiley Analytical Science. <https://doi.org/10.1002/was.0110273>
- Bureau, S., Ruiz, D., Reich, M., Gouble, B., Bertrand, D., Audergon, J.-M., & Renard, C. M. (2009). Application of ATR-FTIR for a rapid and simultaneous determination of sugars and organic acids in apricot fruit. *Food Chemistry*, *115*(3), 1133–1140.
- Caladcad, J. A., Cabahug, S., Catamco, M. R., Villaceran, P. E., Cosgafa, L., Cabizares, K. N., & Hermosilla, M. (2020). Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal. *Computers and Electronics in Agriculture*, *172*, 105327.
- Camuffo, D. (2019). Measuring temperature. In D. Camuffo (Ed.), *Microclimate for cultural heritage: Conservation and restoration of indoor and outdoor monuments* (pp. 383–429). Elsevier. <https://doi.org/10.1016/C2013-0-00676-7>
- Castro-Giráldez, M., Fito, P. J., Chenoll, C., & Fito, P. (2010). Development of a dielectric spectroscopy technique for the determination of apple (Granny Smith) maturity. *Innovative Food Science & Emerging Technologies*, *11*(4), 749–754.
- Chandra, M. A., & Bedi, S. (2021). Survey on SVM and their application in image classification. *International Journal of Information Technology*, *13*, 1–11.
- Chang, L., Zhang, Z., Ye, L., & Friedrich, K. (2008). Synergistic effects of nanoparticles and traditional tribo-fillers on sliding wear of polymeric hybrid composites. In *Tribology and interface engineering series* (Vol. 55, pp. 35–61). Elsevier.
- Chassagne-Berces, S., Fonseca, F., Citeau, M., & Marin, M. (2010). Freezing protocol effect on quality properties of fruit tissue according to the fruit, the variety and the stage of maturity. *LWT-Food Science and Technology*, *43*(9), 1441–1449.
- Chen, L.-Y., Wong, D.-M., Fang, C.-Y., Chiu, C.-I., Chou, T.-I., Wu, C.-C., Chiu, S.-W., & Tang, K.-T. (2018). Development of an electronic-nose system for fruit maturity and quality monitoring. In *2018 IEEE International Conference on Applied System Invention (ICASI)*.
- Chen, P., McCarthy, M., Kauten, R., Sarig, Y., & Han, S. (1993). Maturity evaluation of avocados by NMR methods. *Journal of Agricultural Engineering Research*, *55*(3), 177–187.
- Cheng, G. W., & Breen, P. J. (1992). Cell count and size in relation to fruit size among strawberry cultivars. *Journal of the American Society for Horticultural Science*, *117*(6), 946–950.
- Cheng, J.-H., Qu, J.-H., Sun, D.-W., & Zeng, X.-A. (2014). Visible/near-infrared hyperspectral imaging prediction of textural firmness of grass carp (*Ctenopharyngodon idella*) as affected by frozen storage. *Food Research International*, *56*, 190–198. <https://doi.org/10.1016/j.foodres.2013.12.009>
- Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, *151*, 61–69.
- Cirilli, M., Bassi, D., & Ciacciulli, A. (2016). Sugars in peach fruit: A breeding perspective. *Horticulture Research*, *3*(1), 1–12.
- Cocetta, G., Beghi, R., Mignani, I., & Spinardi, A. (2017). Nondestructive apple ripening stage determination using the Delta absorbance meter at harvest and after storage. *HortTechnology*, *27*(1), 54–64.
- Costamagna, F., Giordani, L., Costa, G., & Noferini, M. (2012). Use of ad index to define harvest time and characterize ripening variability at harvest in 'Gala' apple. In *EUFRIN thinning working group symposia 998*.
- Crisosto, C. H. (1994). Stone fruit maturity indices: A descriptive. *Postharvest News and Information*, *5*(6), 65–68.

- De-la-Torre, M., Zatarain, O., Avila-George, H., Muñoz, M., Oblitas, J., Lozada, R., Mejía, J., & Castro, W. (2019). Multivariate analysis and machine learning for ripeness classification of cape gooseberry fruits. *Processes*, 7(12), 928.
- Debener, J., Heinke, V., & Kriebel, J. (2023). Detecting insurance fraud using supervised and unsupervised machine learning. *Journal of Risk and Insurance*.
- DeLong, J., Prange, R., Harrison, P., Nichols, D., & Wright, H. (2014). Determination of optimal harvest boundaries for Honeycrisp™ fruit using a new chlorophyll meter. *Canadian Journal of Plant Science*, 94(2), 361–369.
- Dhall, D., Kaur, R., & Juneja, M. (2020). Machine learning: A review of the algorithms and its applications. In *Proceedings of ICRIC 2019: Recent innovations in computing* (pp. 47–63).
- Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion*, 50, 92–111.
- Doerffinger, F. C., Miller, W. B., Nock, J. F., & Watkins, C. B. (2015a). Relationships between starch pattern indices and starch concentrations in four apple cultivars. *Postharvest Biology and Technology*, 110, 86–95.
- Doerffinger, F. C., Miller, W. B., Nock, J. F., & Watkins, C. B. (2015b). Variations in zonal fruit starch concentrations of apples—A developmental phenomenon or an indication of ripening? *Horticulture Research*, 2(1), 1–9.
- dos Santos Neto, J. P., de Assis, M. W. D., Casagrande, I. P., Júnior, L. C. C., & de Almeida Teixeira, G. H. (2017). Determination of ‘Palmer’ mango maturity indices using portable near infrared (VIS-NIR) spectrometer. *Postharvest Biology and Technology*, 130, 75–80.
- Du, K.-L. (2010). Clustering: A neural network approach. *Neural Networks*, 23(1), 89–107.
- Dull, G. (1986). Nondestructive evaluation of quality of stored fruits and vegetables. *Food Technology (USA)*.
- Durrant-Whyte, H. F. (1990). Sensor models and multisensor integration. In *Autonomous robot vehicles* (pp. 73–89). Springer.
- Elhariri, E., El-Bendary, N., Hassanien, A. E., Badr, A., Hussein, A. M., & Snašel, V. (2014). Random forests based classification for crops ripeness stages. In *Proceedings of the fifth international conference on innovations in bio-inspired computing and applications IBICA 2014*.
- El-Mesery, H. S., Mao, H., & Abomohra, A.E.-F. (2019). Applications of non-destructive technologies for agricultural and food products quality inspection. *Sensors*, 19(4), 846.
- Facchini srl. (2021). FT fruit pressure tester. In F. srl (Ed.).
- Fadchar, N. A., & Cruz, J. C. D. (2020). A non-destructive approach of young coconut maturity detection using acoustic vibration and neural network. In *2020 16th IEEE international colloquium on signal processing & its applications (CSPA)*.
- Fahey, T., Pham, H., Gardi, A., Sabatini, R., Stefanelli, D., Goodwin, I., & Lamb, D. W. (2021). Active and passive electro-optical sensors for health assessment in food crops. *Sensors*, 21(1), 171.
- Fan, G.-F., Guo, Y.-H., Zheng, J.-M., & Hong, W.-C. (2019). Application of the weighted k-nearest neighbor algorithm for short-term load forecasting. *Energies*, 12(5), 916.
- Fathizadeh, Z., Aboonajmi, M., & Hassan-Beygi, S. R. (2021). Nondestructive methods for determining the firmness of apple fruit flesh. *Information Processing in Agriculture*.
- Fauvel, M., Tarabalka, Y., Benediktsson, J. A., Chanussot, J., & Tilton, J. C. (2012). Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3), 652–675.
- Fawole, O. A., & Opara, U. L. (2013). Changes in physical properties, chemical and elemental composition and antioxidant capacity of pomegranate (cv. Ruby) fruit at five maturity stages. *Scientia Horticulturae*, 150, 37–46.
- Fuchs, Y., Pesis, E., & Zauberman, G. (1980). Changes in amylase activity, starch and sugars contents in mango fruit pulp. *Scientia Horticulturae*, 13(2), 155–160.
- Gardi, A., Kapoor, R., & Sabatini, R. (2017). Detection of volatile organic compound emissions from energy distribution network leaks by bistatic LIDAR. *Energy Procedia*, 110, 396–401.
- Gardi, A., Sabatini, R., & Ramasamy, S. (2016). Stand-off measurement of industrial air pollutant emissions from unmanned aircraft. In *2016 international conference on unmanned aircraft systems, ICUAS 2016*.
- Gasic, K., Reighard, G., Windham, J., & Ognjanov, M. (2013). Relationship between fruit maturity at harvest and fruit quality in peach. In *VIII international peach symposium 1084*.
- Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion*, 32, 75–89.
- Ghavami, N., Sotiriou, I., & Kosmas, P. (2019). Experimental investigation of microwave imaging as means to assess fruit quality. In *2019 13th european conference on antennas and propagation (EuCAP)*.
- Goliáš, J., Bejček, L., Graetz, P., & Klusáček, S. (2003). Mechanical resonance method for evaluation of peach fruit firmness. *Horticultural Science*, 30, 1–6.

- Gong, L., Meng, X., Liu, N., & Bi, J. (2014). Evaluation of apple quality based on principal component and hierarchical cluster analysis. *Transactions of the Chinese Society of Agricultural Engineering*, *30*(13), 276–285.
- Gorry, P. A. (1990). General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Analytical Chemistry*, *62*(6), 570–573.
- Gun, W., & Chen, K. (2010). Relationship between dielectric properties from 10 to 4 500 MHz and internal quality of peaches. *Nongye Jixie Xuebao - Transactions of the Chinese Society for Agricultural Machinery*, *41*(3), 134–138.
- Guzmán, E., Baeten, V., Pierna, J. A. F., & García-Mesa, J. A. (2015). Determination of the olive maturity index of intact fruits using image analysis. *Journal of Food Science and Technology*, *52*(3), 1462–1470.
- Habaragamuwa, H., Ogawa, Y., Suzuki, T., Shiigi, T., Ono, M., & Kondo, N. (2018). Detecting greenhouse strawberries (mature and immature), using deep convolutional neural network. *Engineering in Agriculture, Environment and Food*, *11*(3), 127–138.
- Harel, B., Parmet, Y., & Edan, Y. (2020). Maturity classification of sweet peppers using image datasets acquired in different times. *Computers in Industry*, *121*, 103274.
- Harker, F., Maingdonald, J., & Jackson, P. (1996). Penetrometer measurement of apple and kiwifruit firmness: Operator and instrument differences. *Journal of the American Society for Horticultural Science*, *121*(5), 927–936.
- Harrill, R. (1998). Using a refractometer to test the quality of fruits and vegetables. *P. PUBLISHING, Éd.) Consulté le July, 20, 2010*.
- Harrison, O. (2018). Machine learning basics with the K-nearest neighbors algorithm. *Towards data science*. <https://towardsdatascience.com/a-simple-introduction-to-k-nearest-neighbors-algorithm-b3519ed98e>
- He, Y., Zhang, Y., Pereira, A. G., Gómez, A. H., & Wang, J. (2005). Nondestructive determination of tomato fruit quality characteristics using VIS/NIR spectroscopy technique. *International Journal of Information Technology*, *11*(11), 97–108.
- Hemmings, H. C., & Hopkins, P. M. (2006). *Foundations of anesthesia: Basic sciences for clinical practice*. Elsevier.
- Huybrechts, C., Deckers, T., & Valcke, R. (2002). Predicting fruit quality and maturity of apples by fluorescence imaging: effect of ethylene and AVG. In *International conference: postharvest unlimited 599*.
- Ibba, P., Tronstad, C., Moscetti, R., Mimmo, T., Cantarella, G., Petti, L., Martinsen, Ø. G., Cesco, S., & Lugli, P. (2021). Supervised binary classification methods for strawberry ripeness discrimination from bioimpedance data. *Scientific Reports*, *11*(1), 1–13.
- Iglesias, I., Echeverria, G., & Soria, Y. (2008). Differences in fruit colour development, anthocyanin content, fruit quality and consumer acceptability of eight ‘Gala’ apple strains. *Scientia Horticulturae*, *119*(1), 32–40.
- Ignat, T., Alchanatis, V., Hoffman, A., Egozi, H., & Schmilovitch, Z. E. (2015). Sensor fusion for maturity prediction of pepper. *Chemical Engineering Transactions*, *44*, 67–72.
- Infante, R., Rubio, P., Contador, L., Noferini, M., & Costa, G. (2011). Determination of harvest maturity of D’Agen plums using the chlorophyll absorbance index. *International Journal of Agriculture and Natural Resources*, *38*(2), 199–203.
- Ip, R. H., Ang, L.-M., Seng, K. P., Broster, J., & Pratley, J. (2018). Big data and machine learning for crop protection. *Computers and Electronics in Agriculture*, *151*, 376–383.
- Isa, N. E. z. M., Amir, A., Ilyas, M. Z., & Razalli, M. S. (2017). The performance analysis of K-nearest neighbors (K-NN) algorithm for motor imagery classification based on EEG signal. In *MATEC web of conferences*.
- Issad, H. A., Aoudjit, R., & Rodrigues, J. J. (2019). A comprehensive review of Data Mining techniques in smart agriculture. *Engineering in Agriculture, Environment and Food*, *12*(4), 511–525.
- Jacobi, K. K., MacRae, E. A., & Hetherington, S. E. (1998). Early detection of abnormal skin ripening characteristics of Kensington’ mango (*Mangifera indica* Linn). *Scientia Horticulturae*, *72*(3–4), 215–225.
- Jahromi, A. H., & Taheri, M. (2017). A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features. In *2017 Artificial Intelligence and Signal Processing Conference (AISP)*.
- Janiesch, C., Zscheck, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695.
- Jantra, C., Slaughter, D. C., Roach, J., & Pathaveerat, S. (2018). Development of a handheld precision penetrometer system for fruit firmness measurement. *Postharvest Biology and Technology*, *144*, 1–8.
- Jarimopas, B., & Kitthawee, U. (2007). Firmness properties of mangoes. *International Journal of Food Properties*, *10*(4), 899–909.

- Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture*, 2, 1–12.
- Jha, S., Chopra, S., & Kingsly, A. (2007). Modeling of color values for nondestructive evaluation of maturity of mango. *Journal of Food Engineering*, 78(1), 22–26.
- Jha, S., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., & Sharma, R. (2014). Non-destructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*, 124, 152–157.
- Jhawar, J. (2016). Orange sorting by applying pattern recognition on colour image. *Procedia Computer Science*, 78, 691–697.
- Jiang, B., Wang, X., & Leng, C. (2018). A direct approach for sparse quadratic discriminant analysis. *Journal of Machine Learning Research*, 19(31), 1–37.
- Jie, D., Xie, L., Rao, X., & Ying, Y. (2014). Using visible and near infrared diffuse transmittance technique to predict soluble solids content of watermelon in an on-line detection system. *Postharvest Biology and Technology*, 90, 1–6.
- Jo, T. (2021). *Machine learning foundations: Supervised, unsupervised, and advanced learning*. Springer.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260.
- Josan, J., & Chohan, G. (1982). Studies on maturity standards of peach (*Prunus persica* Batsch) cv. Floradasun. *Haryana Journal of Horticultural Sciences*.
- Kader, A. A. (1997). Fruit maturity, ripening, and quality relationships. In *International symposium effect of pre- & postharvest factors in fruit storage* 485.
- Kafkas, E., Koşar, M., Paydaş, S., Kafkas, S., & Başer, K. (2007). Quality characteristics of strawberry genotypes at different maturation stages. *Food Chemistry*, 100(3), 1229–1236.
- Kang, X., Duan, P., & Li, S. (2020). Hyperspectral image visualization with edge-preserving filtering and principal component analysis. *Information Fusion*, 57, 130–143.
- Károlyi, A. I., Fullér, R., & Galambos, P. (2018). Unsupervised clustering for deep learning: A tutorial survey. *Acta Polytechnica Hungarica*, 15(8), 29–53.
- Kasampalis, D. S., Tsouvaltzis, P., & Siomos, A. S. (2020). Chlorophyll fluorescence, non-photochemical quenching and light harvesting complex as alternatives to color measurement, in classifying tomato fruit according to their maturity stage at harvest and in monitoring postharvest ripening during storage. *Postharvest Biology and Technology*, 161, 111036.
- Kaul, A., & Raina, S. (2022). Support vector machine versus convolutional neural network for hyperspectral image classification: A systematic review. *Concurrency and Computation: Practice and Experience*, 34(15), e6945.
- Kaur, H., Sawhney, B., & Jawandha, S. (2018). Evaluation of plum fruit maturity by image processing techniques. *Journal of Food Science and Technology*, 55(8), 3008–3015.
- Kaur, K., & Gupta, O. (2017). A machine learning approach to determine maturity stages of tomatoes. *Oriental Journal of Computer Science and Technology*, 10(3), 683–690.
- Khodabakhshian, R., Emadi, B., Khojastehpour, M., & Golzarian, M. R. (2017). Determining quality and maturity of pomegranates using multispectral imaging. *Journal of the Saudi Society of Agricultural Sciences*, 16(4), 322–331.
- Kim, Y. H., Yang, Y. J., Kim, J. S., Choi, D. S., Park, S. H., Jin, S. Y., & Park, J. S. (2018). Non-destructive monitoring of apple ripeness using an aldehyde sensitive colorimetric sensor. *Food Chemistry*, 267, 149–156.
- Kolb, C. A., Wirth, E., Kaiser, W. M., Meister, A., Riederer, M., & Pfündel, E. E. (2006). Noninvasive evaluation of the degree of ripeness in grape berries (*Vitis vinifera* L. cv Bacchus and Silvaner) by chlorophyll fluorescence. *Journal of Agricultural and Food Chemistry*, 54(2), 299–305.
- Kotani, A., Kusu, F., Takamura, K., & Hakamata, H. (2019). A portable voltammetric sensor for determining titratable acidity in foods and beverages. *Journal of the Electrochemical Society*, 167(3), 037517.
- Kotu, V., & Deshpande, B. (2019). *Data science: Concepts and practice*. Morgan Kaufmann.
- Kumar, D. P., Amgoth, T., & Annavarapu, C. S. R. (2019). Machine learning algorithms for wireless sensor networks: A survey. *Information Fusion*, 49, 1–25.
- Kusuma, A., & Putra, M. D. M. (2018). Tomato maturity classification using Naive Bayes Algorithm and histogram feature extraction. *Journal of Applied Intelligent System*, 3(1), 39–48.
- Lakshmi, S., Pandey, A., Ravi, N., Chauhan, O., Gopalan, N., & Sharma, R. (2017). Non-destructive quality monitoring of fresh fruits and vegetables. *Defence Life Science Journal*, 2(2), 103–110.
- LaRue, J. H. (1989). *Peaches, plums, and nectarines: Growing and handling for fresh market* (Vol. 3331). UCANR Publications.

- Lavee, S., & Wodner, M. (2004). The effect of yield, harvest time and fruit size on the oil content in fruits of irrigated olive trees (*Olea europaea*), cvs. Barnea and Manzanillo. *Scientia Horticulturae*, 99(3–4), 267–277.
- Lebrun, M., Plotto, A., Goodner, K., Ducamp, M.-N., & Baldwin, E. (2008). Discrimination of mango fruit maturity by volatiles using the electronic nose and gas chromatography. *Postharvest Biology and Technology*, 48(1), 122–131.
- Lee, D.-J., Archibald, J. K., Chang, Y.-C., & Greco, C. R. (2008). Robust color space conversion and color distribution analysis techniques for date maturity evaluation. *Journal of Food Engineering*, 88(3), 364–372.
- Lehman-Salada, L. (1996). Instrument and operator effects on apple firmness readings. *HortScience*, 31(6), 994–997.
- Li, B., Lecourt, J., & Bishop, G. (2018a). Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction—A review. *Plants*, 7(1), 3.
- Li, C.-H., Kuo, B.-C., Lin, C.-T., & Huang, C.-S. (2011). A spatial-contextual support vector machine for remotely sensed image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3), 784–799.
- Li, X., Wei, Y., Xu, J., Feng, X., Wu, F., Zhou, R., Jin, J., Xu, K., Yu, X., & He, Y. (2018b). SSC and pH for sweet assessment and maturity classification of harvested cherry fruit based on NIR hyperspectral imaging technology. *Postharvest Biology and Technology*, 143, 112–118.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
- Lichtenthaler, H. K. (1987). [34] Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes. *Methods in Enzymology*, 148, 350–382.
- Lindon, J. C., Nicholson, J. K., & Holmes, E. (2011). *The handbook of metabonomics and metabolomics*. Elsevier.
- Liu, H., & Zhang, S. (2012). Noisy data elimination using mutual k-nearest neighbor for classification mining. *Journal of Systems and Software*, 85(5), 1067–1074.
- Ma, J., Ma, Y., & Li, C. (2019). Infrared and visible image fusion methods and applications: A survey. *Information Fusion*, 45, 153–178.
- Machin, M., Sanguesa, J. A., Garrido, P., & Martinez, F. J. (2018). On the use of artificial intelligence techniques in intelligent transportation systems. In *2018 IEEE wireless communications and networking conference workshops (WCNCW)*.
- Macrelli, E., Romani, A., Paganelli, R. P., Sangiorgi, E., & Tartagni, M. (2013). Piezoelectric transducers for real-time evaluation of fruit firmness. Part I: Theory and development of acoustic techniques. *Sensors and Actuators A: Physical*, 201, 487–496.
- Magwaza, L. S., & Opara, U. L. (2015). Analytical methods for determination of sugars and sweetness of horticultural products—A review. *Scientia Horticulturae*, 184, 179–192.
- Magwaza, L. S., & Tesfay, S. Z. (2015). A review of destructive and non-destructive methods for determining avocado fruit maturity. *Food and Bioprocess Technology*, 8(10), 1995–2011.
- Mahesh, S., Jayas, D., Paliwal, J., & White, N. (2015). Comparison of partial least squares regression (PLSR) and principal components regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food and Bioprocess Technology*, 8(1), 31–40.
- Mao, W., & Wang, F.-Y. (2012). Chapter 8—Cultural modeling for behavior analysis and prediction. *New Advances in Intelligence and Security Informatics*, 91–102.
- Marques, E. J. N., de Freitas, S. T., Pimentel, M. F., & Pasquini, C. (2016). Rapid and non-destructive determination of quality parameters in the ‘Tommy Atkins’ mango using a novel handheld near infrared spectrometer. *Food Chemistry*, 197, 1207–1214.
- McGlone, V. A., Jordan, R. B., Seelye, R., & Clark, C. J. (2003). Dry-matter—A better predictor of the post-storage soluble solids in apples? *Postharvest Biology and Technology*, 28(3), 431–435.
- Mendoza, F., Lu, R., & Cen, H. (2012). Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content. *Postharvest Biology and Technology*, 73, 89–98.
- Middleton Spectral Vision. (2021). *FX-10e VNIR Hyperspectral Camera (GigE Interface)*. <https://www.middletonspectral.com/product/fx-10e/>
- Mijwel, M. M. (2018). Artificial neural networks advantages and disadvantages. Retrieved from LinkedIn. <https://www.linkedin.com/pulse/artificial-neuralnet.Work>
- Miki, T., Saito, K., Hayashi, S., Kajikawa, H., Shimada, M., Ogawa, R., Kawate, Y., Ikegaya, D., Kimura, N., & Takabatake, K. (1996). Nondestructive analysis of sugar content on watermelon using MRI device. *TEION KOGAKU (Journal of Cryogenics and Superconductivity Society of Japan)*, 31(5), 258–266.

- Mim, F. S., Galib, S. M., Hasan, M. F., & Jerin, S. A. (2018). Automatic detection of mango ripening stages—An application of information technology to botany. *Scientia Horticulturae*, 237, 156–163.
- Mirzapour, F., & Ghassemian, H. (2015). Improving hyperspectral image classification by combining spectral, texture, and shape features. *International Journal of Remote Sensing*, 36(4), 1070–1096.
- Mishra, P., & Woltering, E. (2023). Semi-supervised robust models for predicting dry matter in mango fruit with near-infrared spectroscopy. *Postharvest Biology and Technology*, 200, 112335.
- Misra, S., & Li, H. (2019). Noninvasive fracture characterization based on the classification of sonic wave travel times. *Machine Learning for Subsurface Characterization*, 243–287.
- Mitchell, F., Mayer, G., & Beede, R. (1977). Studies of various factors affecting postharvest performance of shipping stone fruits. *Final Report, California Tree Fruit Agreement*.
- Mitchell, H. B. (2007). *Multi-sensor data fusion: An introduction*. Springer.
- Mizrach, A., Flitsanov, U., El-Batsri, R., & Degani, C. (1999). Determination of avocado maturity by ultrasonic attenuation measurements. *Scientia Horticulturae*, 80(3–4), 173–180.
- Mohammed, M., Khan, M. B., & Bashier, E. B. M. (2016). *Machine learning: algorithms and applications*. Crc Press.
- Moing, A., Svanella, L., Rolin, D., Gaudillère, M., Gaudillère, J.-P., & Monet, R. (1998). Compositional changes during the fruit development of two peach cultivars differing in juice acidity. *Journal of the American Society for Horticultural Science*, 123(5), 770–775.
- Monash University Malaysia. (2021). *High Performance Liquid Chromatography (HPLC)*. <https://www.monash.edu.my/science/research/research-facilities/applied-chemistry/hplc-autosampler>
- Montavon, G., Samek, W., & Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73, 1–15.
- Moradi, G., Shamsi, M., Sedaaghi, M. H., & Alsharif, M. R. (2012). Pomegranate MR image analysis using fuzzy clustering algorithms. *Agricultural Engineering International: CIGR Journal*, 14(3), 152–160.
- Morrison, D., & Abeyratne, U. (2014). Ultrasonic technique for non-destructive quality evaluation of oranges. *Journal of Food Engineering*, 141, 107–112.
- Mostaco, G. M., De Souza, I. R. C., Campos, L. B., & Cugnasca, C. E. (2018). AgronomoBot: a smart answering Chatbot applied to agricultural sensor networks. In *14th international conference on precision agriculture*.
- Murase, H. (2000). Artificial intelligence in agriculture. *Computers and Electronics in Agriculture*, 29(1/2).
- Mushtaq, M. S., & Mellouk, A. (2017). *Quality of experience paradigm in multimedia services: Application to OTT video streaming and VoIP services*. Elsevier.
- Ncama, K., Magwaza, L. S., Poblete-Echeverría, C. A., Nieuwoudt, H. H., Tesfay, S. Z., & Mditshwa, A. (2018). On-tree indexing of ‘Hass’ avocado fruit by non-destructive assessment of pulp dry matter and oil content. *Biosystems Engineering*, 174, 41–49.
- Neto, A. F., Olivier, N. C., Cordeiro, E. R., & de Oliveira, H. P. (2017). Determination of mango ripening degree by electrical impedance spectroscopy. *Computers and Electronics in Agriculture*, 143, 222–226.
- Nicolai, B. M., Verlinden, B. E., Desmet, M., Saevels, S., Saeys, W., Theron, K., Cubeddu, R., Pifferi, A., & Torricelli, A. (2008). Time-resolved and continuous wave NIR reflectance spectroscopy to predict soluble solids content and firmness of pear. *Postharvest Biology and Technology*, 47(1), 68–74.
- Nielsen, S. S. (2017). *Food analysis*.
- Nor, F. M., Clarkson, M., & Othman, H. (2014). An improved ring method for calibration of hydrometers. *Measurement*, 48, 1–5.
- Offermann, S., Bicanic, D., Claude Krapez, J., Balageas, D., Gerkema, E., Chirtoc, M., Egee, M., Keijzer, K., & Jalink, H. (1998). Infrared transient thermography for non-contact, nondestructive inspection of whole and dissected apples and of cherry tomatoes at different maturity stages. *Instrumentation Science & Technology*, 26(2–3), 145–155.
- Olarewaju, O. O., Bertling, I., & Magwaza, L. S. (2016). Non-destructive evaluation of avocado fruit maturity using near infrared spectroscopy and PLS regression models. *Scientia Horticulturae*, 199, 229–236.
- Pacheco, W. D. N., & López, F. R. J. (2019). Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-Means Clustering. In *2019 XXII symposium on image, signal processing and artificial vision (STSIVA)*.
- Palmer, J. W., Harker, F. R., Tustin, D. S., & Johnston, J. (2010). Fruit dry matter concentration: A new quality metric for apples. *Journal of the Science of Food and Agriculture*, 90(15), 2586–2594.
- Pandian, A. P. (2021). Performance evaluation and comparison using deep learning techniques in sentiment analysis. *Journal of Soft Computing Paradigm*, 3(2), 123–134.
- Papachristoudis, G. (2019). *The Bias-Variance Tradeoff*. <https://towardsdatascience.com/the-bias-variance-tradeoff-8818f41e39e9>

- Parvathi, S., & Selvi, S. T. (2021). Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosystems Engineering*, 202, 119–132.
- Pascual, M., Villar, J., Rufat, J., Rosell, J., Sanz, R., & Arno, J. (2009). Evaluation of peach tree growth characteristics under different irrigation strategies by LIDAR system: Preliminary results. In *VI International symposium on irrigation of horticultural crops* 889.
- Pathan, M., Patel, N., Yagnik, H., & Shah, M. (2020). Artificial cognition for applications in smart agriculture: A comprehensive review. *Artificial Intelligence in Agriculture*.
- Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153, 69–81.
- Paudel, D., de Wit, A., Boogaard, H., Marcos, D., Osinga, S., & Athanasiadis, I. N. (2023). Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*, 206, 107663.
- Paz, P., Sánchez, M.-T., Pérez-Marín, D., Guerrero, J.-E., & Garrido-Varo, A. (2008). Nondestructive determination of total soluble solid content and firmness in plums using near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry*, 56(8), 2565–2570.
- Pech, J.-C., Bouzayen, M., & Latché, A. (2008). Climacteric fruit ripening: Ethylene-dependent and independent regulation of ripening pathways in melon fruit. *Plant Science*, 175(1–2), 114–120.
- Peirs, A., Scheerlinck, N., De Baerdemaeker, J., & Nicolai, B. M. (2003). Starch index determination of apple fruit by means of a hyperspectral near infrared reflectance imaging system. *Journal of near Infrared Spectroscopy*, 11(5), 379–389.
- Peleg, K., Ben-Hanan, U., & Hinga, S. (1990). Classification of avocado by firmness and maturity. *Journal of Texture Studies*, 21(2), 123–140.
- Peng, Y., & Lu, R. (2008). Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. *Postharvest Biology and Technology*, 48(1), 52–62.
- Pereira, L. F. S., Barbon, S., Jr., Valous, N. A., & Barbin, D. F. (2018). Predicting the ripening of papaya fruit with digital imaging and random forests. *Computers and Electronics in Agriculture*, 145, 76–82.
- Pereira, T., Tijkskens, L., Vanoli, M., Rizzolo, A., Eccher Zerbini, P., Torricelli, A., Spinelli, L., & Filgueiras, H. (2009). Assessing the harvest maturity of Brazilian mangoes. In *International symposium postharvest pacifica 2009-pathways to quality: V international symposium on managing quality in 880*.
- Perez, J. A., Deligianni, F., Ravi, D., & Yang, G.-Z. (2018). Artificial intelligence and robotics. arXiv preprint [arXiv:1803.10813](https://arxiv.org/abs/1803.10813), 147.
- Perkins-Veazie, P., Clark, J., Huber, D., & Baldwin, E. (2000). Ripening physiology in Navaho thornless blackberries: Color, respiration, ethylene production, softening, and compositional changes. *Journal of the American Society for Horticultural Science*, 125(3), 357–363.
- Pham, H., Gardi, A., Lim, Y., Sabatini, R., & Pang, E. (2019). UAS mission design for early plant disease detection. In *AIAC18: 18th Australian international aerospace congress (2019): HUMS-11th defence science and technology (DST) international conference on health and usage monitoring (HUMS 2019): ISSFD-27th international symposium on space flight dynamics (ISSFD)*.
- Plapinger, T. (2017). *What is a Decision Tree?* <https://towardsdatascience.com/what-is-a-decision-tree-22975f00f3e1>
- Podrazka, M., Baczynska, E., Kundys, M., Jelen, P. S., & Nery, E. W. (2018). Electronic Tongue—A Tool for All Tastes? *Biosensors*, 8(1), 3. <https://doi.org/10.3390/bios8010003>
- Prabha, D. S., & Kumar, J. S. (2015). Assessment of banana fruit maturity by image processing technique. *Journal of Food Science and Technology*, 52(3), 1316–1327.
- Prasad, K., Jacob, S., & Siddiqui, M. W. (2018). Fruit maturity, harvesting, and quality standards. In *Pre-harvest modulation of postharvest fruit and vegetable quality* (pp. 41–69). Elsevier.
- Pu, Y.-Y., Sun, D.-W., Buccheri, M., Grassi, M., Cattaneo, T. M., & Gowen, A. (2019). Ripeness classification of bananito fruit (*Musa acuminata*, AA): A comparison study of visible spectroscopy and hyperspectral imaging. *Food Analytical Methods*, 12(8), 1693–1704.
- Raghavan, V. V., Gudivada, V. N., Govindaraju, V., & Rao, C. R. (2016). *Cognitive computing: Theory and applications*. Elsevier.
- Rajkumar, P., Wang, N., Elmasry, G., Raghavan, G., & Garipey, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering*, 108(1), 194–200.
- Rani, P. M. N., Rajesh, T., & Saravanan, R. (2011). Expert systems in agriculture: A review. *Journal of Computer Science and Applications*, 3(1), 59–71.
- Rawle, A. F. (2017). *Refractive index measurements*. In J. C. Lindon, G. E. Tranter, & D. W. Koppenaal, (Eds.). Elsevier. <https://doi.org/10.1016/B978-0-12-409547-2.12090-6>
- Reid, M. S. (2002). Maturation and maturity indices. *Postharvest Technology of Horticultural Crops*, 3, 55–62.


- Reyes, S. O. L. (2020). Artificial intelligence in precision health: Systems in practice. In *Artificial intelligence in precision health* (pp. 499–519). Elsevier.
- Richmond, M. L., Brandao, S. C., Gray, J. I., Markakis, P., & Stine, C. M. (1981). Analysis of simple sugars and sorbitol in fruit by high-performance liquid chromatography. *Journal of Agricultural and Food Chemistry*, 29(1), 4–7.
- Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582–599.
- Saeed, O. M. B., Sankaran, S., Shariff, A. R. M., Shafri, H. Z. M., Ehsani, R., Alfatni, M. S., & Hazir, M. H. M. (2012). Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. *Computers and Electronics in Agriculture*, 82, 55–60.
- Sahu, D., & Poddar, R. M. (2017). Defect identification and maturity detection of mango fruits using image analysis. *American Journal of Artificial Intelligence*, 1(1), 5–14.
- Salcedo-Sanz, S., Ghamisi, P., Piles, M., Werner, M., Cuadra, L., Moreno-Martínez, A., Izquierdo-Verdiguier, E., Muñoz-Marí, J., Mosavi, A., & Camps-Valls, G. (2020). Machine learning information fusion in Earth observation: A comprehensive review of methods, applications and data sources. *Information Fusion*, 63, 256–272.
- Saranwong, S., Sornsrivichai, J., & Kawano, S. (2004). Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. *Postharvest Biology and Technology*, 31(2), 137–145.
- Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160.
- Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In *Emerging technology in modelling and graphics* (pp. 99–111). Springer.
- Shrivastava, D., Sanyal, S., Maji, A. K., & Kandari, D. (2020). Bone cancer detection using machine learning techniques. In *Smart healthcare for disease diagnosis and prevention* (pp. 175–183). Elsevier.
- Skobelev, P., Budaev, D., Gusev, N., & Voschuk, G. (2018). Designing multi-agent swarm of uav for precise agriculture. In *International conference on practical applications of agents and multi-agent systems*.
- Song, J., Deng, W., Beaudry, R. M., & Armstrong, P. R. (1997). Changes in chlorophyll fluorescence of apple fruit during maturation, ripening, and senescence. *HortScience*, 32(5), 891–896.
- Srivastava, R., Talluri, S., Beebi, S. K., & Kumar, B. R. (2018). Magnetic resonance imaging for quality evaluation of fruits: A review. *Food Analytical Methods*, 11(10), 2943–2960.
- Srivastava, S., Vani, B., & Sadistap, S. (2020). Handheld, smartphone based spectrometer for rapid and nondestructive testing of citrus cultivars. *Journal of Food Measurement and Characterization*, 1–13.
- Stefanouadaki, E., Kotsifaki, F., & Koutsaftakis, A. (1999). Classification of virgin olive oils of the two major Cretan cultivars based on their fatty acid composition. *Journal of the American Oil Chemists' Society*, 76(5), 623–626.
- Steinmetz, V., Roger, J., Molto, E., & Blasco, J. (1999a). On-line fusion of colour camera and spectrophotometer for sugar content prediction of apples. *Journal of Agricultural Engineering Research*, 73(2), 207–216.
- Steinmetz, V., Sevilla, F., & Bellon-Maurel, V. (1999b). A methodology for sensor fusion design: Application to fruit quality assessment. *Journal of Agricultural Engineering Research*, 74(1), 21–31.
- Subedi, P., & Walsh, K. B. (2011). Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. *Postharvest Biology and Technology*, 62(3), 238–245.
- Sugiyama, J., & Tsuta, M. (2010). Visualization of sugar distribution of melons by hyperspectral technique. In *Hyperspectral imaging for food quality analysis and control* (pp. 349–368). Elsevier.
- Sumridetchkajorn, S., & Intaravanne, Y. (2013). Two-dimensional fruit ripeness estimation using thermal imaging. In *ICPS 2013: International conference on photonics solutions*.
- Suresha, M., Kumar, K., & Kumar, G. S. (2012). Texture features and decision trees based vegetables classification. *International Journal of Computer Applications*, 975, 8878.
- Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture*.
- Tantanawat, T., Srimongkol, S., Yuttawiriya, R., Haewchin, A., Sangkapes, T., & Hattha, E. (2020). Development of a method for measuring oil content in oil palm mesocarp using a single-outlet piston press: A feasibility study. *Journal of Food Measurement and Characterization*, 14(1), 207–219.
- Thomopoulos, R., Croitoru, M., & Tamani, N. (2015). Decision support for agri-food chains: A reverse engineering argumentation-based approach. *Ecological Informatics*, 26, 182–191.
- Thompson, N. C., Greenwald, K., Lee, K., & Manso, G. F. (2020). The computational limits of deep learning. arXiv preprint [arXiv:2007.05558](https://arxiv.org/abs/2007.05558)

- Thornton, C. (2021). *Machine Learning—Lecture 4: The Naive Bayes Classifier*. <http://users.sussex.ac.uk/~christ/crs/ml/lec02b.html>
- Torres, I., Pérez-Marín, D., De la Haba, M.-J., & Sánchez, M.-T. (2015). Fast and accurate quality assessment of Raf tomatoes using NIRS technology. *Postharvest Biology and Technology*, *107*, 9–15.
- Travers, S., Bertelsen, M. G., Petersen, K. K., & Kucheryavskiy, S. V. (2014). Predicting pear (cv. Clara Frijs) dry matter and soluble solids content with near infrared spectroscopy. *LWT-Food Science and Technology*, *59*(2), 1107–1113.
- Trebolazabala, J., Maguregui, M., Morillas, H., de Diego, A., & Madariaga, J. M. (2017). Portable Raman spectroscopy for an in-situ monitoring the ripening of tomato (*Solanum lycopersicum*) fruits. *Spectrochimica Acta Part a: Molecular and Biomolecular Spectroscopy*, *180*, 138–143.
- Trilles, S., Torres-Sospedra, J., Belmonte, Ó., Zarazaga-Soria, F. J., González-Pérez, A., & Huerta, J. (2020). Development of an open sensorized platform in a smart agriculture context: A vineyard support system for monitoring mildew disease. *Sustainable Computing: Informatics and Systems*, *28*, 100309.
- Trimble, S. (2019). Spectroscopy in Agriculture. <https://felixinstruments.com/blog/spectroscopy-in-agriculture/>
- Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–1231.
- Tu, S., Xue, Y., Zheng, C., Qi, Y., Wan, H., & Mao, L. (2018). Detection of passion fruits and maturity classification using Red-Green-Blue Depth images. *Biosystems Engineering*, *175*, 156–167.
- Tyl, C., & Sadler, G. D. (2017). pH and titratable acidity. In *Food analysis* (pp. 389–406). Springer.
- Udomkun, P., Nagle, M., Mahayothee, B., & Müller, J. (2014). Laser-based imaging system for non-invasive monitoring of quality changes of papaya during drying. *Food Control*, *42*, 225–233.
- Vanoli, M., & Buccheri, M. (2012). Overview of the methods for assessing harvest maturity. *Stewart Postharvest Review*, *8*(1), 1–11.
- Varga, L. A., Makowski, J., & Zell, A. (2021). Measuring the ripeness of fruit with hyperspectral imaging and deep learning. arXiv preprint [arXiv:2104.09808](https://arxiv.org/abs/2104.09808)
- Wajid, A., Singh, N. K., Junjun, P., & Mughal, M. A. (2018). Recognition of ripe, unripe and scaled condition of orange citrus based on decision tree classification. In *2018 international conference on computing, mathematics and engineering technologies (iCoMET)*.
- Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*, *146*, 43–50.
- Wang, Y.-W., Wang, J., Yao, C., & Lu, Q.-J. (2009). Firmness measurement of peach by impact force response. *Journal of Zhejiang University SCIENCE B*, *10*(12), 883–889.
- Washington State University. (2021). *WSU Tree Fruit*. Washington State University. <http://treefruit.wsu.edu/orchard-management/harvest/>
- Wendel, A., Underwood, J., & Walsh, K. (2018). Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture*, *155*, 298–313.
- Wisskirchen, G., Biacabe, B. T., Bormann, U., Muntz, A., Niehaus, G., Soler, G. J., & von Brauchitsch, B. (2017). Artificial intelligence and robotics and their impact on the workplace. *IBA Global Employment Institute*, *11*(5), 49–67.
- Xia, Y. (2020). Correlation and association analyses in microbiome study integrating multiomics in health and disease. *Progress in Molecular Biology and Translational Science*, *171*, 309–491.
- Yossy, E. H., Pranata, J., Wijaya, T., Hermawan, H., & Budiharto, W. (2017). Mango fruit sortation system using neural network and computer vision. *Procedia Computer Science*, *116*, 596–603.
- Yu, K.-H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, *2*(10), 719–731.
- Zakaria, A., Shakaff, A. Y. M., Masnan, M. J., Saad, F. S. A., Adom, A. H., Ahmad, M. N., Jaafar, M. N., Abdullah, A. H., & Kamarudin, L. M. (2012). Improved maturity and ripeness classifications of *Magnifera indica* cv. harumanis mangoes through sensor fusion of an electronic nose and acoustic sensor. *Sensors*, *12*(5), 6023–6048.
- Zhang, B., Peng, B., Zhang, C., Song, Z., & Ma, R. (2017). Determination of fruit maturity and its prediction model based on the pericarp index of absorbance difference (I_{AD}) for peaches. *PLoS ONE*, *12*(5), e0177511.
- Zhang, H., Weng, T.-W., Chen, P.-Y., Hsieh, C.-J., & Daniel, L. (2018). Efficient neural network robustness certification with general activation functions. arXiv preprint [arXiv:1811.00866](https://arxiv.org/abs/1811.00866)
- Zhang, L., & McCarthy, M. J. (2013). Assessment of pomegranate postharvest quality using nuclear magnetic resonance. *Postharvest Biology and Technology*, *77*, 59–66.

- Zhang, P., Wei, Y., Xu, F., Wang, H., Chen, M., & Shao, X. (2020). Changes in the chlorophyll absorbance index (I AD) are related to peach fruit maturity. *New Zealand Journal of Crop and Horticultural Science*, 48(1), 34–46.
- Zhang, Y., & Wu, L. (2012). Classification of fruits using computer vision and a multiclass support vector machine. *Sensors*, 12(9), 12489–12505.
- Zhu, X., Ma, H., Ji, J., Jin, X., Zhao, K., & Zhang, K. (2020). Detecting and identifying blueberry canopy fruits based on Faster R-CNN. *Journal of Southern Agriculture*, 51(6), 1493–1501.
- Zion, B., Chen, P., & McCarthy, M. J. (1995). Nondestructive quality evaluation of fresh prunes by NMR spectroscopy. *Journal of the Science of Food and Agriculture*, 67(4), 423–429.
- Ziosi, V., Noferini, M., Fiori, G., Tadiello, A., Trainotti, L., Casadoro, G., & Costa, G. (2008). A new index based on vis spectroscopy to characterize the progression of ripening in peach fruit. *Post-harvest Biology and Technology*, 49(3), 319–329.
- Zude, M., Herold, B., Roger, J.-M., Bellon-Maurel, V., & Landahl, S. (2006). Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. *Journal of Food Engineering*, 77(2), 254–260.
- Zulkifli, Z. M., Hashim, F. H., Raj, T., & Huddin, A. B. (2018). A rapid and non-destructive technique in determining the ripeness of oil palm fresh fruit bunch (FFB). *Jurnal Kejuruteraan*, 30(1), 93–101.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Maidul Islam^{1,5} · Suraj Bijjahalli² · Thomas Fahey^{1,5} · Alessandro Gardi^{1,3,5}  · Roberto Sabatini^{1,3,5} · David W. Lamb^{4,5}

✉ Roberto Sabatini
roberto.sabatini@ku.ac.ae

¹ School of Engineering, RMIT University, Melbourne, VIC 3000, Australia

² University of Sydney, Sydney, NSW 2000, Australia

³ Present Address: Department of Aerospace Engineering, Khalifa University, Abu Dhabi, United Arab Emirates

⁴ Precision Agriculture Research Group, University of New England, Armidale, NSW 2351, Australia

⁵ Food Agility CRC Ltd., 81 Broadway, Sydney, NSW 2007, Australia