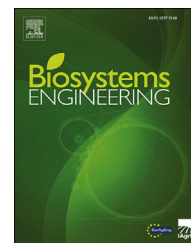


Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Review

Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review



Jana Wieme^a, Kaveh Mollazade^{b,c}, Ioannis Malounas^d,
Manuela Zude-Sasse^c, Ming Zhao^e, Aoife Gowen^e,
Dimitrios Argyropoulos^e, Spyros Fountas^d, Jonathan Van Beek^{a,*}

^a Technology and Food Science Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO),
Burgemeester Van Gansberghelaan 115 bus 1, 9820 Merelbeke, Belgium

^b Department of Biosystems Engineering, Faculty of Agriculture, University of Kurdistan, 66177-15175 Sanandaj,
Iran

^c Department of Horticultural Engineering, Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB),
Max-Eyth-Allee 100, 14469 Potsdam-Bornim, Germany

^d Agricultural University of Athens (AUA), Iera Odos 75, 11855 Athens, Greece

^e School of Biosystems and Food Engineering, University College Dublin (UCD), Stillorgan Rd, Belfield, Dublin 4,
Ireland

ARTICLE INFO

Article history:

Received 25 February 2022

Received in revised form

11 July 2022

Accepted 22 July 2022

Published online 5 September 2022

Keywords:

Hyperspectral imaging

Fruit

Vegetables

Mushrooms

Artificial intelligence

Quality assessment

Over the last two decades, research in hyperspectral imaging has been increasing and its use in horticulture is expected to be spreading in the coming years. The emerging techniques are currently gaining interest of the research community. However, there are still challenges to the applicability. In this review we demonstrate that hyperspectral imaging can be used as an effective tool for fruit, vegetables and mushrooms in assessing quality parameters related to well defined variables that can be analysed in the laboratory, as well as complex properties such as maturity, ripeness, detection of biotic defects, physiological disorders, mechanical damages, and sensory quality. Therefore, this paper starts by giving an overview of the quality concept of produce, measuring principles, theory and analysis of hyperspectral imaging systems. Then, emerging techniques to monitor and assess quality parameters, both pre- and postharvest, are described, as well as applications of these are reviewed and discussed. Afterwards, this review proceeds by illustrating the current and potential use of artificial intelligence and its subdomains, machine learning and deep learning, for hyperspectral imaging analysis in horticulture. Lastly, some challenges and considerations for future research are highlighted, including improvement of data availability, possible solutions for an improved integration of artificial intelligence and the transfer of knowledge from research parameters to parameters relevant for industrial stakeholders.

© 2022 The Authors. Published by Elsevier Ltd on behalf of IAGrE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail addresses: jana.wieme@ilvo.vlaanderen.be (J. Wieme), k.mollazade@uok.ac.ir (K. Mollazade), gmalounas@aua.gr (I. Malounas), mzude@atb-potsdam.de (M. Zude-Sasse), ming.zhao@ucd.ie (M. Zhao), aoife.gowen@ucd.ie (A. Gowen), dimitrios.argyropoulos@ucd.ie (D. Argyropoulos), sfountas@aua.gr (S. Fountas), jonathan.vanbeek@ilvo.vlaanderen.be (J. Van Beek).

<https://doi.org/10.1016/j.biosystemseng.2022.07.013>

1537-5110/© 2022 The Authors. Published by Elsevier Ltd on behalf of IAGrE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Nomenclature			
R_c^2	Coefficient of determination in calibration	PC	Parallelepiped classification
R_p^2	Coefficient of determination in prediction	PCA	Principal Component Analysis
AI	Artificial Intelligence	PCIP	Pseudo-color image processing method
ANN	Artificial Neural Network	PCR	Principal Component Regression
CARS	Competitive Adaptive Reweighted Sampling	PLS-DA	Partial Least Squares Discriminant Analysis
CLS	Classical Least Squares	PLSR	Partial Least Squares Regression
CNN	Convolutional Neural Network	POD	Peroxidase
DL	Deep learning	PPO	Polyphenol oxidase
DMC	Dry matter content	QDA	Quadratic Discriminant Analysis
EAU	Enzyme activity unit	RBF	Radial Basis Function
ECR	Extended collaborative representation	RF	Random Forest
EMCVS	Ensemble Monte Carlo Variable Selection	RFr	Random frog
ExG	Excess Green	RMSE _c	Root mean square in calibration
Faster R-CNN	Faster region based convolutional neural network	RMSE _p	Root mean square in prediction
FCNN	Fully connected network	ROC	Receiver operating characteristic
GMM	Gaussian Mixture Model	ROI	Region of interest
InGaAs	Indium Gallium Arsenide	SAE	Sparse autoencoder
IP-T	Image processing-thresholding	SFS	Sequential forward selection
kNN	k-Nearest Neighbors	SLRS	Stepwise Linear Regression Selection
LDA	Linear Discriminant Analysis	SPA	Successive Projections Algorithm
LEDs	Light-Emitting Diodes	SSAE	Stacked sparse autoencoder
LS-SVM	Least squares-support vector machine	SSC	Soluble solid content
Mask R-CNN	Mask region based convolutional neural network	SVM	Support vector machine
MCR	Multivariate Curve Resolution	SVM-DA	Support Vector Machine Discriminant Analysis
MC-UVE	Monte Carlo-uninformative variable elimination	SVMR	Support Vector Machine Regression
MFA	Multifactorial analysis	SWIR	Short wave infrared
ML	Machine learning	TSS	Total soluble solids
MLP	Multilayer Perceptron	Tukey's HSD	Honestly significant difference
MLR	Multiple Linear Regression	VGG	Visual Geometry Group
MNF	Minimum noise fraction	VI	Vegetation indices
mRMR	Minimum redundancy maximum relevance	VIP	Variable Importance in Projection
OCA	Overall classification accuracy	VIS-NIR	Visible to near infrared
		XGBoost	Extreme gradient boost
		ϵ -SVMs	Epsilon-support vector machines

1. Introduction

Fresh fruit and vegetables are essential food products for human nutrition due to the balanced supply of nutrients such as sugars, organic acids, vitamins, pro-vitamins, and minerals, as well as non-nutritional, beneficial compounds such as fibres and secondary metabolites. Commercially, the global market of fresh produce has a high economic importance. Challenges along the supply chain of fresh produce occur in all stages. Production processes need to be managed sustainably, while producing high quality products. At harvest, the maturity stage is typically an important factor not only affecting the product quality at harvest, but also susceptibility of the product in postharvest. In postharvest, the main task is maintaining the good quality of perishable products as long as possible to avoid food waste and economical losses. Consequently, knowledge of factors related to produce quality in the pre-harvest, harvest, and postharvest stages is relevant for

producer, harvest manager, storage manager, in packaging facilities, direct marketing on the farm, local markets, global distributor, wholesaler, and consumer. Throughout the production to consumption chain, the term quality is used continuously, but its meaning varies depending on the stage in the chain. However, in all these stages, the term “quality” is related to the degree of excellence and the absence of defect on the product (absence of defect and blemishes, cultivar-typical ripeness, freshness, non-harmful amount of residues considering pesticides and other chemicals, and cleanliness). It frequently describes sensory properties (appearance, colour, texture, taste, and aroma) and nutritional properties (Valero & Serrano, 2010). Quality can therefore be examined from the perspective of product or from the perspective of manager and consumer.

Considering fresh fruit, in 2022, sales revenue and average per capita consumption of fresh fruit are expected to be \$ 630,683 million and 33.9 kg, respectively. Furthermore, the estimated annual growth rate of this market is 5.56% (Statista,

2021a). Quality keeping properties depend largely on the properties of the skin serving as barrier for water vapour and against mechanical stress, as well as on the physiological climacteric response of the fruit. As a result, the postharvest life varies between a few days (e.g., for soft berries) to one year (e.g., apple). From the consumer's point of view, the quality is sensory and based on the appearance of the fruit and assessed by uniformity in size, peel colour, skin gloss, perfect shape and peel, and absence of disease (Civille & Oftedal, 2012). In terms of product quality, parameters related to the physiology and maturity of fruit (such as moisture content, firmness, soluble solid content, acidity, chlorophyll, carotenoids and flavonoids pigments, ethylene production, respiration rate, etc.) are important for keeping quality measures. Such properties should be examined by analytical methods (Chen, 2015). However, appearance and physical dimensions can be analysed inline. This procedure is limited only by the mechanical susceptibility of the perishable product and economical viewpoints.

The edible parts of vegetables are commonly classified based on the portion of the plant that is consumed. Those classes are leaves (e.g., lettuce), stems (e.g., asparagus), roots (e.g., parsnip), bulbs (e.g., onion) and flowers (e.g., broccoli). The importance of vegetables can be illustrated by their expected average per capita consumption, 71.0 kg in 2022, an expected global revenue of \$ 941,818 million in 2022 and its expected annual growth rate of 5.40% (Statista, 2021b). Vegetable quality is therefore of high importance for all stakeholders across the supply chain. The chemical and physical basis for fruit and vegetables quality are colour, size, shape, wholeness, presence of defects, flavour, aroma, taste, texture, nutritional value, and the consistency of those parameters (Barrett et al., 2010). However, points for checking the product quality can differ between classes of vegetables, e.g., flower vegetables should not contain brown brushing marks while stem vegetables should have a crisp texture.

Edible mushrooms have been consumed worldwide since ancient times as a delicacy due to their specific aroma and texture, or as a component of staple sustenance during food shortage. Whereas in recent times, edible mushrooms have been considered as a boost to healthy nutrition of human beings due to its low energy level, high proportion of dietary fibre content and other supply of nutrients such as minerals, vitamins, lipids, and proteins (Kalac, 2016). While over the 2000 edible mushroom species found on the earth, only 100 can be commercially cultivated, and only 10–20 can be cultivated on an industrial scale (Kalac, 2016). Until 2019, the total world production of cultivated mushrooms reached nearly 11 million metric tons. *Agaricus bisporus* is one of the most cultivated species (Kalac, 2019) and is marketed both in immature (white button mushrooms, brown mushrooms, crimini/cremini or chestnut mushrooms) and mature state (portobello mushrooms) (Kalac, 2016). Conversely to fruit and vegetables, mushrooms are the most perishable food items. The lifetime of their fruiting bodies in nature is around 10–14 days, while shelf-life can be even shorter than 3 days at ambient temperature due to their high water content, high respiration rate, water loss under low humidity storage condition, lack of physical protection from mechanical damage during transportation and other microbial attack or enzymatic

oxidation happened during the pre- or postharvest stages (Kalac, 2019; Lin et al., 2019). From the consumer point of view, the quality of mushrooms is determined by colour, texture, cleanliness, and sensory parameters. Colour changes caused by perturbation or enzymatic browning can be detected by the consumers directly and critically affect the commercial values of mushrooms in the market (Lin et al., 2019). Therefore, monitoring mushroom quality become critical for the manufacturers and dealers.

Hyperspectral imaging sensor could be used as an alternative for time consuming and costly destructive sampling. A review of the literature shows that hyperspectral imaging can be used as an effective tool for fruit, vegetables and mushrooms in assessing quality parameters both on defined variables (e.g., colour, firmness, acidity, sugar) and complex properties (e.g., maturity, ripeness, detection of biotic defects, physiological disorders, mechanical damages, sensory quality). A literature search in the title, abstract, and keywords sections of journal articles was performed in the Web of Science Core Collection using the term “hyperspectral imaging” and its various variations such as “spectral imaging”, “chemical imaging”, and “imaging spectroscopy” as well as the term “fruit”, “vegetables” or “mushrooms”. Results show that 1767 documents are indexed based on the word combinations in the years 2000–2020 (Science, 2021). The publication of documents in this area increased in recent years (Fig. 1). It is expected that the use of hyperspectral imaging technology in the horticultural industry will be considered even more in the coming years. Actually a high percentage of 5.2% (93 articles) of the papers in the first search, on the use of hyperspectral imaging for evaluating the fruit, vegetables or mushroom quality published between 2000 and 2020, are review papers. Most review papers, however, limit the focus on fruit quality parameters (Li et al., 2015; Lu et al., 2020; Pathmanaban et al., 2019; Wang et al., 2016) or more general on agricultural applications (Benelli et al., 2020). The present review aims to close the few gaps, examining the application of hyperspectral imaging in predicting quality-related indices, as well as the detection of defects, namely, physiological disorders, mechanical damages, and fungal infections.

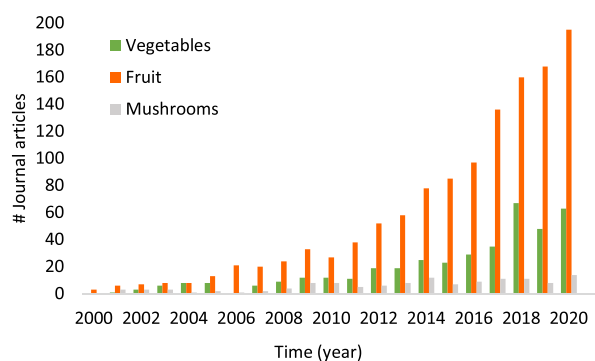


Fig. 1 – Number of published journal articles between 2000 and 2020 using the term “hyperspectral imaging” and its various variations such as “spectral imaging”, “chemical imaging”, and “imaging spectroscopy”, as well as the term “fruit”, “vegetable” or “mushroom” (Science, 2021).

In addition, this review also focuses on the added value of the rapidly developing capabilities of Artificial Intelligence (AI) in processing hyperspectral imaging for quality assessment of vegetables, fruit and mushrooms, as well as the associated challenges. Due to the recent technological evolution, AI made tremendous progress in both research and the business world and its techniques are already integrated in various applications (e.g., voice assistance, smart cars, facial recognition, search engines, etc.). AI techniques also caught the attention of the horticultural, agricultural and food domain, introducing new possibilities for the development of smart systems, especially since the emergence of Machine Learning (ML). ML is a subbranch of AI, which deals with algorithms that attempt to learn to recognise patterns in data, and use these to make decisions (Nturambirwe & Opara, 2020). As vast amounts of data are produced in hyperspectral imaging systems, the analysis and interpretation of these data pose many challenges and remain a bottle neck in many horticultural and agricultural applications (Signoroni et al., 2019; Weersink et al., 2018). The majority of current applications in these sectors use RGB images (Dhiman et al., 2022), as the computing power needed to process such databases is substantially lower than the one needed for hyperspectral imaging systems. Furthermore, they are generally easier and cheaper to obtain. However, as computing power becomes cheaper and more accessible, the use of AI and ML in research of hyperspectral imaging systems begins to increase. Behind this success lies, on the one hand, the advantage of hyperspectral imaging over RGB cameras, i.e. the large amount of information captured in a single image, which in many cases is not visible to the human eye. AI techniques applied to hyperspectral imaging data can be used to reveal correlations with the aforementioned quality parameters, as it can exploit the richness of the spectral and spatial information. On the other hand, nowadays, there is also the rise of a relatively new and promising subbranch of ML, i.e., Deep Learning (DL). DL gained a lot of importance as it has improved the efficiency and quality of diverse applications in comparison with traditional machine learning models (Jaiswal et al., 2021). For the quality control of fresh produce, the combination of hyperspectral data with innovative techniques within AI therefore opens up many possibilities.

An analogous literature search in the Web of Science database using the term “hyperspectral imaging” and its variations in combination with the term “artificial intelligence”, or its subdomains “machine learning” and “deep learning”, and “fruit”, “vegetables” or “mushrooms”, yields a much lower number of results. Respectively 28 documents in 2021, 21 documents in 2020 and in the years before only 0 to 7 results, with a first search result of 2012 (Science, 2022). Even though the explicit focus on the application of AI for quality assessment of fruit, vegetables and mushrooms with hyperspectral imaging received limited attention in the past, the recent increase in the number of studies illustrates that this topic is expected to become more important in the near future. Overall, both searches show that the number of works about fruit has a clear majority compared to vegetables or mushrooms.

Moreover, a similar trend to the first literature search is observed here, namely the presence of various review papers.

However, these review papers have a different focus within the contexts of AI for quality assessment. In general, the focus on hyperspectral imaging is more narrow, as the presence of RGB analysis with AI is more prevalent (Dhiman et al., 2022; Meenu et al., 2021; Naranjo-Torres et al., 2020; Nturambirwe & Opara, 2020). In addition, some review papers have a broader context, giving a general overview of AI applications in food and agricultural products (Fracarolli et al., 2021; Megeto et al., 2021), or on multiple hyperspectral imaging applications (Jaiswal et al., 2021). As above, many reviews also focus mainly on fruit aspects, both quality (Dhiman et al., 2022; Jaiswal et al., 2021) and, for example, fruit detection (Koirala et al., 2019; Naranjo-Torres et al., 2020). In this review, the focus is on the combination of the potential of hyperspectral imaging in context of quality assessment in fruit, vegetables and mushrooms, with the possibilities and applications of AI in this context.

The remainder of this review paper is therefore organised in four sections. Section 2 introduces hyperspectral imaging, both from a theoretical and application point of view. This section also provides an overview of the most common wavelengths in the reviewed literature, as well as a short summary of the models typically used. Section 3 provides a thorough overview of the state-of-the-art parameters related to quality assessment of fruit, vegetables and mushrooms, ranging from qualitative parameters to defect detection. To illustrate the potential of hyperspectral imaging analysis with AI, Section 4 presents a high-level description of the recent developments in AI, focussing on the machine learning branch and its subbranch deep learning. In this section, the models that were already mentioned in previous sections, are also discussed in more detail in combination with associated tasks and applications in agriculture, horticulture and food. Finally, based on the lessons learned throughout the paper, Section 5 pinpoints some challenges and future considerations to improve the analysis and applicability of hyperspectral imaging for quality assessment of fruit, vegetables and mushrooms.

2. Hyperspectral imaging

2.1. Theory

Hyperspectral imaging was originally developed for remote sensing applications around 1985 (Goetz et al., 1985). During the last two decades, it was gradually employed in diverse research areas of agriculture and foods to acquire complementary information on both images and chemical compositions of research objects (Gowen, O'Donnell, Taghizadeh, Cullen, et al., 2008). Hyperspectral imaging is an advanced technique that integrates conventional imaging and spectroscopy (e.g., infrared, Raman, fluorescence, etc.) of an object to attain a data form of a three-dimensional (3-D) hypercube including the two-dimensional (2-D) digital images ($X \times Y$) attached with the third dimensional spectral information (λ) (Amigo et al., 2013). Therefore, hyperspectral images are constructed of a deck of spatial image planes aligned along the spectral wavelength; each image plane assigned to each single wavelength band visually depicts information of the pixelwise

product at each wavelength. Consequently, each pixel in an image contains a spectrum representing the light absorbing, reflecting, emitting and/or scattering properties of the spatial region represented by that pixel, although each pixel spectrum may be influenced by its neighboring pixels due to various optical, instrumental, and background effects. The resulting spectrum acts like a fingerprint, which can be used to estimate the composition of that pixel.

Hyperspectral imaging technology can engage with specular reflected, transmitted, emitted, and diffusely scattered light. Hyperspectral imaging technology is commonly limited to wavelength ranges of 400–2500 nm. Although other wavelength ranges could also be imaged, this review was limited to this wavelength range. Hyperspectral images can be obtained by using cameras with charge coupled device or complementary metal oxide semiconductor sensors for the visible to near infrared (VIS-NIR) wavelength range of 400–1000 nm, and by using Indium Gallium Arsenide (InGaAs) detectors for the longer wavelength range of 950–2500 nm. Commonly available illumination systems used for hyperspectral image acquisition are halogen lamps, light-emitting diodes (LEDs), and laser. In addition, there are three typical ways of scanning configuration including whisker-broom (i.e., point scan), push-broom (i.e., line scan), and staring imager configuration (i.e., plane scan) (Amigo & Grassi, 2020). In order to obtain reliable images, spatial calibration and spectral reflectance calibration are critical in the image acquisition process. Spatial calibration is required to fix the angles and the distances of illumination systems to the objects, to adjust the frame period and integration time, and to decide on the frame size and resolution of image acquisition. The spectral reflectance calibration accounts for the background spectral response of the instrument and the “dark” camera response. Thus, the corrected reflectance values of measured signals are calculated on pixel-by-pixel basis (Amigo et al., 2013).

After data acquisition, pre-processing of hyperspectral imaging data aims to carry out spatial and spectral operations to remove irrelevant information and noise from the images and the spectra collected, and to reduce the dimensionality of data for further analysis. As shown in Fig. 2, the spatial operations include the removal of dead pixels and spikes by image filtering methods, selection of the region of interest (ROI), creating masks for background information removal using thresholds and classifiers (e.g., principal component analysis (PCA), K-means, etc.), and reducing image sizes by performing data binning or interpolation.

Then the spatially corrected data cube is unfolded into a 2-D data frame and spectral operations involve spectral data pre-processing algorithms, including baseline correction, derivative conversion, normalisation, centering, etc., to enhance the main spectral features. As the collected spectral information in the full spectral range may be redundant to relate to specific chemical imaging, spectral compression methods including PCA, Successive Projections Algorithm (SPA), Competitive Adaptive Reweighted Sampling (CARS), Stepwise Linear Regression Selection (SLRS), etc., can be used for further hyperspectral data compression. Furthermore, spectral variable selection algorithms, such as Variable Importance in Projection (VIP), SLRS, Ensemble Monte Carlo Variable Selection (EMCVS), CARS, etc., can be employed with Partial

Least Squares Regression (PLSR), Principal Component Regression (PCR), Multiple Linear Regression (MLR), Partial Least Squares Discriminant Analysis (PLS-DA), etc., to select the most relevant spectral variables to maintain a smaller and even more condensed size of the variable group. On the other hand, the enhanced spectral variables can be deployed in the development of multispectral imaging techniques to speed up the data acquisition processes in real applications. A more condensed size allows for a faster computational progress in the next step of either supervised and unsupervised classification and/or regression modelling by for example PCA, K-means clustering, PLS-DA, Support Vector Machine Discriminant Analysis (SVM-DA), Linear Discriminant Analysis (LDA), Artificial Neural Network (ANN), etc.; or quantification of target compounds by quantisation and resolution analysis strategies including PLSR, Support Vector Machine Regression (SVMR), ANN for regression, MLR, Multivariate Curve Resolution (MCR), Classical Least Squares (CLS), etc. (Amigo et al., 2013). Consequently, the developed models or algorithms are applied on each acquired hypercube to generate chemical distribution maps. Additionally, domain statistical methods are performed on the distribution maps to demonstrate the model performances.

2.2. Sensor technology, wavelength selection and analysis

Table 1 provides an overview on the hyperspectral research carried out in the fields of fruit, vegetables and mushrooms within the last decade. Studies and research adjacent to hyperspectral imaging were omitted, e.g., the detection of pathogens with Raman spectroscopy for onions (Gan et al., 2017), downy mildew detection through chlorophyll fluorescence for lettuce (Bauriegel et al., 2014), detecting internal rot of onions through transmittance by avoiding wavelengths with strong absorption peaks of chlorophyll and water (Sun et al., 2018).

The sensors used in these studies were diverse and ranging from commercial hyperspectral sensors, such as Cubert, Imec, HySpex, Bayspec, Specim and Headwall, towards more experimental setups. A recent review of commercial hyperspectral sensors and prices was carried out by Appeltans et al. (2020). Within this review, the most abundantly used wavelength range was 900–1050 nm, over 90% of journal articles used this region of the electromagnetic spectrum. In addition, the VIS-NIR region from 450 to 850 nm was also used in 68–88% of studies. Sensors capturing shorter (<400 nm) and longer wavelengths (between 1100 and 2500 nm) were only used in less than 30% of hyperspectral studies in the fields of fruit, vegetables and mushrooms. The use of wavelength ranges is similar in both fruit, vegetables and mushrooms research and thus appear related to the availability of sensor technology and relevant quality parameters. Lower availability and higher cost of other sensors (Appeltans et al., 2020) are the most probable cause of this distribution. This is especially true for hyperspectral imaging sensors used in field conditions, where lower signal to noise ratio of SWIR sensors limits the potential in field research.

Papers highlighted in Table 1 were screened for results on spectral variable selection to point out meaningful

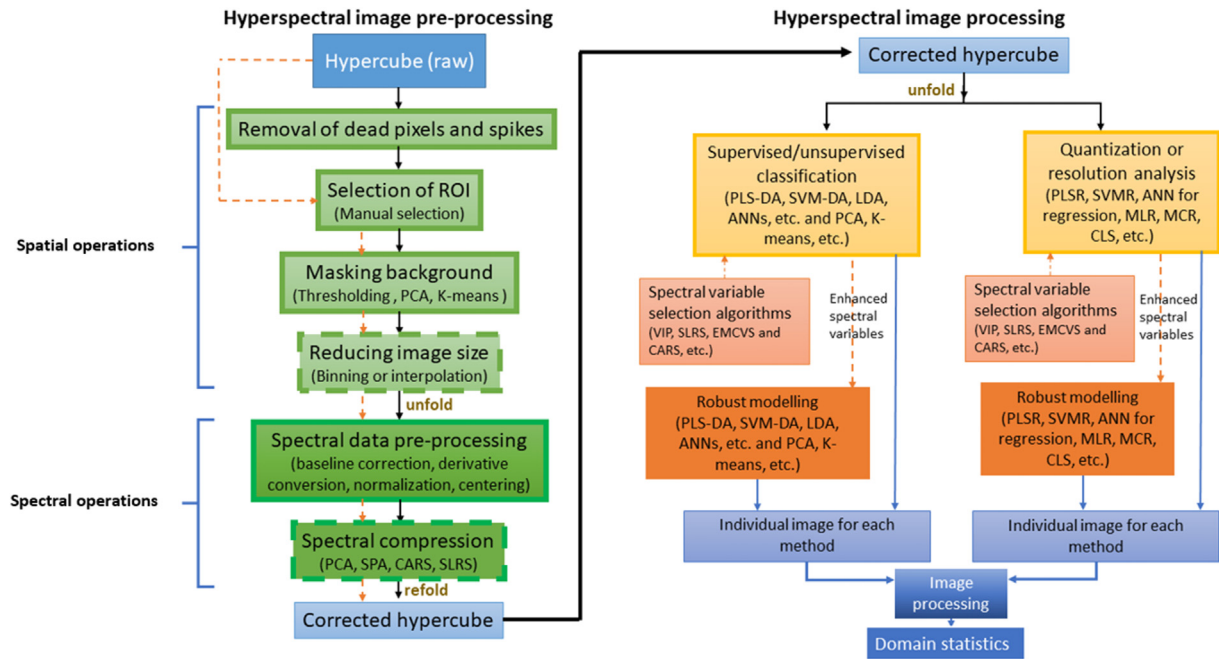


Fig. 2 – Flowchart of hyperspectral image pre-processing and processing.

wavelengths. The number of journal articles that mention relevant wavelength in a specific region were depicted in Fig. 3. The most abundantly used wavelengths regions were 601–850 nm, which is mentioned in over 50% of all studies. For quality parameters, 651–700 nm was the most studied region with over 13 mentions within separate studies. As shown in Fig. 3, for all other wavelengths regions, more research was carried out on defect detection compared to analysis of quality parameters.

The number of journal articles related to hyperspectral imaging research for fruit was more abundant in most of the wavelength regions. However, for vegetables the regions between 300–450, 501–550, 901–950 and 1001–2500 nm were mentioned more frequently. In addition, the wavelength range between 451 and 500 nm was shown relevant in studies related to quality of mushrooms: half of all studies specifying this range were studies on mushrooms ($n = 4$).

Based on the studies presented in Table 1, 118 prediction models for one parameter were published in the fields of fruit, vegetables and mushrooms research within the last decade. The most prevalent prediction algorithm in the reviewed studies is PLSR, used 41 times with an average performance indicator coefficient of determination in the prediction (R_p^2) of 0.78 ($n = 35$). Other regression models are used less, namely PCR in 14 models with an average R_p^2 of 0.83 ($n = 11$) and MLR in 7 models with an average R_p^2 of 0.81 ($n = 7$). PLS-DA is the most used classification prediction model, with 11 occasions and an average overall classification accuracy (OCA) of 89.4% ($n = 10$), followed by LDA on 4 occasions with an average OCA of 86.8% ($n = 4$). Prediction models used for both regression and classification were SVM, ANN and Convolutional Neural Network (CNN). SVM was used in 24 models and had an average OCA of 87% ($n = 10$) and R_p^2 of 0.82 ($n = 13$), ANN was used for 9

prediction models and had an average OCA of 95.3% ($n = 6$) and R_p^2 of 0.81 ($n = 3$) and CNN for 11 prediction models with an average OCA of 87.3% ($n = 7$) and R_p^2 of 0.81 ($n = 4$). Overall, SVM and PLSR were the most used algorithms for predicting fruit, vegetable and mushroom quality from hyperspectral information in the last decade. Section 4 takes a closer look at the capabilities of these models, associated tasks and the potential evolution to newer models.

3. Quality assessment and parameters

As mentioned in Section 2, parameters for the quality assessment of fruit, vegetables and mushrooms were divided into two categories, namely (i) qualitative properties (including ripeness and maturity), and (ii) parameters related to defects.

Maturity is an important term in the quality assessment and has two different definitions. On the one hand, physiological maturity is achieved when the fruit is able to ripen after harvest. On the other hand, horticultural maturity (or commercial maturity) means that the fruit has ripened to the point that it can be sold in terms of appearance and edibility. Maturity in the fruit has its own complexities which led to the development of a large number of physicochemical indices for its diagnosis. In the following, physiological maturity and horticultural maturity will both be assessed through quality variables.

The most relevant quality-related parameters are pigments, firmness and elastic modulus, moisture content and dry matter, soluble solids content (SSC), and acidity. These parameters all influence shelf life and resistance to decay and mechanical damage. Defects in fruit, vegetables and mushrooms, especially those obtained during the storage process,

Table 1 – Crops vs quality parameters, ripeness parameters and defects, limited to studies published in the last 10 years.

Crop	Industrial parameter	Prediction model ^a	Performance	References
<i>Quality/ripeness</i>				
Apple	Firmness	PLSR	$R_p^2 = 0.87$	Te Ma et al. (2021b)
Apple	Moisture content	LS-SVM	$R_p^2 = 0.96$; $RMSE_p = 0.45$	Dong and Guo (2015)
Avocado	Dry matter	SVM	$R_p^2 = 0.90$; $RMSE_p = 2.6$ g/kg	Diaz et al. (2021)
Banana	Chlorophyll	PCR	$R_p^2 = 0.79$; $RMSE_p = 5.98 \times 10^{-4}\%$	Saputro et al. (2018)
Banana	TSS, moisture and firmness	MLR	$R_p^2 = 0.85, 0.87$ and 0.91	Rajkumar et al. (2012)
Banana	Carotenoids	PLSR	$R_p^2 = 0.94$; $RMSE_p = 0.79$	Permata et al. (2017)
Grape	Anthocyanin	PLSR	$R_p^2 = 0.86$	Diago et al. (2016)
Grape	SSC	ϵ -SVMs	$R_p^2 = 0.92$	Gutierrez et al. (2019)
Grape	Titrateable acidity	PLSR	$R_p^2 = 0.82$ – 0.95	Baiano et al. (2012)
Honey peach	Chlorophyll	PLS-DA	$R_p^2 = 0.74$	Sun, Wang, et al. (2017)
Kiwifruit	Acidity	PLSR	$R_p^2 = 0.64$; $RMSE_c = 0.14$	Ma et al. (2021a)
Mango	Titrateable acidity	MLR	$R_p^2 = 0.81$; $RMSE_p = 0.24\%$	Rungpichayapichet et al. (2017)
Mango	Dry matter content	CNN, PLS-DA	$F_1 = 0.97$	Wendel et al. (2018)
Orange	SSC	ANN	$R_p^2 = 0.55$ $RMSE_p = 0.86\%$	Riccioli et al. (2021)
Pear	SSC, firmness	SAE-FCNN	$R_p^2 = 0.89, 0.92$; $RMSE_p = 1.81$ N, 0.22%	Yu et al. (2018)
Persimmon	SSC, firmness	PLS	$R_p^2 = 0.76, 0.88$; $RMSE_p = 1.40$ °Brix	Wei et al. (2020)
Plum	Firmness	MLR	$R_p^2 = 0.69$; $RMSE_p = 0.63$ kg/cm ²	Meng et al. (2021)
Strawberry	Anthocyanin	PLSR	$R_p^2 = 0.63$ – 0.87	Cho et al. (2021)
Strawberry	Ripeness	CNN AlexNet	OCA = 98.6%	Gao et al. (2020)
Asparagus	Shelf life	PLS-DA	OCA = 82–91%	Sánchez et al. (2013)
Broccoli	Total glucosinolates	PLSR	SECV = 1.75 μ mol/g DM	Hernández-Hierro et al. (2014)
Lettuce	Chlorophyll/anthocyanin content	–	NR	Simko et al. (2016)
Pepper	TSS, dry matter, osmotic potential, ascorbic acid, total chlorophyll, carotenoids content and firmness	PLSR, PCR, SVM, Kernel	$R_p^2 = 0.93, 0.92, 0.9, 0.77, 0.91, 0.92$ and 0.65	Ignat et al. (2014)
Pepper	TSS, total chlorophyll, carotenoid and ascorbic acid content	PLSR	$R_p^2 = 0.79, 0.92, 0.93$ and 0.61 ; $RMSE_p = 1.1$ Brix %	Schmilovitch et al. (2014)
Spinach	N, water content	PLS	$R_p^2 = 0.41, 0.87$; $RMSE_p = 0.26\%, 0.85\%$	Corti et al. (2017)
Tomato	Elasticity	ANN	$R_p^2 = 0.90$; $RMSE_p = 0.11$	Mollazade et al. (2015)
Tomato	Chlorophyll, carotene	PLS	$R_p^2 = 0.73, 0.82$; $RMSE_p = 0.30, 0.25$	Polder et al. (2004)
Tomato	Colour (L^* , a^* , b^* , H, Ch)	PLSR, PCR	$R_p^2 = 0.86, 0.93, 0.42, 0.95$ and 0.51	van Roy et al. (2017)
Tomato	Ripeness (6 and 3 classes)	PLS-DA	OCA = 80.2–88.4% and 84.4–92.1%	Zhu et al. (2015)
Tomato	SSC, pH	PLS	$R_p^2 = 0.38$ and 0.59 ; $RMSE_p = 0.5$ and 0.17	Huang, Lu, Hu, et al. (2018)
Mushroom	Bolete sp. identification	CNN ResNet	OCA = 99.76%	Dong et al. (2021)
<i>Defects</i>				
Apple	Bruising	PLS-DA	OCA > 90%	Luo et al. (2012)
Blueberry	Mechanical damage	CNN ResNet/ResNeXt	OCA = 88.4, 87.8%; $F_1 = 0.90, 0.89$	Wang et al. (2018)
Date	Microbial infection	LDA, QDA	OCA = 91–100%	Teena et al. (2014)
Jujube	Bruising	CARS-PLS-DA	OCA = 91.1%	Yuan et al. (2021)
Jujube	Chilling injury	LDA	OCA = 98.3%	Lu et al. (2018)
Kiwifruit	Bruising	PC	OCA = 85.5%	Lu and Tang (2012)

Loquat	Bruising, browning	XGBoost	OCA = 95.9%	Munera et al. (2021)
Mango	Mechanical damage	kNN	OCA = 98%	Rivera et al. (2014)
Olive	Verticillium wilt	LDA/SVM	OCA = 59%–79.2%	Calderon et al. (2015)
Orange	Scarring	IP-T	OCA = 91.5%	Li et al. (2011)
Orange	Fungal diseases	PCIP	OCA = 98.6%	Li et al. (2016)
Honey peach	Chilling injury	ANN	OCA = 96.87%	Sun, Gu, et al. (2017)
Honey peach	Fungal diseases	PLSR	OCA = 98.7%	Sun, Wang, et al. (2017)
Peach	Bruising	MNF	OCA = 87.5%	Zhang et al. (2015)
Peach	Bruise, Sound	Watershed	OCA = 96.5% (bruise); 97.5 (sound)	Li et al. (2018)
Capsicum	Tomato Spotted Wilt virus	SVM	OCA = 93.6% (VNIR), 91.5% (SWIR)	Moghadam et al. (2017)
Cucumber	Chilling injury	SVM	OCA = 90.5%	Cen et al. (2016)
Cucumber	Bruising	–	–	Lu et al. (2011)
Cucumber	Subsurface yellow spots	SVM	OCA = 98%	Lu and Lu (2019)
Cucumber	Disease	ECR (SVM, LDA)	OCA > 94.7%	Li et al. (2020)
Cucumber	Defects	CNN-SSAE	OCA = 88.3–91.1%	Liu et al. (2018)
Lettuce	Biological contamination	VI	OCA > 91%	Mo et al. (2017)
Onion	Sour skin, neck rot	–	–	Wang et al. (2014)
Spinach	E. coli	ANN	$R_p^2 = 0.97$	Siripatrawan et al. (2011)
Squash	Powdery mildew	RBF	OCA = 89% early disease stage OCA = 96% late disease stage	Abdulridha, Ampatzidis, Roberts, et al. (2020)
Tomato	Puncture force, slope and flesh firmness	PLS	$R_p^2 = 0.85, 0.90$ and 0.74 ; $RMSE_p = 1.59$ N, 0.52 N/mm and 1.00 N	Huang, Lu, Xu, et al. (2018)
Tomato	Fusarium, root crown disease	SVM	OCA = 74–90%	Abu-Khalaf (2015)
Tomato	Bacterial Spot	ANN	OCA = 98%	Abdulridha, Ampatzidis, Kakarla, et al. (2020)
Tomato	Target Spot	ANN	OCA > 96%	Abdulridha, Ampatzidis, Kakarla, et al. (2020)
Mushroom	Brown blotch	PLS-DA	OCA > 95%	Gaston et al. (2010b)
Mushroom	Enzymatic browning, casing soil (Colour Lab)	PLS-DA	OCA = 98–100%	Taghizadeh et al. (2011b)
Mushroom	Colour (Lab)	PLSR	$RMSE_c = 1.5$	Taghizadeh et al. (2011a)
Mushroom	Bruising (moisture content)	PLS-DA	OCA = 100%	Esquerre et al. (2012)
Mushroom	Cobweb disease	SVM	OCA = 77.27–100%	Parrag et al. (2014)
Mushroom	Mechanical damage (Colour)	SVM, ANN	OCA = 92.48–94.19%	Rojas-Moraleda et al. (2016)
Mushroom	Browning (Colour)	PLS-DA	OCA = 69.4–94.5%	Mollazade (2017)

^a ANN (Artificial neural network), CARS (Competitive adaptive reweighted sampling), CNN (Convolutional neural network), EAU (Enzyme activity unit), ECR (Extended collaborative representation), EMCVS (Ensemble Monte Carlo variable selection), Faster R-CNN (Faster region based convolutional neural network), FCNN (Fully connected network), GMM (Gaussian mixture model), IP-T (Image processing-thresholding), kNN (k-Nearest neighbors), LDA (Linear discriminant analysis), LS-SVM (Least squares-support vector machine), MFA (Multifactorial analysis), MC-UVE (Monte Carlo-uninformative variable elimination), MLR (Multiple linear regression), MNF (Minimum noise fraction), mRMR (Minimum redundancy maximum relevance), NR (not reported), OCA (Overall classification accuracy), PC (Parallelepiped classification), PCA (Principal component analysis), PCIP (Pseudo-color image processing method), PCR (Principal component regression), PLS (Partial least squares), PLS-DA (Partial least squares discriminant analysis), PLSR (Partial least squares regression), QDA (Quadratic discriminant analysis), R_c^2 (Coefficient of determination in cross-validation), R_p^2 (Coefficient of determination in prediction), RBF (Radial basis function), RF (Random forest), RFr (Random frog), $RMSE_c$ (Root mean square in calibration), $RMSE_p$ (Root mean square in prediction), ROC (Receiver operating characteristic), SAE (Sparse autoencoder), SFS (Sequential forward selection), SPA (Successive projections algorithm), SSAE (Stacked sparse autoencoder), SSC (Soluble solid content), SVM (Support vector machine), SLRS (Stepwise linear regression selection), TSS (Total soluble solids), Tukey's HSD (honestly significant difference), VI (Vegetation Indices), XGBoost (Extreme gradient boost), ϵ -SVMs (Epsilon-support vector machines).

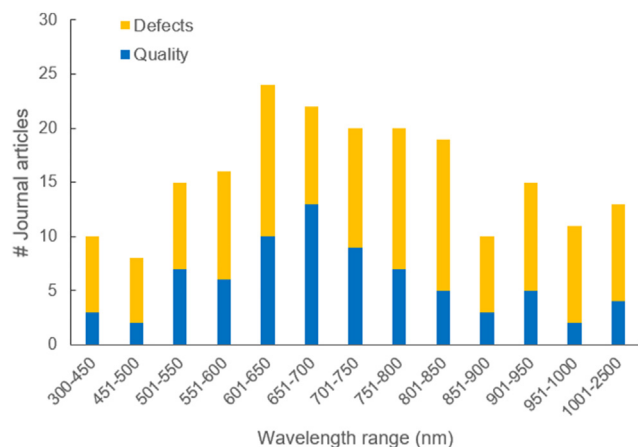


Fig. 3 – Number of published journal articles contained within this review (Table 1) that highlight relevant wavelengths within the specified wavelength range, to estimate quality parameters or defects.

can be the result of one or more different parameters, including physiological disorders, mechanical damage, and fungal infections. Most of the horticultural applications of hyperspectral imaging on quality parameters were carried out for fruit and edible vegetable fruits and done postharvest, a review by Lu et al. (2020) provides good examples.

3.1. Pigments

Chlorophyll, which is responsible for photosynthesis, is a green pigment found in most fruits. This pigment absorbs the main part of blue and red light and reflects the green light through the electromagnetic spectrum. Most fruit change colour as part of the ripening process. Colour changes in the fruit are due to the enhanced degradation rate of chlorophyll and sometimes production of carotenoids or flavonoids (Walsh et al., 2020) (Fig. 4). Carotenoids are the second common group of pigments in nature and cause yellow to red colour. Flavonoids are water-soluble pigments mostly located in the vacuole, that are responsible for a red or purple colour.

Saputro et al. (2018) developed an optical system based on hyperspectral imaging to detect the ripening of banana by predicting the amount of chlorophyll pigment in the peel of fruit. After wavelength selection with PCA, PCR and SVMs were used to predict the amount of chlorophylls and determine the ripeness of banana at three levels (immature, mature and very mature). In this study, very good results were obtained in predicting chlorophyll content ($R_p^2 = 0.79$ and Root Mean Squared Error of the prediction ($RMSE_p$) = $5.98 \times 10^{-4}\%$) and determining the ripening level. Permata et al. (2017) reported similar results in the same wavelength range in predicting the ripening level of banana through the estimation of carotenoids pigment ($R_p^2 = 0.94$ and $RMSE_p = 0.79\%$). In another study, hyperspectral imaging the wavelength range of 400–1700 nm was used to predict anthocyanins, at different stages of maturity of strawberry (Cho et al., 2021). The highest accuracy (86%) was obtained using spectral information in

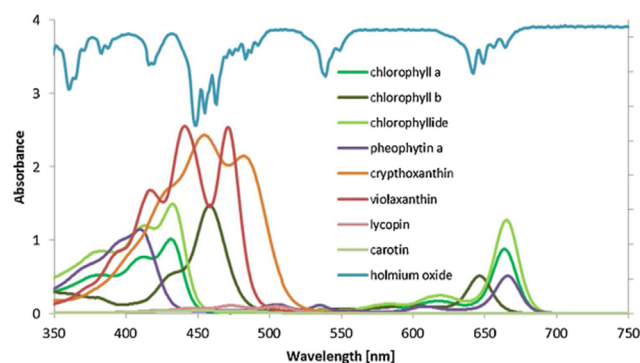


Fig. 4 – Pigment standards: absorption spectra of pigments in acetone (Walsh et al., 2020).

order to create spatial distribution maps of this pigment on the surface of strawberry.

Similar to fruit, colour is an as important qualitative parameter for vegetables and pigments play a key role in quality determination. As a result, the ability to determine and quantify pigments is a popular research topic. Simko et al. (2016) successfully developed a non-destructive method using hyperspectral imaging to determine chlorophyll and anthocyanin content for lettuce (R_p^2 values of 0.81 and 0.74, respectively). Moreover, Simko et al. (2016) used sensor fusion with linear and non-linear regression methods to determine total chlorophyll and carotenoid content. Their research aimed to predict the maturity stages of bell peppers utilising optical absorbance at 470, 649 and 664 nm. Results regarding total chlorophyll content were moderate achieving an R_p^2 of 0.44, however, for carotenoids, the results were more than promising with an R_p^2 of 0.87. The same wavelengths for determining chlorophyll and carotenoid content in bell pepper were also used by Schmilovitch et al. (2014) in combination with PLSR with results being only slightly better for carotenoid content, R_p^2 of 0.93 compared to 0.92 for chlorophyll. Polder et al. (2004) also investigated pigments in tomatoes by measuring chlorophyll and carotenes in the range of 400–700 nm in combination with PLSR, achieving an R_p^2 of 0.73 and 0.82, respectively.

In addition to the estimation of pigments, research on hyperspectral imaging focussing exclusively on colour were conducted as well. van Roy et al. (2017) successfully used hyperspectral imaging to measure the colour of vine tomatoes ($R_p^2 = 0.9$). Moreover, Huang et al. (2014) used hyperspectral imaging towards predicting edamame beans colour during the drying process achieving an R_p^2 of 0.94 while using images acquired between 400 and 1000 nm.

Due to the lack of protection cuticles on the surface, mushrooms are very perishable compared with fruit and other vegetables. Colour change on the surface of white button mushrooms from bright white to brown manifests the trend of mushroom quality deterioration. Browning effects on mushrooms are mainly caused by a complex series of enzymatic reactions of polyphenol oxidase (PPO) and peroxidase (POD) enzymes. Other important factors influencing colour

change include pre-harvest conditions, postharvest treatments and storage conditions, mechanical damages during handling and microbial/fungal infections (Lin et al., 2019). Hunter L-value measurement and conventional RGB imaging were widely applied, non-destructive methods for monitoring colour change on mushroom surface until hyperspectral imaging techniques within the VIS-NIR were implemented in mushroom studies by Gowen, O'Donnell, Taghizadeh, Cullen, et al. (2008). PLSR was applied to successfully predict L-values of mushrooms with an RMSE_p of 1.36 by Taghizadeh et al. (2011b). Furthermore, MLR and PCR models were shown to predict the measured L*- and b*-values of sliced white button mushrooms under various storage conditions ($R_p^2 = 0.95$ and 0.75 ; RMSE_p = 0.47 and 0.66) (Gowen, O'Donnell, Taghizadeh, Gaston, et al., 2008). In addition, different studies investigated the detection of discolouration caused by the PPO enzymatic reaction using hyperspectral imaging (Gaston et al., 2010a; Mollazade, 2017). The brown colour patterns were detected using both PLS-DA (Mollazade, 2017) and PCR models (Gaston et al., 2010a). However, the colour references of mushrooms used for the model development were obtained using a colorimeter to measure the selected areas of mushroom cap surface, which is limited in a small measurement area on the mushroom surface; thus bias of colour prediction is a common issue using the developed models.

3.2. Firmness and elasticity

Tissue changes occur naturally during fruit growth on the tree, while after harvest and during storage, these changes continue. Fruit tissue changes occur due to chemical changes in the middle septum and primary cell wall components, such as pectins, cellulose, and hemicellulose, which accelerate the softening of the flesh. Firmness and elasticity are two important indices to assess changes in fruit tissue (Mehl et al., 2002).

Firmness and elasticity are quality parameters frequently approached with spectroscopy in the visible and short wave infrared wavelength range, but there is common understanding that mechanical properties cannot be measured with point spectroscopy (Walsh et al., 2020). However, other imaging technology have been employed with more success. Studies show that firmness can be measured by means of time-resolved and spatially-resolved approaches. Recently, a method called multifiber-based spatially resolved spectra collection was developed to measure the firmness of the fruit (Ma et al., 2021b). In this method, a number of optical fibres, located at certain distances from the point of entry of light into the samples, are used to transfer the returned photons from inside the fruit to the hyperspectral imaging system. The acquired spectra in the wavelength range of 600–1100 nm were analysed using the CARS and PLSR combination. The results showed that the region 750–960 nm predicted firmness with high accuracy ($R_p^2 = 0.87$).

Compared to fruit, elasticity and firmness of (leafy) vegetables and mushrooms have not drawn so much interest, although researchers tried to develop solutions in order to predict them. Mollazade et al. (2015) used multispectral imaging, ANN and images captured at 660 nm to estimate tomato elasticity with R_p^2 values of 0.90. Furthermore, Huang, Lu, Xu,

et al. (2018) predicted tomato firmness with satisfactory results ($R_p^2 = 0.74$) using spectra in the range of 550–1650 nm. Thus, both publications found around elasticity and firmness in vegetables, deal with tomatoes, which are anatomically challenging considering the light passes through the tissue.

3.3. Moisture content and dry matter

Fresh fruit contain 80–95% water, depending on the type of product. During transpiration, water vapour is transferred from the intercellular space of fruit tissue to the surrounding air. Decreasing moisture leads to wilting and lack of brittleness of fruit with adverse effects on the appearance, tissue, taste, and volume of the product. Decreasing moisture content is also a primary cause of postharvest damages such as physiological disorders (Kader, 2002).

In some fruits, such as avocados, it is necessary to use internal quality indicators to determine the degree of ripeness (Clark et al., 2007). Studies show that the dry matter content (DMC) of avocados can be measured with acceptable accuracy using hyperspectral images acquired in the wavelength range of 400–1000 nm (Girod et al., 2008). Because the first to third overtones (vibration absorption bands) of the water molecule are in the wavelength range of 700–1200 nm, it should be possible to predict the moisture content in fruit tissue in this wavelength range (Siregar et al., 2017). Also, since a strong absorption peak of the water molecule related with the O–H stretching first overtone occurs at the wavelength of 1450 nm, moisture content was predicted with high accuracy ($R_p^2 = 0.96$ and RMSE_p = 0.45) in this region for kiwifruit and apple (Dong & Guo, 2015; Liu & Guo, 2015).

As for vegetables, only for spinach, which has a moisture content between 90 and 99%, research on hyperspectral imaging for the determination of moisture content and dry matter was found. Corti et al. (2017) estimated moisture content using PLS modelling and hyperspectral imaging in the range of 400–1000 nm.

Fresh mushrooms normally contain over 90% of moisture, and are therefore very sensitive to environmental temperature and humidity. Mushrooms stored under high temperature and low humidity will result in moisture loss, and consequently discolouration and deformation of mushroom surfaces. Both the spectral wavelength ranges of 400–1000 nm and 950–1700 nm can be used for determining moisture content in mushrooms. Gowen, O'Donnell, Taghizadeh, Gaston, et al. (2008) developed MLR and PCR models using the acquired hyperspectral data over 400–1000 nm for the prediction of moisture content in mushrooms. Lin et al. (2019) successfully predicted moisture contents in whole or sliced white button mushrooms by PLSR models developed using hyperspectral imaging data over the wavelength range of 950–1655 nm. The attained statistic results revealed R_p^2 values of 0.89 and 0.71 for the moisture content prediction in mushroom slices and whole mushrooms, respectively.

3.4. Soluble solids content

In many fruits, starch is accumulating during the growth of fruit, which is then converted to sugar during the ripening

process. The amount of sugar increases as the fruit ripens. Since most SSC in the fruit is sugar and a direct and close relationship exists between both, the percentage of total SSC is therefore measured to determine the total amount of sugars (OECD, 2018).

Measuring SSC level using hyperspectral imaging was successful in indoor conditions for various fruits, such as orange (Zhang et al., 2020), persimmon (Wei et al., 2020), and apple (Tian et al., 2019). Gutierrez et al. (2019) recently used a portable NIR spectrophotometer in the wavelength range of 400–1000 nm to predict the amount of SSC by ϵ -SVM. The results showed the possibility of hyperspectral imaging for real-time SSC predictions in grape berry ($R_p^2 = 0.92$, $RMSEP_p = 1.27$ °Brix). This study opened new possibilities for hyperspectral imaging in the field of precision horticulture to evaluate the maturity and quality of fruit on-the-go.

Regarding vegetables or mushrooms, SSC appears to be of low importance to the scientific community. The only publication found was again about tomatoes. For them, SSC is one of the most important parameters for quality and maturity. However, results of Huang, Lu, Hu, et al. (2018) determining SSC using spectroscopy in the range of 550–1300 nm PLS models were poor ($R_p^2 = 0.38$).

3.5. Acidity

In most fruit, the ratio of sugar to acid determines the taste, and as a result, this ratio is used as an indicator to determine the horticultural maturity of fruit. At the beginning of the ripening process, due to the low sugar content and the high acidity of fruit, the sugar to acid ratio is low and the fruit tastes sour. During ripening, the acidity of fruit decreases and the amount of sugar and consequently the ratio of sugar to acidity increases. The acidity of overripe fruit is very low and therefore the fruit loose taste (OECD, 2018).

Baiano et al. (2012) investigated the possibility of predicting titratable acidity and sensory index of taste in table grape. The results showed a correlation between the actual titratable acidity values and the values obtained from the PLSR model generated by the pre-processed reflectance spectra ($R_p^2 > 0.82$). Recently, Ma et al. (2021a) developed a method to make 3-D (360-degree) spatial distribution maps of the acidity in the fruit from hyperspectral images in the wavelength range of 1002–2300 nm ($R_c^2 = 0.64$ and $RMSE_c = 0.14$).

Similar to SCC and elasticity and firmness, tomatoes were the only vegetables of which researchers assessed the acidity using hyperspectral imaging. Tomato acidity defines the ripening of the fruit while also being of particular importance to the food processing industry (specific pH requirements to avoid the growth of pathogens). Huang, Lu, Hu, et al. (2018) used a spectroscopy system to predict pH in tomatoes with moderate success ($R_p^2 = 0.59$).

3.6. Physiological disorders

Physiological disorders are the result of dysfunction of physiological processes within fruit tissues. It is very difficult to distinguish physiological disorders from each other. Some expand before harvesting the fruit and increase sharply

during cold storage. However, some are caused by unfavourable storage conditions. Low temperature breakdown (chilling and freezing) is the most common type of physiological disorder of horticultural products. Chilling occurs due to storage at low temperatures (10–13 °C). Freezing is caused by the formation of ice crystals in the tissues of products stored at a temperature below its freezing point. Due to low temperature breakdown, cell membranes, including mitochondria and chloroplasts, collapse and, as a result, a series of secondary reactions such as the accumulation of toxins (e.g., ethanol and acetaldehyde), as well as change in cell structure occurs (Kratsch & Wise, 2000).

Hyperspectral imaging can detect low temperature breakdowns in the wavelength range of 400–1000 nm for several fruits, such as jujube (Lu et al., 2018) and peach (Pan et al., 2016). Pan et al. (2016) reported that chill-damaged peaches can be separated from healthy ones with an OCA of 95.8% using ANN. Sun, Gu, et al. (2017) detected chilling injuries in peach fruit by monitoring fruit tissue changes and moisture content. The optimal wavelengths found in the visible area correspond to the absorption of light by the colours yellow, orange, and red. The other optimal wavelengths were 675 nm and 970 nm, which are respectively related to the absorption peak of chlorophyll *a* and the second overtone O–H stretching of water molecule.

For vegetables, one of the most common researched physiological disorders is chilling injury, particularly in cucumbers. Both Moghadam et al. (2017) and Cen et al. (2016) investigated the detection of chilling injury in cucumbers using hyperspectral imaging, with the latter achieving an OCA of 91.6% and higher for two- and three-class classification when applying an SVM classifier. Moreover, Lu and Lu (2019) used SVM and hyperspectral imaging in the range of 700–900 nm to successfully detect cucumbers with subsurface yellowish spots, a physiological disorder whose cause is still unknown (OCA > 98%). Another example of a vegetable crop that suffers physiological disorders during refrigeration is lettuce. Discolouration and browning can occur, and although these disorders do not result in loss of nutrients, nor cause health issues, they make the products unappealing to consumers. Mo et al. (2015) managed to achieve an OCA above 99.9% using hyperspectral imaging to classify lettuce based on discolouration and browning during refrigerator storage.

For mushrooms, Gowen et al. (2009) applied PCA combined with LDA to detect discolouration caused by freeze damage. The LDA model applied to the PCA scores was used to classify pixels into undamaged and damaged classes and achieved an OCA above 95%; highlighting the spectral variability that can arise from the use of pixel spectra in modelling.

3.7. Mechanical damage

Mechanical damage is considered as a type of stress that occurs during harvesting and after harvesting of produce. This stress is associated with physiological and morphological changes if the mechanical force applied is greater than its elastic threshold, the cell wall of fruit tissue is destroyed and tissue cohesion is reduced. As a result, intracellular contents leak out of the cell and enter the intercellular space. Enzymes in secreted intracellular substances, such as POD and PPO

enzymes, cause the breakdown of tissues, and as a result, bruising or browning can occur. Tissues that are mechanically damaged are a key factor in determining the attack of pathogens that cause flesh rot (Martinez-Romero et al., 2004; Miller, 1992).

Several studies reported a higher reflectance in the range of 530–930 nm of both peel and flesh of healthy fruits compared to that of mechanically damaged fruits, e.g., for loquat (Munera et al., 2021), orange (Li et al., 2011), and peach (Zhang et al., 2015). This difference in reflectance is due to changes in concentration of chlorophylls and carotenoids pigments that absorb light at wavelengths of 680 nm and 500 nm, respectively, and indicates a change in colour due to mechanical stress on fruit tissue cells. Conversely, in the wavelength range of 940–970 nm, light absorption in the bruising area decreases. The reason for this is attributed to changes in the structure of the fruit tissue or some degree of dryness in the peel, caused by the rupture of the cell wall – due to mechanical damages – and a resulting drop in moisture and hardening of fruit tissue. Studies also showed that the range of 730–830 nm can be used for early detection of bruising because in the damaged area the concentration of interstitial water increases and water molecules cause greater absorption of light (Luo et al., 2012; Zhang et al., 2015; Zhu & Li, 2019).

For vegetables, Lu et al. (2011) tried to detect bruising on cucumbers. They caused the damage by applying mechanical load and acquired images in the spectral range of 700–1000 nm. Their research suggests that such defects could be detected by analysing enhanced scattering coefficient during optical evaluation.

Bruise damage on white button mushrooms through hyperspectral imaging was performed by Gowen, O'Donnell, Taghizadeh, Cullen, et al. (2008). PCA on spectral data in the 400–1000 nm wavelength range achieved an OCA above 79% on the discoloured patterns. In addition, both PLS-DA model in combination with EMSCV (Esquerre et al., 2012) and ANN/SVM (Rojas-Moraleda et al., 2016) were successful in discriminating mechanical damaged mushrooms.

3.8. Fungal and bacterial infections

Mature fruit are more sensitive to damage during harvest and therefore, to the attack of pathogens that need damaged tissue for penetrating. Fruit may be contaminated before harvest in the field or after harvest during transportation and storage processes. Postharvest diseases that cause fruit loss are diverse, but fungal infections are considered the most common type of disease in ripe fruits due to their low acidity and high moisture content (Valero & Serrano, 2010). Hyperspectral imaging was used very limitedly to detect fungal infections in fruits. Sun, Wang, et al. (2017) found that the amount of chlorophyll decreases significantly when pathogens infect the peach fruit. It is therefore possible to indirectly detect fungal contamination in fruit using spectral data at wavelengths of 617, 675, and 818 nm, which correspond to the absorption of light by chlorophyll pigment. The results showed that unhealthy samples can be separated from healthy ones with 98.8% accuracy using PLSR. Mehl et al. (2002) found significant differences between the reflectance spectra of healthy and infected apples in the visible area of the electromagnetic

spectrum, and also in the area corresponding to the chlorophyll absorption peak (700 nm). These optimal wavelengths were able to distinguish healthy apples from infected apples with an OCA of 76%, 85%, and 95% for Red Delicious, Golden Delicious, and Gala cultivars, respectively. Calderon et al. (2015) used hyperspectral imagery and SVM to classify verticillium wilt-infected olive trees, with promising results: an OCA between 59 and 79%.

On the other hand, hyperspectral imaging found many applications in vegetable fungal infection detection. Wang et al. (2014) used wavelengths between 550 and 1650 to detect *Botrytis aclada*-infected onions and concluded that spectral differences were present. Abdulridha, Ampatzidis, Roberts, et al. (2020) used hyperspectral imaging in the 380–1020 nm range to classify healthy and powdery mildew infected squash with excellent results. The classification of asymptomatic squash in the field was 89% while for late-stage infected vegetables the accuracy reached 96%. Abu-Khalaf (2015) and Abdulridha, Ampatzidis, Kakarla, et al. (2020) both used hyperspectral imaging to classify fungal tomato diseases. Abu-Khalaf (2015) investigated *Fusarium oxysporum* and *Rhizoctonia solani* while Abdulridha, Ampatzidis, Kakarla, et al. (2020) focused on target spot. Both achieved impressive accuracies, i.e. 90% for *Fusarium* and *Rhizoctonia* and 99% for target spot.

Mushrooms can be easily attacked by fungal or bacterial diseases. Coweb disease, caused by *Cladobotryum dendroides* subspecies, is one of the most notable fungal infections on white button mushrooms. Cobweb disease can cause cap spotting and brownish rot on the mushroom sporocarp. Parrag et al. (2014) investigated push-broom hyperspectral imaging configuration over 900–1700 nm for the detection of the cobweb disease on white button mushrooms in the early stages of infection. In this study, SVM was employed to discriminate fungal infections on the surface of mushroom sporocarp before they became obviously visible, and achieved over 75% of correct classification. Brown blotch disease is another fungal infection of white button mushrooms and is mainly caused by *Pseudomonas tolaasii*. Gaston et al. (2010b) used PLS-DA modelling strategy on VIS-NIR hypercube data to discriminate the brown blotches caused by bacterial infection, the browning effects from mechanical damage and the undamaged mushrooms with an OCA above 95%.

3.9. Other uses of hyperspectral imaging for quality assessment

The applications of hyperspectral imaging in the vegetable sector are numerous and not limited to the categories of parameters presented above. Sánchez et al. (2013) accurately predicted the harvest month and growing method of asparagus with accuracies ranging between 82 and 91% for the growing method and between 87 and 98% for the harvest month. Hernández-Hierro et al. (2014) reported opportunities of hyperspectral imaging to quantify glucosinolates in broccoli in the wavelength range of 950–1650 nm. Besides, Susic et al. (2018) used hyperspectral imaging to distinguish between abiotic (water-stress) and biotic stress (nematode infestation) in tomatoes. The accuracy for differentiating between well-watered and water-deficient plants was up to

100%, and between 90 and 100% when classifying nematode-infested plants. Moreover, hyperspectral imaging was used to detect contaminants in vegetables. For example, [Mo et al. \(2017\)](#) developed a discrimination method for biological contaminants (worms) in fresh-cut lettuce with excellent accuracies ranging between of 97% and 100%, and [Siripatrawan et al. \(2011\)](#) proposed a detection system for *Escherichia coli* in packaged fresh spinach by combining hyperspectral imaging with chemometrics, thus allowing for a rapid and easy interpretation. Furthermore, other diseases also attracted the interest of researchers. [Moghadam et al. \(2017\)](#) classified tomato wilt spotted virus-infected capsicum plants and [Luo et al. \(2012\)](#) used hyperspectral imaging in the range of 400–2500 nm to recognise anthracnose and brown spot on cucumber. The evaluation of ripeness was also studied. [Zhu et al. \(2015\)](#) evaluated the ripeness of tomatoes using PLS-DA by classifying them into one of six ripeness grades and achieved an OCA between 80.2% and 88.4% depending on the ripeness grade.

The application of hyperspectral imaging for leafy vegetables, such as lettuce, celery, spinach, and brassica, tends to focus more on phenotyping and breeding ([Tripodi et al., 2018](#)). Important quality parameters could be determined with hyperspectral imaging, e.g., dry matter, total soluble solids and polyphenols in leeks ([Golubkina et al., 2018](#)) and celery ([Golubkina et al., 2020](#)). Current applications are limited to nitrogen and water content estimation for spinach with R^2_p values of 0.41 and 0.87, respectively ([Corti et al., 2017](#); [Diezma et al., 2013](#)).

Within this review, studies were mostly carried out on fresh plant material. Information on quality parameters and biochemical constitution could also be obtained through dry material with the advantage of a homogeneous product and a repeatable measurement and setup compared to fresh material. Although the applications are limited, some examples are the estimation of capsaicinoids in hot peppers ([Park & Bae, 2008](#)) or glucosinolates for broccoli ([Hernández-Hierro et al., 2014](#)).

4. Application of machine learning for hyperspectral data analysis

This section is organised into two parts. First, typical tasks in machine learning are discussed in more detail. Several ML models were already presented in the previous sections, and are now further illustrated in the context of the various tasks within ML. Then, in the second part, the focus shifts to a relatively new sub-branch of ML: deep learning. Here, a brief explanation is given of what DL is, and illustrates the state of the art with some examples of applications of DL in the horticultural sector.

4.1. Machine learning tasks

ML can be used in different kinds of tasks and applications, but in general classification and regression are the most common one. These tasks are also often close to each other and are even combined in certain applications, as was actually already evident from the applications of the various parameters in Section 3.

4.1.1. Classification

A first and very common task in ML is classification. It aims to predict discrete responses and to classify data into different categories by assigning a class label to each data point. In recent years, different ML techniques were implemented for classification problems concerning the quality assessment of fruits and vegetables. In the reviewed works, the most common models for classification tasks are (i) SVM and (ii) PLS-DA).

A common application of classification in the horticultural and agricultural domain concerns the classification between healthy and affected plants or fruits. Among the reviewed papers ([Table 1](#)), SVM reached accuracies ranging from 79.2% ([Calderon et al., 2015](#)) to 100% ([Susic et al., 2018](#)), depending on the specific classification problem, in these cases verticillium wilt in olive trees and nematodes in tomatoes, respectively. Similar accuracies were found for PLS models, also depending on the problem and the various parameters. [Yuan et al. \(2021\)](#) used a CARS-PLS-DA model to discriminate between internal bruised and non-bruised Lingwu long jujube, resulting in an accuracy for the calibration and prediction set of 87% and 91%, respectively ([Yuan et al., 2021](#)). A PLS-DA model for the classification of potato tubers infected with blackspot was used by [Lopez-Maestresalas et al. \(2016\)](#) resulting in an accuracy of 98.56%. [Wendel et al. \(2018\)](#) also used a PLS-DA for the classification of mango maturity (dry matter content) and compared it with a CNN approach (Section 4.2), resulting in an F_1 -score of more than 97%.

Although SVM and PLS(-DA) are strongly represented for classification tasks with hyperspectral imaging, other ML techniques are also used with high accuracies for some tasks, including LDA ([Teena et al., 2014](#)), Quadratic Discriminant Analysis (QDA) ([Teena et al., 2014](#)), k-Nearest Neighbors (kNN) ([Rivera et al., 2014](#)), Adaptive Bayes in combination with a Gaussian Mixture Model (GMM) ([Bauer et al., 2011](#)), Random Forest (RF) ([Loggenberg et al., 2018](#)), Radial Basis Function (RBF) ([Abdulridha, Ampatzidis, Roberts, et al., 2020](#)) and ANN such as Multilayer Perceptron (MLP) ([Abdulridha, Ampatzidis, Kakarla, et al., 2020](#); [Gutierrez et al., 2018](#)). In the aforementioned work of [Calderon et al. \(2015\)](#) of verticillium wilt in olive trees, LDA performed worse than SVM with an accuracy of 59% compared to 79%. [Gutierrez et al. \(2018\)](#) compared models based on SVM and MLP for the classification of grapevine varieties, and achieved the best results for the models trained with MLP. In general, it can be concluded that the employment of different data, models and tasks, as well as the use of different metrics, complicates comparison of the results between the different works ([Kamilaris & Prenafeta-Boldu, 2018](#)). As was also already seen in Section 3, most studies deal with classification of fruits, while (leafy) vegetables are less represented.

4.1.2. Regression

Regression is a second important task in ML, used to model the correlation between dependent variable and independent variables. Although there is an enormous variation of applications with regression for quality detection in fruit and vegetables, two applications seem more common, namely the modelling of SSC (e.g., of pepper ([Ignat et al., 2014](#)), persimmon ([Wei et al., 2020](#)), kiwifruit ([Ma et al., 2021a](#)), apple ([Ma et al., 2021b](#)), mango ([Rungpichayapichet et al., 2017](#))), and

of firmness (e.g., of banana (Rajkumar et al., 2012), tomato (Huang, Lu, Xu, et al., 2018), mango (Rungpichayapichet et al., 2017)). However, as already mentioned, there are many other uses for regression models as well, such as modeling/prediction of water and nitrogen contents (spinach (Corti et al., 2017)), titratable acidity (mango (Rungpichayapichet et al., 2017)), pH (tomato (Huang, Lu, Xu, et al., 2018)), maturity estimation (mango (Wendel et al., 2018)). Analogous to previous observations, fruit is more represented than (leafy) vegetables in the various applications.

Some of the most widely used ML models for this type of applications are PLS(R) (Corti et al., 2017; Huang, Lu, Hu, et al., 2018; Huang, Lu, Xu, et al., 2018; Ma et al., 2021a, 2021b; Wei et al., 2020) and MLR (Rajkumar et al., 2012).

4.1.3. Other ML tasks

Of course, the possible tasks in ML are not limited to classification and regression, although these do account for the largest share in the application of ML in horticulture and agriculture. Another ML task is object detection, where the goal is to localise and detect certain objects in an image (e.g., humans, cars, animals, etc.). In horticultural sector, this is mainly reflected in applications such as localising fruits for automatic fruit counting (Song et al., 2014) or to develop a picking robot (Peng et al., 2019). However, most of these applications are (still) based on RGB data. A more common use case of object detection with hyperspectral data is to detect defects such as mechanical damage (Rivera et al., 2014) and diseases (Lopez-Maestresalas et al., 2016). This task is more-over closely related to classification, as in most cases it also involves assigning a class or type to the detected object. Analogous models are used, e.g., kNN for the detection of mechanical damage in mangos (Rivera et al., 2014) or PLS-DA for the detection of blackspot in potato (Lopez-Maestresalas et al., 2016), with a comparable OCA of 98% and 94%, respectively.

In addition, segmentation can also be seen as a typical ML task, related to both classification and detection task. Segmentation aims to label each pixel in an image as a pointwise classification. In this way, similarly labelled pixels can be grouped, creating segments in the image, to simplify an image and give an exact outline of the objects inside of it. Li et al. (2018) used an improved watershed segmentation on multi-spectral images of peaches to discriminate bruised from sound regions.

It can be concluded, based on the various studies mentioned, that there is a wide range of ML models available. Therefore, given a specific problem, determining the most appropriate model to replace the more traditional manual methods to assess fruit and vegetable quality, is challenging.

4.2. Deep learning

DL techniques use deeper ANN through the use of more neurons, more layers and more complex ways of connections between layers. The advantage of these deeper techniques lies in their ability to learn features automatically rather than relying on manual feature engineering, as is the case with more traditional ML techniques (LeCun et al., 2015). In this way, the system can solve more complex problems, is flexible,

and is therefore able to capture more variability in data, which is interesting in for horticultural applications. A common type of network for multidimensional data is a CNN, as fully connected neural networks (FCNN), such as MLP, are not optimal for this kind of data. In a FCNN, a neuron in a given layer is connected to all other neurons in the next layer, resulting in a huge amount of input for multidimensional data. Moreover, in this data, information is typically spread across the different dimensions, which makes linear processing not an optimal strategy. After all, if a neuron contains a particular piece of information, it is more likely that the environment surrounding it will also contain relevant, spatial information. Neurons are thus not linearly connected in a convolutional layer, but rather through a local system of receptive fields over the depth of the input data. A CNN typically consists of multiple convolutional layers, each of which attempts to extract a certain type of features from the data. Using these layers, different representations of the input data are created, ranging from more general representations in the first layers to more specific representations deeper in the network.

In the last years, DL was widely studied and improved the state of the art in many different sectors (LeCun et al., 2015). DL research in different sectors opened a window of opportunity for applications in horticulture, agriculture and food. Kamilaris and Prenafeta-Boldu (2018) provide an overview of the first applications of DL in precision agriculture, showing its importance in recent agricultural applications such as crop and/or weed detection and classification, disease detection, fruit counting, yield estimation, and plant phenotyping, being closely related to horticultural applications. The interest in and the use of DL in the food domain is also rising. Zhou et al. (2019) present an overview of the research progress of DL in food, with applications such as food recognition and classification, calorie estimation and quality detection. Naranjo-Torres et al. (2020) specifically focus on the use of CNN for fruit image processing (e.g., fruit classification, detection, quality control) and observes that in this research domain CNNs were not yet represented before 2019. In addition, Fracaroli et al. (2021) state that one of the main objectives of future agriculture will be increasing food quality, and that DL can play an important role in this. Although the number of studies with DL is increasing, there are still many unexplored opportunities, including the combination of DL with hyperspectral data for the quality assessment of fruits and vegetables.

To the best of our knowledge, most of the works involving DL for applications with fruit, mushrooms and vegetables make use of RGB data. It is typically used for the detection of the fruits in a bigger image as a vital element of an autonomous robotic platform. For example, Peng et al. (2019) proposed a classification algorithm using a CNN to classify images into either a tomato or not. Liu et al. (2019) also used a convolutional network, but proposed a pixel-wise segmentation method using Mask Region Based Convolutional Neural Network (Mask R-CNN), based on Excess Green (ExG) for cucumbers. Nonetheless, there are already a few cases where DL was successful in combination with hyperspectral imaging for the quality assessment of fruit, mushrooms and vegetables, most of which date back to the last five years. It demonstrates strong capabilities for the different ML tasks, addressing

different variants of a CNN. In the eight papers reviewed about this topic, only one paper does not mention a convolutional network. In this work, [Yu et al. \(2018\)](#) applied a two-step method to predict firmness and SSC of postharvest Korla fragrant pear by extracting deep spectral features using a Sparse Autoencoder (SAE) and feeding those to a FCNN.

In all other works, both common CNN architectures such as AlexNet, Visual Geometry Group (VGG), ResNet etc. ([Dong et al., 2021](#); [Gao et al., 2020](#); [Garillos-Manliguez & Chiang, 2021](#); [Wang et al., 2018](#)), and self-composed CNN ([Liu et al., 2018](#); [Nagasubramanian et al., 2018](#)) emerge. The self-composed architecture of [Nagasubramanian et al. \(2018\)](#) consists of convolutional and pooling layers for the classification of healthy and charcoal rot-infected soybean. [Liu et al. \(2018\)](#) also proposed a self-composed architecture, in combination with a Stacked SAE (SSAE), constructed by stacking multiple SAEs hierarchically on top of each other and so creating a two-step method CNN-SSAE, inverted compared to the one of [Yu et al. \(2018\)](#). In the cascaded system of [Liu et al.](#), the CNN is used to localise surface defective regions based on the RGB extracted wavelengths of the hyperspectral image. The SSAE then uses the image parts of the defective regions as input for classification.

In contrast to self-composed architectures, numerous variants of existing CNN architectures have been proposed in recent years. These can be used, with appropriate modifications and whether or not pretrained, for various applications and datasets, including hyperspectral ones. [Wang et al. \(2018\)](#), [Gao et al. \(2020\)](#) and [Dong et al. \(2021\)](#) respectively used ResNet/ResNeXt, AlexNet and ResNet in their systems to detect mechanical damage in blueberries, strawberry ripeness and species identification of bolete mushrooms. [Garillos-Manliguez and Chiang \(2021\)](#) compared different architectures, i.e., MobileNet/MobileNetV2, ResNet50/ResNeXt50, VGG16/VGG19 and AlexNet for maturity estimation of papaya, which resulted in a corresponding range of accuracies between 55% and 88%, making AlexNet the best option in this case. The accuracies of the self-composed architectures range from 88.3 to 95.7%, which is comparable to the accuracies of the works using known architectures. However, it is still difficult to compare the results of different works, given the large differences in datasets and metrics used ([Kamilaris & Prenafeta-Boldu, 2018](#); [Meenu et al., 2021](#)). Both strategies have interesting characteristics and many possibilities, but besides the choice between a self-composed or existing CNN, it is mainly important to fine-tune the hyperparameters of the network to the desired application and corresponding datasets, taking into account an acceptable train and validation time with existing hardware.

In addition, two studies used multimodal data to combine the sensitivity of hyperspectral imaging and the precision of RGB images. By using multiple modalities of information, the information of one modality (in this case RGB data) can complement and/or reinforce the other modalities to overcome the limitations of using only one type of data. [Garillos-Manliguez and Chiang \(2021\)](#) used this technique for maturity estimation of papaya by combining the information of RGB-data (e.g., morphological changes) and hyperspectral data (e.g., internal properties). Different popular architectures (i.e., MobileNet, MobileNetV2, ResNet50, ResNeXt50, VGG16,

VGG19, and AlexNet) were adapted toward the multimodal data. [Sa et al. \(2016\)](#) used multimodal data (RGB and NIR) to focus on the detection of fruits (sweet pepper and rock melon, but adaptable to other fruits) with a Faster R-CNN network, analogous to application in the works of [Liu et al. \(2019\)](#) and [Peng et al. \(2019\)](#) as basis for an autonomous harvesting robot. They compared early and late fusion of the multimodal data and concluded that the late fusion method performs better, with an F_1 -score of 0.84 compared to 0.80 for early fusion, 0.82 for RGB only and 0.80 for NIR only.

Some works not only used DL techniques, but also compared them with earlier presented ML techniques. [Wendel et al. \(2018\)](#) compared PLS and CNN for the classification of mango maturity. The different CNNs had slightly higher F_1 -scores of 0.985 and 0.989, than PLS with a mean F_1 -score of 0.972. The same trend is visible in the work of [Wang et al. \(2018\)](#) about the detection of internal mechanical damage in blueberries. The CNNs (ResNet/ResNeXt) resulted in F_1 -scores of 0.8952 and 0.8905, and thus outperformed SMO ($F_1 = 0.83$), linear regression ($F_1 = 0.78$), RF ($F_1 = 0.75$), bagging ($F_1 = 0.73$) and MLP ($F_1 = 0.80$) techniques.

It can be concluded that DL techniques have great potential for hyperspectral imaging systems, as they can simplify the overall processing pipeline. Although there are still many challenges to overcome, which are highlighted in Section 5 as DL will certainly be further explored in the near future. In addition, a similar trend as for ML techniques can be observed about the ratio fruit/vegetables: DL for applications concerning fruit are more present in papers than (leafy) vegetables. Thus, the use of DL, certainly in combination with vegetable data, is an interesting gap to explore in the coming years.

5. Future considerations

In general, the main challenges with hyperspectral imaging for quality assessment of fruit, vegetables and mushrooms, are related to data availability and reliability of the available models and results, especially when considering the applicability. A lot of different studies have already been done on many diverse cases, but in general they tend to lack a more practical focus. Only a few studies ([Li et al., 2011](#); [Riccioli et al., 2021](#); [Schmilovitch et al., 2014](#)) carried out true test set validations outside of calibration/training dataset from a different year or orchard, not simply splitting a single dataset. Moreover, to achieve reliable prediction models and increased knowledge transfer from researchers to horticultural practice, research should focus more on repeatability and the transfer from research parameters (chlorophyll content, firmness, SSC) to relevant industrial parameters. Furthermore, successful non-destructive assessment of quality parameters with hyperspectral imaging on the field, during storage and handling and on the shelf, can lead to the optimal use of production inputs and a reduction of food waste, while ensuring safe and high-quality food.

The end goal of hyperspectral imaging systems is generally not to work towards a practical application. Data acquisition processes for inline applications should be carried out more rapidly. Multispectral imaging techniques that focus on a smaller set of relevant wavelength ranges for the specific

application should be used. Although several techniques already exist to reduce redundant wavelengths in the (statistical) analysis (Section 2.1), the development and standardisation of spectral compression methods and variable selection algorithms should certainly be considered in the near future. Finally, before (hyper-)spectral imaging systems find their way into mainstream applications, (pre-)processing and analysis of spectral imaging system should strive towards standards to handle spectral information and facilitate availability of spectral data. At the moment, a lot of different models and parameters are used in a lot of use cases, it is difficult to generalise toward practice.

DL techniques are powerful ML methods that recently showed their supremacy on various applications (LeCun et al., 2015), however, it has shortcomings as well. More data will generally be needed when a problem becomes more complex (Kamilaris & Prenafeta-Boldu, 2018), and so, the training time will also increase, certainly with hyperspectral imaging (Zhou et al., 2019). Besides that, most of the aforementioned works must be considered a proof of concept, rather than operational methods which are directly applicable in practice. This is because they mostly rely on supervised models, requiring an extensive labelled training dataset, which is very time-intensive and therefore an expensive process, certainly for hyperspectral data. Labelling is a big bottleneck in supervised ML techniques, and in general, DL has an even bigger need for large datasets due to its complexity.

The success of DL methods in various sectors, is partly due to the availability of large-scale, labelled and public datasets, such as ImageNet (Deng et al., 2009). In the case of limited availability of labelled data, transfer learning can provide a solution. Here, a network is first pretrained on a large and generic database such as ImageNet and then further optimised based on the limited dataset (Utkin et al., 2016). The major drawback, however, is that most of the available large datasets consist of RGB images, making the transition to hyperspectral data not evident (Signoroni et al., 2019). Consequently, it would be useful for further research of DL techniques, as well as for more concrete and practical use, to build and share a large-scale database of images capturing various defects in fruits and vegetables (Nturambirwe & Opara, 2020).

Another challenge concerns the intrinsic variability of horticultural data, so the current trend is incorporating more and more labelled datasets is unlikely to be sufficient. After all, this variability makes it difficult to generalise such supervised models to the varying conditions in horticulture (e.g., cultivars, different growth stages, lighting condition, etc.), let alone to other crop species or applications (Slaughter et al., 2008). This immediately leads to the important nuance, as mentioned above, regarding the results discussed in the previous sections: in almost all cases, the test dataset was similar to the training dataset, which does not sufficiently test the generalisation capabilities of the network. Generalisation is an important aspect, certainly taking into account the practical applicability of the models.

Hence, there is a trend towards developing DL methods that avoid the high amounts of labelled data, such as unsupervised and semi-supervised learning techniques in different

(research) sectors (Guo et al., 2019). In unsupervised methods, the system will independently detect patterns in the input data, without an annotated ground truth (Sathya & Abraham, 2013). Although, the drawback of those methods is that there is no way to force the method to detect a desired signal. Weakly or semi-supervised techniques, being a combination of both, can combine the advantages of both techniques. In those systems, only a small part of the input dataset must be labelled, with the bulk of the data remaining unlabelled. The system itself can then be understood as an unsupervised detection of patterns, combined with specialisation on the annotated data. Different attempts were made to reduce the labelling bottleneck with DL, as well as ML, in the broad sector of agriculture and horticulture. However, most examples use RGB data and were applied in arable farming. For example, Potena et al. (2017) introduced an algorithm for automatic selection of a good covering subset to be labelled (cfr. the knapsack problem) to detect weeds in sunflower fields and Di Cicco et al. (2016) proposed a method to create synthetic additional data of weeds. Another interesting research area in the field of weakly supervised learning, is the use of active learning, based on iteratively querying the user. For example, Hu et al. (2018) describe an active learning algorithm (estimated error reduction) using hyperspectral imaging to classify blueberry damage by using only nine labelled samples. An alternative solution to the high need for data, is the use of synthetic data, which will certainly be explored in the near future.

Despite the differences in the reviewed works (e.g., RGB, other crops, limited testing), it may be of interest to further develop and apply these types of techniques and concepts to hyperspectral images in the context of quality assessment for fruit and vegetables. The transfer of ML techniques from other applications and other types of data will therefore be an interesting challenge in the future. To this end, it is important that the current methods become standardised and made comparable as much as possible, whereby the availability of a public dataset will play an important role.

Author contributions

Conceptualisation, S.F., D.A. and J.V.B.; methodology, J.V.B.; investigation, J.W., K.M., I.M., M.Z. and J.V.B.; data curation, J.V.B.; writing—original draft preparation, J.W., K.M., I.M., M.Z.-S. M.Z. and J.V.B.; writing—review and editing, M.Z.-S., J.W., D.A. and A.G.; supervision, D.A., S.F.; project administration, S.F.; funding acquisition, S.F., M.Z.-S., D.A. and J.V.B. All authors have read and agreed to the published version of the manuscript.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 862665 ICT AGRI FOOD, through regional funding provided by GSRI (Greece), BMEL (Germany), DAFM (Ireland) and VLAIO (Belgium).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Abdulridha, J., Ampatzidis, Y., Kakarla, S. C., & Roberts, P. (2020). Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. *Precision Agriculture*, 21(5), 955–978. <https://doi.org/10.1007/s11119-019-09703-4>
- Abdulridha, J., Ampatzidis, Y., Roberts, P., & Kakarla, S. C. (2020). Detecting powdery mildew disease in squash at different stages using UAV-based hyperspectral imaging and artificial intelligence. *Biosystems Engineering*, 197, 135–148. <https://doi.org/10.1016/j.biosystemseng.2020.07.001>
- Abu-Khalaf, N. (2015). *Sensing tomato's pathogen using Visible/Near infrared (VIS/NIR) spectroscopy and multivariate data analysis (MVDA)*.
- Amigo, J. M., & Grassi, S. (2020). Configuration of hyperspectral and multispectral imaging systems. In *Data handling in science and technology* (Vol. 32, pp. 17–34). Elsevier.
- Amigo, J. M., Martí, I., & Gowen, A. (2013). Hyperspectral imaging and chemometrics: A perfect combination for the analysis of food structure, composition and quality. In *Data handling in science and technology* (Vol. 28, pp. 343–370). Elsevier.
- Appeltans, S., Guerrero, A., Nawar, S., Pieters, J., & Mouazen, A. M. (2020). Practical recommendations for hyperspectral and thermal proximal disease sensing in potato and leek fields. *Remote Sensing*, 12, 1939. <https://doi.org/10.3390/rs12121939>
- Baiano, A., Terracone, C., Peri, G., & Romaniello, R. (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computers and Electronics in Agriculture*, 87, 142–151. <https://doi.org/10.1016/j.compag.2012.06.002>
- Barrett, D. M., Beaulieu, J. C., & Shewfelt, R. (2010). Color, flavor, texture, and nutritional quality of fresh-cut fruits and vegetables: Desirable levels, instrumental and sensory measurement, and the effects of processing. *Critical Reviews in Food Science and Nutrition*, 50(5), 369–389. <https://doi.org/10.1080/10408391003626322>
- Bauer, S. D., Korc, F., & Forstner, W. (2011). The potential of automatic methods of classification to identify leaf diseases from multispectral images. *Precision Agriculture*, 12(3), 361–377. <https://doi.org/10.1007/s11119-011-9217-6>
- Bauriegel, E., Brabandt, H., Gärber, U., & Herppich, W. B. (2014). Chlorophyll fluorescence imaging to facilitate breeding of *Bremia lactucae*-resistant lettuce cultivars. *Computers and Electronics in Agriculture*, 105, 74–82. <https://doi.org/10.1016/j.compag.2014.04.010>
- Benelli, A., Cevoli, C., & Fabbri, A. (2020). In-field hyperspectral imaging: An overview on the ground-based applications in agriculture. *Journal of Agricultural Engineering*, 51(3), 129–139.
- Calderon, R., Navas-Cortes, J. A., & Zarco-Tejada, P. J. (2015). Early detection and quantification of verticillium wilt in olive using hyperspectral and thermal imagery over large areas. *Remote Sensing*, 7(5), 5584–5610. <https://doi.org/10.3390/rs70505584>
- Cen, H. Y., Lu, R. F., Zhu, Q. B., & Mendoza, F. (2016). Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biology and Technology*, 111, 352–361. <https://doi.org/10.1016/j.postharvbio.2015.09.027>
- Chen, C. (2015). Overview of plant pigments. In *Pigments in fruits and vegetables* (pp. 1–7). Springer.
- Cho, J. S., Lim, J. H., Park, K. J., Choi, J. H., & Ok, G. S. (2021). Prediction of pelargonidin-3-glucoside in strawberries according to the postharvest distribution period of two ripening stages using VIS-NIR and SWIR hyperspectral imaging technology. *LWT-Food Science and Technology*, 141. <https://doi.org/10.1016/j.lwt.2021.110875>
- Civille, G. V., & Oftedal, K. N. (2012). Sensory evaluation techniques - Make “good for you” taste “good”. *Physiology & Behavior*, 107(4), 598–605. <https://doi.org/10.1016/j.physbeh.2012.04.015>
- Clark, C. J., White, A., Jordan, R. B., & Woolf, A. B. (2007). Challenges associated with segregation of avocados of differing maturity using density sorting at harvest. *Postharvest Biology and Technology*, 46(2), 119–127. <https://doi.org/10.1016/j.postharvbio.2007.05.010>
- Corti, M., Marino Gallina, P., Cavalli, D., & Cabassi, G. (2017). Hyperspectral imaging of spinach canopy under combined water and nitrogen stress to estimate biomass, water, and nitrogen content. *Biosystems Engineering*, 158, 38–50. <https://doi.org/10.1016/j.biosystemseng.2017.03.006>
- Deng, J., Dong, W., Socher, R., Li, L., Kai, L., & Li, F.-F. (20-25 June 2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*.
- Dhiman, B., Kumar, Y., & Kumar, M. (2022). Fruit quality evaluation using machine learning techniques: Review, motivation and future perspectives. *Multimedia Tools and Applications*, 1–23.
- Di Cicco, M., Potena, C., Grisetti, G., & Pretto, A. (2016). *Automatic model based dataset generation for fast and accurate crop and weeds detection*.
- Diago, M. P., Fernandez-Novales, J., Fernandes, A. M., Melo-Pinto, P., & Tardaguila, J. (2016). Use of visible and short-wave near-infrared hyperspectral imaging to fingerprint anthocyanins in intact grape berries. *Journal of Agricultural and Food Chemistry*, 64(40), 7658–7666. <https://doi.org/10.1021/acs.jafc.6b01999>
- Diaz, J. J. V., Aldana, A. P. S., & Zuluaga, D. V. R. (2021). Prediction of dry matter content of recently harvested ‘Hass’ avocado fruits using hyperspectral imaging. *Journal of the Science of Food and Agriculture*, 101(3), 897–906. <https://doi.org/10.1002/jsfa.10697>
- Diezma, B., Lleó, L., Roger, J. M., Herrero-Langreo, A., Lunadei, L., & Ruiz-Altsient, M. (2013). Examination of the quality of spinach leaves using hyperspectral imaging. *Postharvest Biology and Technology*, 85, 8–17. <https://doi.org/10.1016/j.postharvbio.2013.04.017>
- Dong, J. L., & Guo, W. C. (2015). Nondestructive determination of apple internal qualities using near-infrared hyperspectral reflectance imaging. *Food Analytical Methods*, 8(10), 2635–2646. <https://doi.org/10.1007/s12161-015-0169-8>
- Dong, J.-E., Zhang, J., Zuo, Z.-T., & Wang, Y.-Z. (2021). Deep learning for species identification of bolete mushrooms with two-dimensional correlation spectral (2DCOS) images. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 249, Article 119211. <https://doi.org/10.1016/j.saa.2020.119211>
- Esquerre, C., Gowen, A. A., Downey, G., & O'Donnell, C. P. (2012). Wavelength selection for development of a near infrared imaging system for early detection of bruise damage in mushrooms (*Agaricus bisporus*). *Journal of Near Infrared Spectroscopy*, 20(5), 537–546. <https://doi.org/10.1255/jnirs.1014>
- Fracarolli, J. A., Pavarin, F. F. A., Castro, W., & Blasco, J. (2021). Computer vision applied to food and agricultural products. *Revista Ciência Agronômica*, 51.
- Gan, Q., Wang, X., Wang, Y., Xie, Z., Tian, Y., & Lu, Y. (2017). Culture-free detection of crop pathogens at the single-cell

- level by micro-Raman spectroscopy. *Advanced Science*, 4, Article 1700127. <https://doi.org/10.1002/advs.201700127>
- Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y., & Han, X. (2020). Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. *Artificial Intelligence in Agriculture*, 4, 31–38. <https://doi.org/10.1016/j.aiaa.2020.04.003>
- Garillos-Manlriquez, C. A., & Chiang, J. Y. (2021). Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation. *Sensors*, 21(4). <https://doi.org/10.3390/s21041288>
- Gaston, E., Frias, J. M., Cullen, P. J., O'Donnell, C. P., & Gowen, A. A. (2010a). Prediction of polyphenol oxidase activity using visible near-infrared hyperspectral imaging on mushroom (*Agaricus bisporus*) caps. *Journal of Agricultural and Food Chemistry*, 58(10), 6226–6233. <https://doi.org/10.1021/jf100501q>
- Gaston, E., Frias, J. M., Cullen, P. J., O'Donnell, C. P., & Gowen, A. A. (2010b). Visible-near infrared hyperspectral imaging for the identification and discrimination of brown blotch disease on mushroom (*Agaricus bisporus*) caps. *Journal of Near Infrared Spectroscopy*, 18(5), 341–353. <https://doi.org/10.1255/jnirs.894>
- Girod, D., Landry, J., Doyon, G., Osuna-García, J., Salazar-García, S., & Geonaga, R. (2008). Evaluating hass avocado maturity using hyperspectral imaging.
- Goetz, A. F., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *Science*, 228(4704), 1147–1153.
- Golubkina, N. A., Kharchenko, V. A., Moldovan, A. I., Koshevarov, A. A., Zamana, S., Nadezhkin, S., Soldatenko, A., Sekara, A., Tallarita, A., & Caruso, G. (2020). Yield, growth, quality, biochemical characteristics and elemental composition of plant parts of celery leafy, stalk and root types grown in the Northern Hemisphere. *Plants*, 9, 484. <https://doi.org/10.3390/plants9040484>
- Golubkina, N., Seredin, T., Antoshkina, M., Kosheleva, O., Teliban, G., & Caruso, G. (2018). Yield, quality, antioxidants and elemental composition of new leek cultivars under greenhouse organic or conventional system. *Horticulturae*, 4, 39. <https://doi.org/10.3390/horticulturae4040039>
- Gowen, A. A., O'Donnell, C. P., Taghizadeh, M., Cullen, P. J., Frias, J. M., & Downey, G. (2008). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (*Agaricus bisporus*). *Journal of Chemometrics*, 22(3–4), 259–267. <https://doi.org/10.1002/cem.1127>
- Gowen, A. A., O'Donnell, C. P., Taghizadeh, M., Gaston, E., O'Gorman, A., Cullen, P. J., Frias, J. M., Esquerre, C., & Downey, G. (2008). Hyperspectral imaging for the investigation of quality deterioration in sliced mushrooms (*Agaricus bisporus*) during storage. *Sensing and Instrumentation for Food Quality and Safety*, 2(3), 133–143. <https://doi.org/10.1007/s11694-008-9042-4>
- Gowen, A. A., Taghizadeh, M., & O'Donnell, C. P. (2009). Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *Journal of Food Engineering*, 93(1), 7–12. <https://doi.org/10.1016/j.jfoodeng.2008.12.021>
- Guo, L.-Z., Li, Y.-F., Li, M., Yi, J.-F., Zhou, B.-W., & Zhou, Z.-H. (2019). Reliable weakly supervised learning: Maximize gain and maintain safeness.
- Gutierrez, S., Fernandez-Novales, J., Diago, M. P., & Tardaguila, J. (2018). On-the-go hyperspectral imaging under field conditions and machine learning for the classification of grapevine varieties. *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.01102>
- Gutierrez, S., Tardaguila, J., Fernandez-Novales, J., & Diago, M. P. (2019). On-the-go hyperspectral imaging for the in-field estimation of grape berry soluble solids and anthocyanin concentration. *Australian Journal of Grape and Wine Research*, 25(1), 127–133. <https://doi.org/10.1111/ajgw.12376>
- Huang, Y., Lu, R., Hu, D., & Chen, K. (2018). Quality assessment of tomato fruit by optical absorption and scattering properties. *Postharvest Biology and Technology*, 143, 78–85. <https://doi.org/10.1016/j.postharvbio.2018.04.016>
- Huang, Y., Lu, R., Xu, Y., & Chen, K. (2018). Prediction of tomato firmness using spatially-resolved spectroscopy. *Postharvest Biology and Technology*, 140, 18–26. <https://doi.org/10.1016/j.postharvbio.2018.02.008>
- Huang, M., Wang, Q., Zhang, M., & Zhu, Q. (2014). Prediction of color and moisture content for vegetable soybean during drying using hyperspectral imaging technology. *Journal of Food Engineering*, 128, 24–30.
- Hu, M.-H., Zhao, Y., & Zhai, G.-T. (2018). Active learning algorithm can establish classifier of blueberry damage with very small training dataset using hyperspectral transmittance data. *Chemometrics and Intelligent Laboratory Systems*, 172, 52–57. <https://doi.org/10.1016/j.chemolab.2017.11.012>
- Ignat, T., Alchanatis, V., & Schmilovitch, Z. (2014). Maturity prediction of intact bell peppers by sensor fusion. *Computers and Electronics in Agriculture*, 104, 9–17. <https://doi.org/10.1016/j.compag.2014.03.006>
- Jaiswal, G., Sharma, A., & Yadav, S. K. (2021). Critical insights into modern hyperspectral image applications through deep learning. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 11(6), e1426.
- Kader, A. A. (2002). *Postharvest technology of horticultural crops* (Vol. 3311). University of California Agriculture and Natural Resources.
- Kalac, P. (2016). *Edible mushrooms: Chemical composition and nutritional value*. Academic Press.
- Kalac, P. (2019). *Mineral composition and radioactivity of edible mushrooms*. Academic Press.
- Kamilaris, A., & Prenafeta-Boldu, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. (2019). Deep learning—Method overview and review of use for fruit detection and yield estimation. *Computers and Electronics in Agriculture*, 162, 219–234.
- Kratsch, H. A., & Wise, R. R. (2000). The ultrastructure of chilling stress. *Plant, Cell and Environment*, 23(4), 337–350. <https://doi.org/10.1046/j.1365-3040.2000.00560.x>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Li, J., Chen, L., & Huang, W. (2018). Detection of early bruises on peaches (*Amygdalus persica* L.) using hyperspectral imaging coupled with improved watershed segmentation algorithm. *Postharvest Biology and Technology*, 135, 104–113. <https://doi.org/10.1016/j.postharvbio.2017.09.007>
- Li, J. B., Huang, W. Q., Tian, X., Wang, C. P., Fan, S. X., & Zhao, C. J. (2016). Fast detection and visualization of early decay in citrus using Vis-NIR hyperspectral imaging. *Computers and Electronics in Agriculture*, 127, 582–592. <https://doi.org/10.1016/j.compag.2016.07.016>
- Li, J., Huang, W., & Zhao, C. (2015). Machine vision technology for detecting the external defects of fruits—A review. *The Imaging Science Journal*, 63(5), 241–251.
- Li, Y., Luo, Z., Wang, F., & Wang, Y. (2020). Hyperspectral leaf image-based cucumber disease recognition using the extended collaborative representation model. *Sensors (Basel)*, 20(14). <https://doi.org/10.3390/s20144045>
- Lin, X., Xu, J.-L., & Sun, D.-W. (2019). Investigation of moisture content uniformity of microwave-vacuum dried mushroom (*Agaricus bisporus*) by NIR hyperspectral imaging. *LWT*, 109, 108–117. <https://doi.org/10.1016/j.lwt.2019.03.034>
- Li, J. B., Rao, X. Q., & Ying, Y. B. (2011). Detection of common defects on oranges using hyperspectral reflectance imaging.

- Computers and Electronics in Agriculture, 78(1), 38–48. <https://doi.org/10.1016/j.compag.2011.05.010>
- Liu, D. Y., & Guo, W. C. (2015). Identification of kiwifruits treated with exogenous plant growth regulator using near-infrared hyperspectral reflectance imaging. *Food Analytical Methods*, 8(1), 164–172. <https://doi.org/10.1007/s12161-014-9885-8>
- Liu, Z., He, Y., Cen, H., & Lu, R. (2018). Deep feature representation with stacked sparse auto-encoder and convolutional neural network for hyperspectral imaging-based detection of cucumber defects. *Transactions of the ASABE*, 61(2), 425–436. <https://doi.org/10.13031/trans.12214>
- Liu, X., Zhao, D., Jia, W., Ji, W., Ruan, C., & Sun, Y. (2019). Cucumber fruits detection in greenhouses based on instance segmentation. *IEEE Access*, 7, 139635–139642. <https://doi.org/10.1109/ACCESS.2019.2942144>
- Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a Shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing*, 10(2). <https://doi.org/10.3390/rs10020202>
- Lopez-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jaren, C., & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes. by Vis-NIR and SWIR hyperspectral imaging. *Food Control*, 70, 229–241. <https://doi.org/10.1016/j.foodcont.2016.06.001>
- Lu, R., Ariana, D. P., & Cen, H. (2011). Optical absorption and scattering properties of normal and defective pickling cucumbers for 700–1000 nm. *Sensing and Instrumentation for Food Quality and Safety*, 5, 51–56. <https://doi.org/10.1007/s11694-011-9108-6>
- Lu, Y., & Lu, R. (2019). siriTool: A matlab graphical user interface for image analysis in structured-illumination reflectance imaging for fruit defect detection. In *2019 ASABE Annual International Meeting*.
- Luo, X., Takahashi, T., Kyo, K., & Zhang, S. H. (2012). Wavelength selection in vis/NIR spectra for detection of bruises on apples by ROC analysis. *Journal of Food Engineering*, 109(3), 457–466. <https://doi.org/10.1016/j.jfoodeng.2011.10.035>
- Lu, Q., & Tang, M. J. (2012). Detection of hidden bruise on kiwi fruit using hyperspectral imaging and parallelepiped classification. In *2011 International Conference of Environmental Science and Engineering* (Vol 12, Pt B, 12, pp. 1172–1179). <https://doi.org/10.1016/j.proenv.2012.01.044>
- Lu, R., Van Beers, R., Saeys, W., Li, C., & Cen, H. (2020). Measurement of optical properties of fruits and vegetables: A review. *Postharvest Biology and Technology*, 159, Article 111003. <https://doi.org/10.1016/j.postharvbio.2019.111003>
- Lu, H. D., Yu, X. J., Zhou, L. J., & He, Y. (2018). Selection of spectral resolution and scanning speed for detecting green jujubes chilling injury based on hyperspectral reflectance imaging. *Applied Sciences-Basel*, 8(4). <https://doi.org/10.3390/app8040523>
- Martinez-Romero, D., Serrano, M., Carbonell, A., Castillo, S., Riquelme, F., & Valero, D. (2004). Mechanical damage during fruit post-harvest handling: Technical and physiological implications. In *Production practices and quality assessment of food crops* (pp. 233–252). Springer.
- Ma, T., Xia, Y., Inagaki, T., & Tsuchikawa, S. (2021a). Non-destructive and fast method of mapping the distribution of the soluble solids content and pH in kiwifruit using object rotation near-infrared hyperspectral imaging approach. *Postharvest Biology and Technology*, 174. <https://doi.org/10.1016/j.postharvbio.2020.111440>
- Ma, T., Xia, Y., Inagaki, T., & Tsuchikawa, S. (2021b). Rapid and nondestructive evaluation of soluble solids content (SSC) and firmness in apple using Vis-NIR spatially resolved spectroscopy. *Postharvest Biology and Technology*, 173, Article 111417. <https://doi.org/10.1016/j.postharvbio.2020.111417>
- Meenu, M., Kurade, C., Neelapu, B. C., Kalra, S., Ramaswamy, H. S., & Yu, Y. (2021). A concise review on food quality assessment using digital image processing. *Trends in Food Science & Technology*, 118, 106–124.
- Megeto, G. A. S., Silva, A. G.d., Bulgarelli, R. F., Bublitz, C. F., Valente, A. C., & Costa, D. A. G.d. (2021). Artificial intelligence applications in the agriculture 4.0. *Revista Ciência Agronômica*, 51.
- Mehl, P. M., Chao, K., Kim, M., & Chen, Y. R. (2002). Detection of defects on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied Engineering in Agriculture*, 18(2), 219–226.
- Meng, Q. L., Shang, J., Huang, R. S., & Zhang, Y. (2021). Determination of soluble solids content and firmness in plum using hyperspectral imaging and chemometric algorithms. *Journal of Food Process Engineering*, 44(1). <https://doi.org/10.1111/jfpe.13597>
- Miguel Hernández-Hierro, J., Esquerre, C., Valverde, J., Villacreces, S., Reilly, K., Gaffney, M., Lourdes González-Miret, M., & Heredia, F. J. (2014). Preliminary study on the use of near infrared hyperspectral imaging for quantitation and localisation of total glucosinolates in freeze-dried broccoli. *Journal of Food Engineering*, 126, 107–112. <https://doi.org/10.1016/j.jfoodeng.2013.11.005>
- Miller, A. (1992). Physiology, biochemistry and detection of bruising (mechanical stress) in fruits and vegetables. *Postharvest News and Information (United Kingdom)*.
- Moghadam, P., Ward, D., Goan, E., Jayawardena, S., Sikka, P., & Hernandez, E. (2017). Plant disease detection using hyperspectral imaging. In *2017 International Conference on Digital Image Computing - Techniques and Applications (Dicta)* (pp. 384–391).
- Mo, C., Kim, G., Kim, M. S., Lim, J., Lee, S. H., Lee, H. S., & Cho, B. K. (2017). Discrimination methods for biological contaminants in fresh-cut lettuce based on VNIR and NIR hyperspectral imaging. *Infrared Physics & Technology*, 85, 1–12. <https://doi.org/10.1016/j.infrared.2017.05.003>
- Mo, C., Kim, G., Lim, J., Kim, M., Cho, H., & Cho, B.-K. (2015). Detection of lettuce discoloration using hyperspectral reflectance imaging. *Sensors*, 15, 29511–29534. <https://doi.org/10.3390/s151129511>
- Mollazade, K. (2017). Non-destructive identifying level of browning development in button mushroom (*Agaricus bisporus*) using hyperspectral imaging associated with chemometrics. *Food Analytical Methods*, 10(8), 2743–2754. <https://doi.org/10.1007/s12161-017-0845-y>
- Mollazade, K., Omid, M., Tab, F. A., Kalaj, Y. R., & Mohtasebi, S. S. (2015). Data mining-based wavelength selection for monitoring quality of tomato fruit by backscattering and multispectral imaging. *International Journal of Food Properties*, 18(4), 880–896. <https://doi.org/10.1080/10942912.2013.835822>
- Munera, S., Gomez-Sanchis, J., Aleixos, N., Vila-Frances, J., Colelli, G., Cubero, S., Soler, E., & Blasco, J. (2021). Discrimination of common defects in loquat fruit cv. ‘Algerie’ using hyperspectral imaging and machine learning techniques. *Postharvest Biology and Technology*, 171. <https://doi.org/10.1016/j.postharvbio.2020.111356>
- Nagasubramanian, K., Jones, S., Singh, A., Singh, A., Ganapathysubramanian, B., & Sarkar, S. (2018). Explaining hyperspectral imaging based plant disease identification: 3D CNN and saliency maps.
- Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., & Valenzuela, A. (2020). A review of convolutional neural network applied to fruit image processing. *Applied Sciences*, 10(10), 3443.
- Nturambirwe, J. F. I., & Opara, U. L. (2020). Machine learning applications to non-destructive defect detection in horticultural products. *Biosystems Engineering*, 189, 60–83. <https://doi.org/10.1016/j.biosystemseng.2019.11.011>

- OECD. (2018). *OECD fruit and vegetables scheme: Guidelines on objective tests to determine quality of fruits and vegetables, dry and dried produce*.
- Pan, L. Q., Zhang, Q., Zhang, W., Sun, Y., Hu, P. C., & Tu, K. (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry*, 192, 134–141. <https://doi.org/10.1016/j.foodchem.2015.06.106>
- Park, T. S., & Bae, Y. M. (2008). Analysis of capsaicinoids from hot red pepper powder by near-infrared spectroscopy. <https://doi.org/10.13031/2013.25077>.
- Parrag, V., Felföldi, J., Baranyai, L., Geösel, A., & Firtha, F. (2014). Early detection of cobweb disease infection on *Agaricus bisporus* sporocarps using hyperspectral imaging. *Acta Alimentaria*, 43(Suppl 1), 107–113. <https://doi.org/10.1556/AAlim.43.2014.Suppl.16>
- Pathmanaban, P., Gnanavel, B., & Anandan, S. S. (2019). Recent application of imaging techniques for fruit quality assessment. *Trends in Food Science & Technology*, 94, 32–42.
- Peng, T. O., Aizzat Zakaria, M., Fakhri, A., Nasir, A., Majeed, A. P. P. A., Young, T. C., Ng, L., & Yew, C. (2019). Autonomous tomato harvesting robotic system in greenhouses: Deep learning classification. *Mekatronika-Journal of Intelligent Manufacturing & Mechatronics*, 1, 80–86. <https://doi.org/10.15282/mekatronika.v1.1148>
- Permata, Y. C., Handayani, W., & Saputro, A. H. (2017). Carotenoid profile map system based on hyperspectral technique in banana (*Musa sp.*). In *2017 5th International Conference on Instrumentation, Control, and Automation (Ica)* (pp. 163–166).
- Polder, G., Van Der Heijden, G. W. A. M., Van Der Voet, H., & Young, I. T. (2004). Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry. *Postharvest Biology and Technology*, 34, 117–129. <https://doi.org/10.1016/j.postharvbio.2004.05.002>
- Potena, C., Nardi, D., & Pretto, A. (2017). Fast and accurate crop and weed identification with summarized train sets for precision agriculture (Vol. 531). https://doi.org/10.1007/978-3-319-48036-7_9.
- Rajkumar, P., Wang, N., Elmasry, G., Raghavan, G. S. V., & Garipey, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering*, 108(1), 194–200. <https://doi.org/10.1016/j.jfoodeng.2011.05.002>
- Riccioli, C., Perez-Marin, D., & Garrido-Varo, A. (2021). Optimizing spatial data reduction in hyperspectral imaging for the prediction of quality parameters in intact oranges. *Postharvest Biology and Technology*, 176. <https://doi.org/10.1016/j.postharvbio.2021.111504>
- Rivera, N. V., Gomez-Sanchis, J., Chanona-Perez, J., Carrasco, J. J., Millan-Giralolo, M., Lorente, D., Cubero, S., & Blasco, J. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosystems Engineering*, 122, 91–98. <https://doi.org/10.1016/j.biosystemseng.2014.03.009>
- Rojas-Moraleda, R., Valous, N. A., Gowen, A., Esquerre, C., Härtel, S., Salinas, L., & O'Donnell, C. (2016). A frame-based ANN for classification of hyperspectral images: Assessment of mechanical damage in mushrooms. *Neural Computing & Applications*, 28(S1), 969–981. <https://doi.org/10.1007/s00521-016-2376-7>
- van Roy, J., Keresztes, J. C., Wouters, N., De Ketelaere, B., & Saey, W. (2017). Measuring colour of vine tomatoes using hyperspectral imaging. *Postharvest Biology and Technology*, 129, 79–89. <https://doi.org/10.1016/j.postharvbio.2017.03.006>
- Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijtjaru, P., Mahayothee, B., & Muller, J. (2017). Prediction mapping of physicochemical properties in mango by hyperspectral imaging. *Biosystems Engineering*, 159, 109–120. <https://doi.org/10.1016/j.biosystemseng.2017.04.006>
- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., & McCool, C. (2016). DeepFruits: A fruit detection system using deep neural networks. *Sensors*, 16, 1222. <https://doi.org/10.3390/s16081222>
- Sánchez, M. T., Garrido-Varo, A., Guerrero, J. E., & Pérez-Marín, D. (2013). NIRS technology for fast authentication of green asparagus grown under organic and conventional production systems. *Postharvest Biology and Technology*, 85, 116–123. <https://doi.org/10.1016/j.postharvbio.2013.05.008>
- Saputro, A. H., Juansyah, S. D., & Handayani, W. (2018). Banana (*Musa sp.*) maturity prediction system based on chlorophyll content using visible-NIR imaging. In *2018 International Conference on Signals and Systems (Icsigsys)* (pp. 64–68).
- Sathya, R., & Abraham, A. (2013). Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence*, 2. <https://doi.org/10.14569/IJARAI.2013.020206>
- Schmilovitch, Z. e., Ignat, T., Alchanatis, V., Gatker, J., Ostrovsky, V., & Felföldi, J. (2014). Hyperspectral imaging of intact bell peppers. *Biosystems Engineering*, 117, 83–93. <https://doi.org/10.1016/j.biosystemseng.2013.07.003>
- Science, W. o (2021). Web of Science database search. Retrieved 06/15 from <https://apps.webofknowledge.com>.
- Signoroni, A., Savardi, M., Baronio, A., & Benini, S. (2019). Deep learning meets hyperspectral image analysis: A multidisciplinary review. *Journal of Imaging*, 5(5). <https://doi.org/10.3390/jimaging5050052>
- Simko, I., Hayes, R. J., & Furbank, R. T. (2016). Non-destructive phenotyping of lettuce plants in early stages of development with optical sensors. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01985>
- Siregar, S. T. W., Handayani, W., & Saputro, A. H. (2017). Bananas moisture content prediction system using visual-NIR imaging. In *2017 5th International Conference on Instrumentation, Control, and Automation (Ica)* (pp. 89–92).
- Siripatrawan, U., Makino, Y., Kawagoe, Y., & Oshita, S. (2011). Rapid detection of *Escherichia coli* contamination in packaged fresh spinach using hyperspectral imaging. *Talanta*, 85, 276–281. <https://doi.org/10.1016/j.talanta.2011.03.061>
- Slaughter, D. C., Giles, D. K., & Downey, D. (2008). Autonomous robotic weed control systems: A review. *Computers and Electronics in Agriculture*, 61(1), 63–78. <https://doi.org/10.1016/j.compag.2007.05.008>
- Song, Y., Glasbey, C. A., Horgan, G. W., Polder, G., Dieleman, J. A., & van der Heijden, G. (2014). Automatic fruit recognition and counting from multiple images [Article]. *Biosystems Engineering*, 118, 203–215. <https://doi.org/10.1016/j.biosystemseng.2013.12.008>
- Statista. (2021a). Fresh fruits: Worldwide. Retrieved 01/25 from <https://www.statista.com/outlook/40160100/100/fresh-fruits/worldwide#market-revenue>.
- Statista. (2021b). Vegetables: Worldwide. Retrieved 06/21 from <https://www.statista.com/outlook/cmo/food/vegetables/worldwide>.
- Sun, Y., Gu, X. Z., Sun, K., Hu, H. J., Xu, M., Wang, Z. J., Tu, K., & Pan, L. Q. (2017). Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches. *LWT-Food Science and Technology*, 75, 557–564. <https://doi.org/10.1016/j.lwt.2016.10.006>
- Sun, J., Künemeyer, R., McGlone, A., & Tomer, N. (2018). Optical properties of healthy and rotten onion flesh from 700 to 1000 nm. *Postharvest Biology and Technology*, 140, 1–10. <https://doi.org/10.1016/j.postharvbio.2018.02.006>
- Sun, Y., Wang, Y. H., Xiao, H., Gu, X. Z., Pan, L. Q., & Tu, K. (2017). Hyperspectral imaging detection of decayed honey peaches

- based on their chlorophyll content. *Food Chemistry*, 235, 194–202. <https://doi.org/10.1016/j.foodchem.2017.05.064>
- Susic, N., Zibrat, U., Sirca, S., Strajnar, P., Razinger, J., Knapic, M., Voncina, A., Urek, G., & Stare, B. G. (2018). Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging. *Sensors and Actuators B-Chemical*, 273, 842–852. <https://doi.org/10.1016/j.snb.2018.06.121>
- Taghizadeh, M., Gowen, A. A., & O'Donnell, C. P. (2011a). Comparison of hyperspectral imaging with conventional RGB imaging for quality evaluation of *Agaricus bisporus* mushrooms. *Biosystems Engineering*, 108(2), 191–194. <https://doi.org/10.1016/j.biosystemseng.2010.10.005>
- Taghizadeh, M., Gowen, A. A., & O'Donnell, C. P. (2011b). The potential of visible-near infrared hyperspectral imaging to discriminate between casing soil, enzymatic browning and undamaged tissue on mushroom (*Agaricus bisporus*) surfaces. *Computers and Electronics in Agriculture*, 77(1), 74–80. <https://doi.org/10.1016/j.compag.2011.03.010>
- Teena, M., Manickavasagan, A., Ravikanth, L., & Jayas, D. (2014). Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits. *Journal of Stored Products Research*, 59, 306–313. <https://doi.org/10.1016/j.jspr.2014.09.005>
- Tian, X., Li, J. B., Wang, Q. Y., Fan, S. X., Huang, W. Q., & Zhao, C. J. (2019). A multi-region combined model for non-destructive prediction of soluble solids content in apple, based on brightness grade segmentation of hyperspectral imaging. *Biosystems Engineering*, 183, 110–120. <https://doi.org/10.1016/j.biosystemseng.2019.04.012>
- Tripodi, P., Massa, D., Venezia, A., & Cardi, T. (2018). Sensing technologies for precision phenotyping in vegetable crops: Current status and future challenges. *Agronomy*, 8. <https://doi.org/10.3390/agronomy8040057>
- Utkin, L. V., Zaborovskii, V. S., & Popov, S. G. (2016). Detection of anomalous behavior in a robot system based on deep learning elements. *Automatic Control and Computer Sciences*, 50(8), 726–733. <https://doi.org/10.3103/S0146411616080319>
- Valero, D., & Serrano, M. (2010). *Postharvest biology and technology for preserving fruit quality*. CRC Press.
- Walsh, K. B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR 'point'spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biology and Technology*, 168, Article 111246.
- Wang, Z., Hu, M., & Zhai, G. (2018). Application of deep learning architectures for accurate and rapid detection of internal mechanical damage of blueberry using hyperspectral transmittance data. *Sensors*, 18(4), 1126.
- Wang, W., Li, C., & Gitaitis, R. D. (2014). Optical properties of healthy and diseased onion tissues in the visible and near-infrared spectral region. *Transactions of the ASABE*, 57(6), 1771–1782. <https://doi.org/10.13031/trans.57.10815>
- Wang, N.-N., Sun, D.-W., Yang, Y.-C., Pu, H., & Zhu, Z. (2016). Recent advances in the application of hyperspectral imaging for evaluating fruit quality. *Food Analytical Methods*, 9(1), 178–191.
- Weersink, A., Fraser, E., Pannell, D., Duncan, E., & Rotz, S. (2018). Opportunities and challenges for big data in agricultural and environmental analysis. *Annual Review of Resource Economics*, 10(1), 19–37. <https://doi.org/10.1146/annurev-resource-100516-053654>
- Wei, X., He, J., Zheng, S., & Ye, D. (2020). Modeling for SSC and firmness detection of persimmon based on NIR hyperspectral imaging by sample partitioning and variables selection. *Infrared Physics & Technology*, 105, Article 103099. <https://doi.org/10.1016/j.infrared.2019.103099>
- 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. <https://doi.org/10.1016/j.compag.2018.10.021>
- Yuan, R. R., Liu, G. S., He, J. G., Wan, G. L., Fan, N. Y., Li, Y., & Sun, Y. R. (2021). Classification of Lingwu long jujube internal bruise over time based on visible near-infrared hyperspectral imaging combined with partial least squares-discriminant analysis. *Computers and Electronics in Agriculture*, 182. <https://doi.org/10.1016/j.compag.2021.106043>
- Yu, X., Lu, H., & Wu, D. (2018). Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging. *Postharvest Biology and Technology*, 141, 39–49. <https://doi.org/10.1016/j.postharvbio.2018.02.013>
- Zhang, B. H., Li, J. B., Fan, S. X., Huang, W. Q., Zhao, C. J., Liu, C. L., & Huang, D. F. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). *Computers and Electronics in Agriculture*, 114, 14–24. <https://doi.org/10.1016/j.compag.2015.03.015>
- Zhang, H., Zhan, B., Pan, F., & Luo, W. (2020). Determination of soluble solids content in oranges using visible and near infrared full transmittance hyperspectral imaging with comparative analysis of models. *Postharvest Biology and Technology*, 163, Article 111148.
- Zhou, L., Zhang, C., Liu, F., Qiu, Z. J., & He, Y. (2019). Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety*, 18(6), 1793–1811. <https://doi.org/10.1111/1541-4337.12492>
- Zhu, Q., He, C., Lu, R., Mendoza, F., & Cen, H. (2015). Ripeness evaluation of 'Sun Bright' tomato using optical absorption and scattering properties. *Postharvest Biology and Technology*, 103, 27–34. <https://doi.org/10.1016/j.postharvbio.2015.02.007>
- Zhu, X. L., & Li, G. H. (2019). Rapid detection and visualization of slight bruise on apples using hyperspectral imaging. *International Journal of Food Properties*, 22(1), 1709–1719. <https://doi.org/10.1080/10942912.2019.1669638>