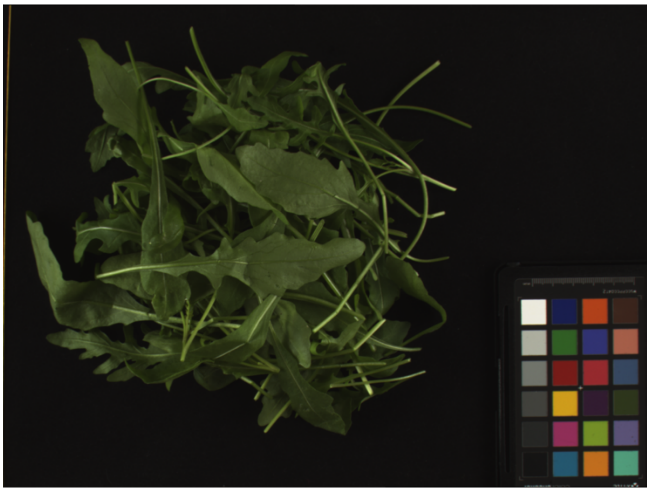
UNIVERSITÀ DI FOGGIA

Dipartimento di Scienze Agrarie, Alimenti, Risorse Naturali e Ingegneria (DAFNE)

Doctoral Thesis in
Management of Innovation in the Agricultural and Food Systems of the Mediterranean Region
– XXXV cycle –

COMPUTER VISION SYSTEM FOR NON-DESTRUCTIVE AND CONTACTLESS EVALUATION OF QUALITY TRAITS IN FRESH ROCKET LEAVES (Diplotaxis tenuifolia L.)




Candidate:
Michela Palumbo

Tutor:
Prof. Giancarlo Colelli

Co-tutor:
Dr. Maria Cefola

UNIVERSITÀ DI FOGGIA

Dipartimento di Scienze Agrarie, Alimenti, Risorse Naturali e Ingegneria (DAFNE)

Doctoral Thesis in
Management of Innovation in the Agricultural and Food Systems
of the Mediterranean Region
– XXXV cycle –

COMPUTER VISION SYSTEM FOR NON-DESTRUCTIVE AND CONTACTLESS EVALUATION OF QUALITY TRAITS IN FRESH ROCKET LEAVES (Diplotaxis tenuifolia **L.***)*

Candidate:	Tutor:	Co-tutor:
Michela Palumbo	*Prof. Giancarlo Colelli*	*Dr. Maria Cefola*

This work was part of the project Prin 2017 “SUS&LOW-Sustaining low-impact practices in horticulture through non-destructive approach to provide more information on fresh produce history and quality” (grant number: 201785Z5H9), funded by the Italian Ministry of Education University.

The research activity was carried out by the University of Foggia (project leader) and two institutes of National Research Council of Italy (CNR): the Institute of Science of Food Production (ISPA) and the Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA).

SUSTAINING ***LOW***-IMPACT PRACTICES IN HORTICULTURE THROUGH NON-DESTRUCTIVE APPROACH TO PROVIDE MORE INFORMATION ON FRESH PRODUCE HISTORY & QUALITY

UNIVERSITÀ DI FOGGIA

Dipartimento di Scienze Agrarie, Alimenti, Risorse Naturali e Ingegneria (DAFNE)

Doctoral program in Management of Innovation in the Agricultural and Food Systems of the Mediterranean Region
– XXXV cycle –

Doctoral thesis on "**Computer vision system for non-destructive and contactless evaluation of quality traits in fresh rocket leaves (*Diplotaxis tenuifolia* L.)**" discussed at Università di Foggia

Candidate:

Michela Palumbo

Tutor:

Prof. Giancarlo Colelli (*Università di Foggia, Italy*)

Co-tutor:

Dr. Maria Cefola (*Istituto di Scienze delle Produzioni Alimentari, ISPA-CNR, Italy*)

Committee members:

Prof. Roberto Beghi (*Università degli Studi di Milano, Italy*)

Prof. Michele Mattetti (*Alma Mater Studiorum Università di Bologna, Italy*)

Prof. Roberto Romaniello (*Università di Foggia, Italy*)

To Gaetano and our children, with sincere love

"…Affrontare il male senza mormorare,
Con pazienza e gioia saper sopportare.
Aver vinto su te stesso,
Sappi, questa è la letizia."

Angelo Branduardi
'La predica della perfetta letizia'
(cfr. San Francesco d'Assisi)

COMPUTER VISION SYSTEM FOR NON-DESTRUCTIVE AND CONTACTLESS EVALUATION OF QUALITY TRAITS IN FRESH ROCKET LEAVES (*Diplotaxis tenuifolia* L.)

Contents

Reference: Palumbo, M., Cefola, M., Pace, B., Attolico, G., & Colelli, G. (2023). Computer vision system based on conventional imaging for non-destructively evaluating quality attributes in fresh and packaged fruit and vegetables. *Postharvest Biology and Technology*, 200, 112332.

Reference: Palumbo, M., Pace, B., Cefola, M., Montesano, F. F., Serio, F., Colelli, G., & Attolico, G. (2021). Self-configuring CVS to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. *Agronomy*, 11(7), 1353.

EXTENDED ABSTRACT

The demand for healthy, safe, high-quality and ready-to-eat fresh produce has increased in the last few years with the change in the lifestyle of modern consumers. Food quality is related to a determined maturity stage, where the composition, or the combination of physical and chemical attributes, has the maximum acceptance by consumers. Some of these quality traits can be perceived with the use of senses, while others, such as nutritional value, chemical and microbial safety, and degree of sustainability of the cultivation approach, cannot be judged directly by consumers. The acceptability of foods could be increased by providing this information to consumers.

The quality of fruit and vegetables is usually assessed by sensorial and subjective determination, using scoring rating scales. Moreover, conventional destructive methods are used to support the sensory evaluation by the assessment of desired chemical and physical attributes related to the quality of the product. Even if they are still widely used, analytical and destructive techniques are time consuming and expensive, adversely impact the environment, require sophisticated equipment and need careful sample preparations. Finally, they are not suitable for in-line applications where speed, accuracy and sustainability are required.

Recent researches have been focused on the use of contactless, non-destructive, rapid and accurate techniques, as well as non-polluting, for quality assessment of fruit and vegetables with the aim to predict their sensory and desired compositional traits in an objective and consistent way. Non-destructive (ND) techniques do not need sample preparation once the final model is developed making the prediction process rapid and objective. Additionally, the growing awareness of modern consumers toward the economic, social and environmental sustainability of production processes has prompted many researchers to develop ND tools for the discrimination of cultivation approach, in order to better support the added value of the products.

Nevertheless, it is important to know that, even if these non-invasive methods provide more significant advantages than analytical and destructive analysis, they cannot completely replace the conventional methods, as, at the moment, they can only be a support in saving time, also reducing impact on the environment.

The thesis focused on the analysis of non-destructive technologies available for the control quality of agri-food products, along the whole supply chain. In particular, the thesis concerns the application of computer vision system to evaluate the quality of fresh rocket leaves. The thesis is structured in three parts (introduction, experimental applications and conclusions) and in 5 chapters, the first and second focused on non-destructive technologies and in particular on computer vision systems for monitoring the quality of agri-food products, respectively. The third, quarter, and fifth chapters aim to assess the rocket leaves based on the estimation of quality aspects, considering different aspects: (i) the variability due to the different agricultural practices, (ii) the senescence of packed and unpacked products, and (iii) development and exploitation of the advantages of new models simpler than the machine learning used in the previous experiments.

The research work of this doctoral thesis was carried out by the University of Foggia, the Institute of Science of Food Production (ISPA) and the Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA) of National Research Council (CNR). It was conducted within the Project SUS&LOW (Sustaining Low-impact Practices in Horticulture through Non-destructive Approach to Provide More Information on Fresh Produce History & Quality), funded by MUR-PRIN 2017, and aimed at sustaining quality of production and of the environment using low input agricultural practices (LIP) and ND quality evaluation. According to the main focus of SUS&LOW, ND evaluation could (i) provide evidence about the inner quality of production, (ii) be considered an additional tool for the discrimination of fresh produce obtained with LIP, and (iii) enable for the estimation of quality and shelf-life of fresh produce also when in the plastic package. This information may be used in order to design strategies to ensure better marketing conditions for fresh fruit and vegetables obtained by LIP. Details about the objective, main results, and general conclusions of activities carried out within SUS&LOW are published in Advances in Horticultural Science and reported in Chapter 1 of this Thesis.

Among innovative and ND methodologies commonly studied and used in quality assessment of fruit and vegetables, image analysis by computer vision systems (CVS) represents an innovative and contactless technology suitable for in-line grading. Different types of CVS have been developed, based on conventional, multispectral and hyperspectral imaging.

The solutions based on conventional imaging (CVS-CI) use RGB colour cameras that are sensible to the visible wavelengths of the electromagnetic spectrum. Their simplicity, flexibility and cost allow for a continuous monitoring of quality along the entire supply chain, from harvest to final consumer, thus reducing wastes and losses. They typically acquire images using a setup composed of the combination of a digital camera, an illumination system, and a personal computer that extracts classification features and builds appropriate models, using statistical methods or machine learning approaches. Besides the automatic external inspection of fresh produce based on morphological traits (size, shape, defects and colour changes), CVS-CI has also shown to be effective in determining some internal characteristics related to the nutritional quality. Moreover, the integration of machine learning methodologies in these systems significantly increases their efficiency and simplifies their design, development and deployment.

CVS-CI technologies may evaluate only internal quality traits strictly related to external visible changes as physical alteration or colour changes (e.g., soluble solids content, chlorophyll content, ammonia content, enzymatic activity, phenols content, antioxidant activity). These limitations have prompted researchers to improve the image analysis performed by CVS-CI through the studying and developing of more advanced solutions. A detailed description of CVS-CI technology, all the advantages and limitations in the quality assessment of fresh and packaged fruit and vegetables and future perspectives have been discussed in Chapter 2 of this Thesis. The content of this chapter was published on Postharvest Biology and Technology including the current state of art about the quality evaluation of fruit and vegetables by CVS-CI, highlighting its potentiality for a continuous and efficient monitoring of quality along the entire distribution chain.

The research work described in this Thesis is aimed to develop and validate predictive models based on the use of CVS-CI for the assessment of the quality level (QL) and the main quality parameters of fresh rocket leaves (*Diplotaxis tenuifolia* L.) with or without packaging. Rocket is a common name which denotes many species of green leaves belonging to the *Brassicaceae* family and it is distinguished by a pungent smell, a bitter taste and a wide range of phytonutrients, such as provitamin A, vitamin C, flavonoids, glucosinolates, fibers, potassium and sulphur. It is widely consumed raw in the Mediterranean countries alone or in mixed leafy salads and, among the variability of rocket species, *Eruca sativa* L. and *Diplotaxis tenuifolia* L. are the ones commonly present in the market. Because of the increase of rocket consumption in the recent years, this vegetable is mainly marketed as ready-to-eat product. The storage conditions (temperature, atmosphere and packaging) and processing operations may limit its shelf-life accelerating some typical degradation processes such as wilting, yellowing, loss of nutritional properties and of sensorial attributes. The possibility of monitoring the shelf-life of rocket leaves from farm to fork, identifying and predicting the QL of the product at each step of the supply chain, could reduce wastes and enhance the sustainability of the production process.

The proposed CVS was able to automatically select, without human intervention, the most relevant colour traits strictly related to the quality of rocket leaves using the Random Forest as machine learning model ensuring a consistent prediction of the product shelf-life.

During the first study, CVS was applied to fresh rocket leaves obtained by LIP, to objectively assess its QL during the storage at 10 °C according to a 5 to 1 rating scale and to discriminate the fertilization and irrigation management applied during the cultivation. Three colour correction techniques (i.e. white balance, linear correction, and polynomial correction) were evaluated and compared in terms of classification performance, in order to identify the best solution for providing consistent colour measurements. Among them, linear colour correction proved to be the best trade-off between efficacy and efficiency in making consistent colour measurements. Promising results showed an accuracy of 95 % in the QL assessment and of about 65-70 % in the discrimination of the cultivation approach. This research activity is published on a special issue of Agronomy and reported in Chapter 3 of the Thesis.

In the second study, five experiments were conducted to validate the CVS in estimating internal quality traits (chlorophyll and ammonia content) related to senescence of rocket leaves, in packaged and unpackaged samples. The same CVS, using its machine learning components, was able to build effective models for both the classification (visual quality level assignment) and the regression (estimation of senescence indicators such as chlorophyll and ammonia contents) problems by just changing the training data. The results, published on Postharvest Biology and Technology and reported in Chapter 4, showed similar achievements, with a negligible performance loss, on packaged (Pearson's linear correlation coefficient of 0.84 for chlorophyll and 0.91 for ammonia) and unpackaged products (0.86 for chlorophyll and 0.92 for ammonia). Moreover, three PLS models were compared to estimate the QL of rocket leaves using the chlorophyll contents obtained by destructive methods (Model I), by CVS on packaged (Model II) and by CVS on unpackaged products (Model III) as predictors. Those estimated non-destructively and contactless by the CVS (Model II and Model III) provided better performances in the QL prediction (R^2_v of 0.77, 0.80, respectively) than the ones measured by destructive analysis (R^2_v of 0.70).

The third study of this Thesis was the exploring a clustering approach to identify relevant and representative colour cues and to construct simpler algorithms for the prediction of the QL of rocket leaves than the ones obtained by the Random Forest model.

Machine learning techniques may have significant computational costs and often produce models not easily understandable by humans. An interesting area of research consists in exploiting the advantages of learning while keeping the solutions simple, fast and interpretable by humans. The research paper, submitted for publication in Journal of Food Engineering and reported in Chapter 5, takes new steps in this direction, based on results obtained in the previous experiments. By analysing the Random Forest model already used to classify the QL and to estimate chlorophyll and ammonia contents in rocket leaves, new methods were proposed to (i) identify relevant clusters of colours that are informative about the properties of the product at hand, (ii) further select the clusters more significant to estimate the desired properties, and, (iii) describe shape and size of regions of the *ab*-plane in the CIELab colour representation corresponding to the clusters of interest. Comparing the results obtained in the previous experiments, these findings provided objective bases for the design of different computational schemes with different execution times enabling the best trade-off between efficacy and efficiency. In detail, two of the considered methods (M3b-C9-P and M1-C5-R) provided good prediction of chlorophyll and ammonia contents, assessing the quality of rocket leaves in an objective and robust way. Additionally, the two methods have computational time of 3 and 1 ms (respectively for M3b-C9-P and M1-C5-R) much shorter than the time required by the Random Forest model (not less than 20 ms).

Overall, results reported in this doctoral thesis could have a significant impact on advanced applications of the traditional vision systems commonly used for the inspection of fruit and vegetables. In particular, the ND and contactless CVS applied on fresh rocket leaves represents a valid alternative to destructive, expensive and time-consuming analyses in the laboratory and can be effectively and extensively used along the whole supply chain, even on packaged leaves which cannot be analysed by traditional tools.

PART I: INTRODUCTION

Chapter 1

SUSTAINING LOW-IMPACT PRACTICES IN HORTICULTURE THROUGH NON-DESTRUCTIVE APPROACH TO PROVIDE MORE INFORMATION ON FRESH PRODUCE HISTORY & QUALITY: THE SUS&LOW PROJECT

Maria Luisa Amodio[1], Giovanni Attolico[2], Lucia Bonelli[3], Maria Cefola[4], Hassan Fazayeli[1], Francesco Fabiano Montesano[3], Bernardo Pace[4], Michela Palumbo[1,4], Francesco Serio[3], Antonio Stasi[1], Giancarlo Colelli[1]

[1] Department of Agriculture, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Via Napoli 25, 71122 Foggia
[2] Institute on Intelligent Industrial Systems and Technologies for Advanced Manufacturing, CNR, Via G. Amendola, 122/O, 70126 Bari
[3] Institute of Sciences of Food Production, CNR, Via G. Amendola, 122/O, 70126 Bari
[4] Institute of Sciences of Food Production, CNR, c/o CS-DAT, Via Michele Protano, 71121 Foggia

ABSTRACT

The general aim of the PRIN project SUS&LOW is to increase the sustainability of fresh produce by testing and implementing low-input agricultural practices (LIP) with positive impact on product quality with the support of non-destructive (ND) tools for real-time quality assessment and for product discrimination. Additionally, new marketing strategies are generated to better support the added value of the products and to satisfy the final consumers' preferences. The SUS&LOW project consists of three work packages (WP) and the adopted methodology used two model crops: rocket salad and tomato. The WP1, focused on the reduction of agricultural inputs, showed that sensor-based fertigation management might improve sustainability of soilless cultivation. Results coming from WP2, aimed to the evaluation of ND techniques, outlined the high potentiality of hyperspectral imaging (HSI) and Fourier transformed-near infrared (FT-NIR) techniques for the authentication of sustainable growing methods. Moreover, project activities' proved computer vision system (CVS) as an effective tool for evaluating the product quality also through the bag. The WP3, dealing with marketing strategies, indicated a positive approach of consumers compared to LIP products certified through a visual storytelling platform.

Keywords: sustainability, quality, non-destructive assessment, shelf-life, marketing strategies.

1. INTRODUCTION

Production of vegetable crops under controlled environments (i.e. greenhouses) has expanded considerably over recent decades in Mediterranean areas (FAO, 2013). Initially, research efforts and the related introduction of technical innovations focused on high-quality, healthy products. However, concern with environmentally-sustainable production has risen in the last decade as industrial greenhouse crops are usually seen as entailing high environmental impact (Torrellas et al., 2012). On the other hand, there is also plenty of evidence that greenhouse vegetable production may decrease the environmental impact compared to the field cultivation (Stanghellini, 2014).

The efficient use of resources (water and fertilizers), in irrigated greenhouse agriculture, is a promising and increasingly adopted strategy to achieve better crop performance, improved nutritional and sensorial quality (Montesano et al., 2015; Montesano et al., 2018). With respect to traditional systems, soilless cultivation and, particularly, closed-cycle with recycling of nutrient solution (NS) produce a number of benefits, including the possibility to standardize the production process, to improve plant growth and yield, and to obtain higher efficiency in water and nutrients use. In addition, it is also possible to modulate the regulation of the secondary metabolism of plants through an optimal control of the nutrient solution composition, or by imposing controlled stresses, or through biofortification treatments, generally leading to an improvement in the nutritional value of products (Rouphael et al., 2018; Renna et al., 2022). Innovative technologies based on the use of sensor networks for fertigation management may considerably reduce water and fertilizers consumption and increase the overall use efficiency of those inputs, and may lead to qualitative and quantitative improvements while preventing both under- and over-irrigation.

The most used instrumental techniques to measure quality attributes of fruit and vegetables are destructive and involve a considerable amount of manual work, primarily due to sample preparation. In addition, most of these analytical techniques are time consuming and sometimes may require sophisticated equipment. Finally, they can be performed only on a limited number of specimens (samples) and therefore their statistical relevance may be limited (Amodio et al., 2017a). Research has been focused on developing non-contact, rapid, environmental-friendly, and accurate methods for non-invasive evaluation of quality in fruit and vegetables. Nowadays, there are a few emerging non-destructive analytical instruments and approaches for this task, including spectroscopy, hyperspectral imaging, and computer vision (Liu et al., 2017).

Near infrared spectroscopy has gained wide attention in the food sector due to its capacity of providing fingerprints of different products on the base of the interaction between their molecular structure and the incident light (Workman and Shenk, 2004) which is the result of different pre-harvest factors that also affect the final composition and quality. The feasibility of NIRS-based analysis to evaluate quality attributes of fresh fruits for commercial application have been reported by numerous authors (Arendse et al., 2018; Palumbo et al., 2022a).

Hyperspectral imaging (HSI) combines the principles of spectroscopy and conventional imaging or computer vision. It is mainly used for internal bruise and defect detection in fruit and vegetables (Ariana et al., 2010; Babellahi et al., 2020; Tsouvaltzis et al. 2020) but also to predict the internal composition (Piazzolla et al., 2013 and 2017; Yang et al., 2015; Liu et al., 2017). Amodio et al. (2017a) showed the potentiality of hyperspectral imaging in the Vis-NIR spectral range to predict internal content of soluble solids, phenols, and antioxidant activity of fennel heads. In addition, this technique provided important information about the maturity of fennel heads which may be used to determine the optimal harvest time. Some studies successfully applied these methods for the discrimination of production origin and agricultural practices, as revised in Amodio et al. (2020). NIR and HIS were in fact used for the classification of apples (Guo et al., 2013), persimmon (Khanmohammadi et al., 2014), and arabica coffee (Bona et al., 2017) from different origins. As for production systems (Sánchez et al., 2013) investigated the potentiality of NIRS technologies to discriminate green asparagus grown under organic and conventional methods. More recently, Amodio et al. (2017b) successfully discriminated conventionally and organically grown strawberries, being also able to identify two different types of organic production systems applied to the same genetic material on the same site, soil, unheated tunnel.

All these studies have suggested multispectral and hyperspectral systems as valid tools to evaluate quality of different agricultural products and, more interestingly, as tools for product authentication.

Finally, Computer Vision Systems (CVS) may be applied to extend quality prediction and discrimination along the whole supply chain from harvesting up to consumers. CVS combine mechanics, optical instrumentation, electromagnetic sensing, and digital image processing technology (Patel et al., 2012). Recently, CVSs have been used to assess quality and marketability of tomatoes (Arias et al., 2000), artichokes (Amodio et al., 2011), fresh-cut nectarines (Pace et al., 2011), fresh-cut lettuce (Pace et al., 2014), fresh-cut radicchio (Pace et al., 2015), and rocket leaves (Cavallo et al., 2017). Moreover, they have been applied for the prediction of internal quality of colored carrots (Pace et al., 2013). Even more interesting is the application of these systems during the post-packaging phase and along the whole distribution chain. Despite the relevance of quality evaluation of packaged products, few investigations were reported in literature. Multi-spectral reflective image analysis has been applied to monitor the evolution and spoilage of leafy spinach covered by plastic materials (Lara et al., 2013); more recently, Cavallo et al. (2018) have proposed an application of image analysis by CVS for non-destructive and contactless evaluation of quality of packaged fresh-cut lettuce. Therefore, the interest of investigating the application of CVS to detect quality and shelf-life of packaged products.

Finally, the possibility of using non-destructive technique for increasing the information on product history (e.g. growing location and agricultural practices) may be considered as baseline to develop marketing tools to promote the diffusion of sustainable production system. Cost barrier is an obstacle for choosing low input products instead of the conventional, even if environment is mentioned as a strong commitment (Krystallis and Chryssohoidis, 2005). Therefore, the knowledge about consumer preferences for the adoption of LIP is still matter of debate.

The general aim of the project is to increase the amount of sustainably-produced food by testing and implementing low-input agricultural practices with positive impact on product quality with the support of non-destructive tools for real-time quality assessment and product discrimination, which may inspire new marketing strategies to better support the added value of the products and increase incomes of potential users.

2. PROJECT ACTIVITIES AND MAIN RESULTS

The SUS&LOW project structure consists of three work packages (WP). WP1 focused on research activities aimed to reduce agricultural inputs (water and fertilizers) in greenhouse cultivation, chosen as a strategic high-value sector for Mediterranean agriculture. This WP was also in charge of making available to the project team vegetables products (rocket and tomato) different for the level of sustainability characterizing the cropping system adopted, to be used in other WPs for the related investigations. Then, WP2 was aimed to the quality assessment and to the implementation of new tools to acquire information about quality and history of fresh produce obtained with LIP (WP1). Non-destructive methods (including NIR, hyperspectral imaging and image analysis by CVS) have been used for food authentication, showing interesting and promising results. Finally, WP3 realized ad hoc survey to analyse the consumer behaviour with respect to the possibility of purchasing fruit and vegetables LIP certified (WP1) and identified by ND technologies (WP2) with the aim to implement adequate marketing strategies. In this section, an overview of the research strategies and approaches adopted in the three WPs is provided. The main results are reported and discussed.

2.1. WP1: quality crops through low-impact practices

Based on the overall project structure, this WP was focused on soilless cultivation, since it has the potentiality to achieve extremely high water and fertilizers use efficiency, beside high yield and quality, in intensive cropping systems. However, the adoption of free-drain open cycle with empiric fertigation schedule management operated by timers (the predominant case in Mediterranean area), may compromise the sustainability of soilless culture. Therefore, the adoption of strategies aimed to rational use of water and fertilizers and excess leaching prevention is a key-factor for increased sustainability and reduced environmental impact of soilless culture (Massa et al., 2020). In this context, substrate moisture/EC (electrical conductivity) sensor-based irrigation is a promising and increasingly adopted strategy to reduce water and fertilizers consumption and losses, and to improve the overall crop performance, product quality and production process sustainability in soilless greenhouse cultivation (Palumbo et al., 2021a).

Several experiments were carried out at the Experimental Farm La Noria (Mola di Bari, BA) of the CNR-ISPA (Bari), with the common approach to compare treatments providing traditionally adopted empirical fertigation management techniques with treatments in which advanced sensor-based fertigation management was implemented. The main results of selected experiments carried out during the project are reported hereafter.

The research activities focused on two model species [rocket salad (*Diplotaxis tenuifolia* L.) and tomato (*Solanum lycopersicum* L.)] selected for their relevance in Mediterranean greenhouse vegetable production. In particular, rocket is reported as an emerging leaf vegetable which cultivation is widespread and in further expansion (Schiattone et al., 2017), while tomato is the most important greenhouse crop grown in soilless cultivation systems (Montesano et al., 2015).

A study was carried out to test two irrigation scheduling approaches (timer- or sensor-based) and two fertilization levels (high or low, with reference to the standard dosage range recommended for the specific fertilizers used) of open-cycle soilless rocket in Mediterranean autumn-winter unheated greenhouse conditions (Montesano et al., 2021). Rocket plants (cv. Dallas, Isi Sementi) were grown in a peat:perlite (3:1) mixture in 4.5 L plastic pots. Four treatments were compared: timer with high or low fertilization (T-HF, T-LF), and sensor-based with high or low fertilization (S-HF, S-LF). In timer-based treatments, irrigation schedule was periodically adjusted based on leaching fraction measurements (≈35 % was set as a target, according to common practice). In sensor-based treatments, on-demand irrigation was operated based on substrate EC/temperature/moisture sensors (GS3, Decagon Devices). These were connected to a CR1000 datalogger programmed to automatically open irrigation valves and supply water enough to constantly maintain volumetric water content to a pre-defined set-point (0.35 m^3 m^{-3}, close to maximum water holding capacity), with no leaching. Slow-release fertilizers (Osmocote Exact and CalMag, ICL) were mixed with the substrate at high (3.75 and 1 g L^{-1}, respectively) or low dosage (2.25 and 0.6 g L^{-1}). Yield, quality, water use and substrate parameters trends were evaluated. Sensors improved water use efficiency compared to timer (34.4 vs 21.4 g FW L^{-1}, on average) (Figure 1) matching water supply with plant needs, and preventing leaching (Figure 1) (no interactive effects of fertilization treatments were observed on those parameters).

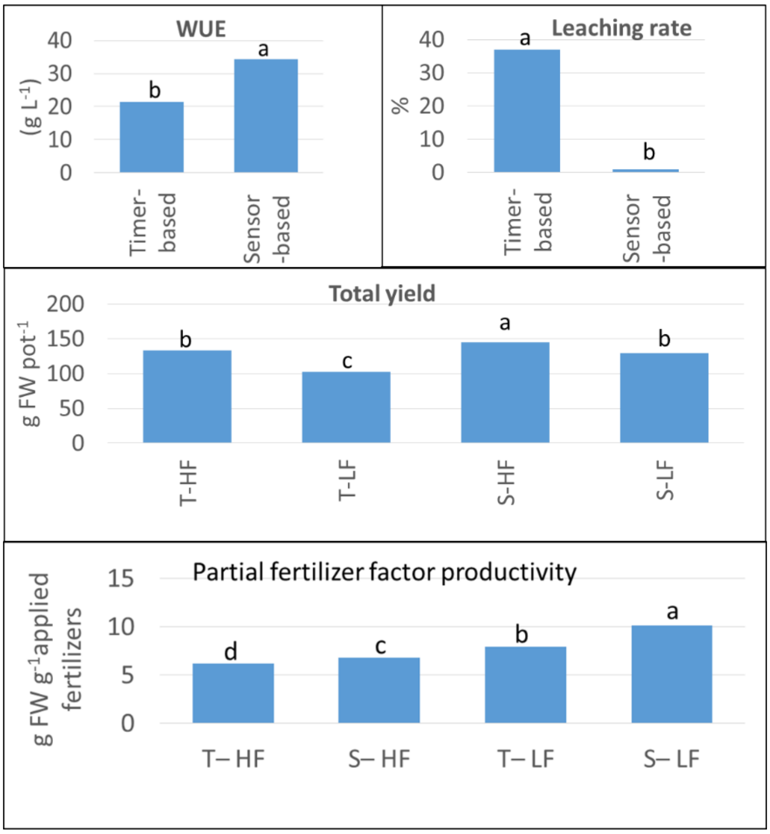


Figure 1. Water use efficiency (WUE), leaching rate, total yield, and partial fertilizer factor productivity of rocket (*Diplotaxis tenuifolia*) grown in open free drain soilless system with timer- (T) or sensor-based (S) irrigation management, and subjected to high (HF) or low (LF) fertilization rate.

Sensor-based irrigation also provided the best plant growth conditions, with interesting interactive effects with fertilization rate. In particular, the highest and the lowest cumulative (three harvests) yield values were obtained in S-HF and T-LF respectively (144.8 and 102.2 g FW pot^{-1}), while similar values were observed in S-LF and T-HF (131.4 g FW pot^{-1}, on average)

(Figure 1). The partial fertilizer factor productivity (g product fresh weight / g fertilizers applied) was higher at low dosage, and, with the same dosage, when the sensors were used (Figure 1). After each harvest time the fresh-cut rocket leaves were immediately transported in refrigerate conditions to the postharvest laboratory (see WP2 section below) (Palumbo et al., 2021b).

In another set of studies, we aimed to apply approaches for the sustainable fertigation management of soilless tomato (semi-closed cycle recirculation; sensor-based nutrient solution supply management) in comparison with a traditional open cycle free-drain nutrient solution management providing the use timer for fertigation schedule. Experiments were conducted with different tomato types (cherry – cv. Carminio, Seminis-Bayer, and intermediate type – cv. Mose, Syngenta), and in different environmental conditions typical of Mediterranean areas (including the use of brackish water for nutrient solution preparation). In general, both approaches (semiclosed-cycle cultivation and open cycle with sensor-based fertigation management) reduced the environmental impact of the production process (reduced water/fertilizers usage; less nutrient solution released into the environment, increased water use efficiency) and positively affected tomato quality traits, compared to empirically management open-free drain cultivation.

2.2. WP2: non-destructive discrimination for low-impact practices and non-destructive quality assessment

2.2.1 NIR spectroscopy and Hyperspectral imaging

In this WP, the objective of the tasks was to assess the potentiality of Fourier transformed-near infrared (FT-NIR) spectrometry and hyperspectral imaging (HSI) to discriminate tomatoes and rocket leaves produced with different level of input as described in WP1, taking also into account the degree of efficiency in water and fertilizers used efficiency (WUE and FUE indexes). A hyperspectral line-scan scanner (Version 1.4, DV srl, Padova, Italy) equipped with two spectrographs, one in the Vis-NIR range, and the second in the NIR range, was used to obtain the HS images. The Vis-NIR spectrograph (400-1000 nm) has a spatial resolution of 1000 × 2000 pixels with a spectral resolution of 5 nm and was connected to a CCD camera. As for the NIR spectrograph (900-1700 nm), the spatial resolution was 600 × 320 pixels with a spectral resolution of 5 nm; and a CMOS (Specim Spectral Imaging Ltd., Oulu, Finland) with 50 frames per second equipped with C-mount lenses was used. As for FT-NIR spectrometry an MPA Multi-Purpose (FT-NIR Analyzer, Bruker Optics, Ettlingen, Germany), was used during spectral acquisition over the range of 800–2777 nm (sphere macrosample resolution 1.71 nm, scanner velocity 10 kHz, sample scan time 64 scans, background scan time 64 scans). After image processing and spectra extraction for the HSI, all spectra belonging to HSI and FT-NIR were tested in discrimination using the agronomic treatments as discriminant classes and Partial Least Squares-Discriminant Analysis (PLS-DA) as classification technique. As for rocket leaves, PLS-DA was conducted with the 4 classes (T-HF; T-LF; S-HF, S-LF) described in the paragraph related to WP1, using 70 percent of samples for calibration purpose and the remaining 30 % for the external validation. The model performance was evaluated based on the accuracy, which is an average of the sensitivity calculated over the various classes, and gives an overall idea of the goodness of the classification. Results indicated HSI as a promising technique for the discrimination of rocket produced with different cultural techniques, with an accuracy of classification in the prediction phase of 97.2 % in Vis-NIR and 99.5 % in NIR range. In Figure 2, the results of the discrimination models can be observed.

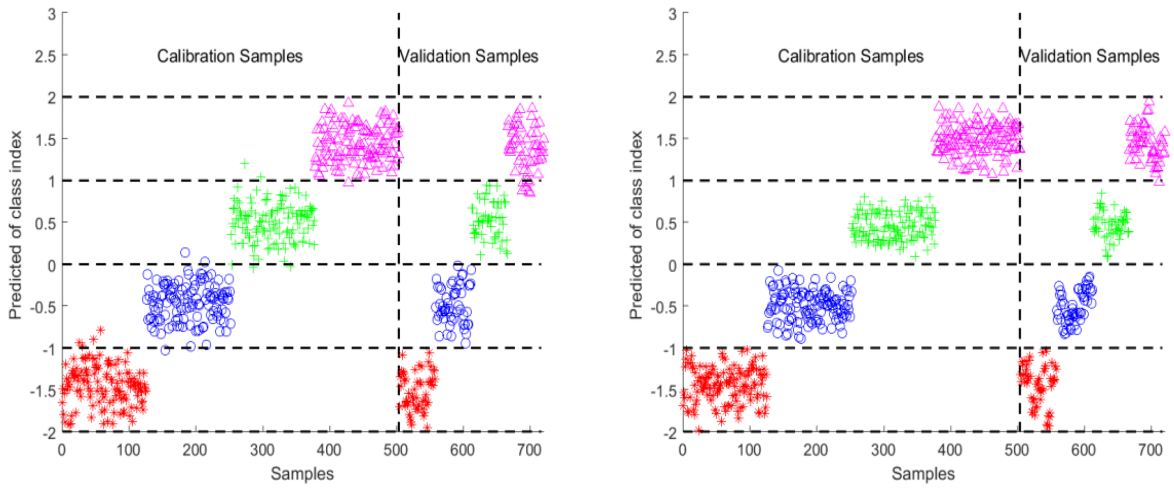


Figure 2. Estimated class index values in the calibration and in the prediction process for the classification based on PLS-DA modes in VNIR range (left) and NIR range (right).

Regarding tomato, where 2 experiments with 2 different varieties were conducted (WP1), for each trial a first PLS-DA was aimed to discriminate the three treatments of cultivation and a second discrimination was performed for different levels of WUE and FUE. According to the efficiency of use of water and fertilizers we could individuate 2 levels (high and low) in each experiment and 3 levels (High, medium, and low) merging the data of both experiments. Therefore, a PLS-DA with 3 levels of WUE (and FUE) was also generated with the full dataset. Among the different non-destructive techniques, FT-NIR and HIS in the VIS-NIR range gave comparable performances in discriminating tomato according to cultural practices and different use of sources. Discrimination for WUE for each variety improved the classification results, respect to the individual treatments, but the highest accuracy was obtained when the discrimination was based on 3 levels of WUE merging the 2 datasets, reaching 92.1 %. In literature there are no studies aimed to discriminate crops for WUE or FUE, while we may find the application of HSI for the classification of water-stressed plants, as for the case of tomatoes (Rinaldi et al., 2015). In comparison to this study, reporting a mean accuracy of around 77 % for discrimination of the two differently irrigated areas, our findings showed higher accuracy, exploring new area of the application for these techniques.

2.2.2. Application of CVS for non-destructive quality evaluation on packaged products

A research activity was carried out to develop and validate an innovative CVS integrating a Random Forest model for classification: this model automatically selects from the image the most relevant colour features for the task of interest. The developed CVS was applied to digital images of fresh-cut rocket leaves cultivated with LIP (WP1) to objectively estimate the evolution of their quality levels (QL) during storage and to discriminate the cultivation approach applied on field. At harvest, rocket leaves were stored at 10 °C in open polypropylene (PP) bags for a number of days required to reach the lowest QL, according to the rating scale from 5 (very good) to 1 (very poor), as reported in Figure 3.

Figure 3. Changes in the sensory quality level (QL) of fresh-cut rocket leaves during the storage at 10 °C according to the 5 to 1 rating scale reported by Palumbo et al. (2021). In detail, QL5=very good; QL4= good; QL3=fair; QL2=poor; QL1=very poor.

At each QL, all the samples were subjected to postharvest quality evaluation, detecting colour parameters by a traditional colorimeter (CR400, Konica Minolta, Osaka, Japan) and physical and chemical parameters, in detail respiration rate (Kader, 2002), electrolyte leakage (Kim et al., 2005) and total chlorophyll content (Cefola and Pace, 2015). Then, images of the same samples were acquired by the CVS for non-destructive quality assessment and for recognizing traits related to the sustainability of the cultivation management used on the field, with specific reference to water and nutrients use (WP1). Image pre-processing was applied: to separate the product from the background; to identify the colour-chart placed in the scene to estimate the effects of lights and of the sensors and to correct colours to minimize these effects. Three colour correction methods (white balance, linear correction, and polynomial correction) with increasing level of complexity were evaluated and compared in terms of consistency of colour measurements and of classification performance. Linear colour correction proved to be the best trade-off between efficacy and efficiency providing a slightly lower performance than polynomial correction with significantly simpler computation. Finally, a Random Forest model was used to train classifiers to assess the QL of rocket leaves and to identify the treatments used during the cultivation.

All the postharvest quality parameters measured by traditional destructive methods were significant in QL assessment of fresh-cut rocket leaves. The proposed classifier based on the Random Forest model was able to identify and select the most relevant colour traits for both the tasks (QL assessment and treatment identification) without human intervention. The accuracy achieved in evaluating QLs of rocket leaves during storage was high (about 95 %), while the performance in discriminating the cultivation approach was lower and not sufficient for practical applications (about 65 -70 %). Indeed, the different cultivation approaches did not significantly affect the visual characteristics of the product and the destructive measures: this task needs further investigations.

Another research activity was carried out to develop and validate the capability of the non-destructive and contactless CVS to assess the visual quality changes during the cold storage of fresh-cut rocket leaves coming from soil and soilless growing systems (WP1) and to estimate some internal quality attributes (chlorophyll and ammonia content) also through the packaging material. Evaluating quality through the package is critical to identify the regions of the bag where the product is visible without shadows or highlights created by illumination: this is mandatory to measure colour properties in a reliable and meaningful way. At harvest, rocket leaves, cultivated on soil or soilless system (WP1), were packed in open PP bags and stored at 10 °C for about 18 d. During storage, all samples were observed to attribute the QL according to the rating scale reported in Figure 3 and the postharvest quality traits were evaluated by destructive conventional methods (colour parameters, chlorophyll content, ammonia content (Fadda et al., 2016) and electrolyte leakage). Then, images of unpackaged and packaged samples were acquired by the CVS. During image acquisition, no constraints were imposed on the position of the product in the bag, on the position of the bag in the scene or on the highlights created by the illumination on the surface of the bag: this was necessary to demonstrate the applicability of this technology into a real industrial line. Colour correction was performed by the linear model, identified as the best trade-off between effectiveness and computational complexity in the previous research activity. Packed and unpacked products were processed using exactly the same phases apart from the artefacts' elimination step applied to the images of packaged products to select the regions where the colour information was meaningful, without interference from light artefacts and reflections. At last, the Random Forest model was used to solve both the classification problem (assessment of the QLs) and the regression problems (estimation of quality marker parameters such as chlorophyll and ammonia contents). The same architecture was used for all the tasks, by simply changing the training data. The histogram of the image, evaluated in the a-b plane of the CIELab colour space, was used as the set of features. The Random Forest model was able to automatically select the subset of values more suitable for solving each task.

All the postharvest quality parameters detected by conventional analysis during the storage of fresh-cut rocket leaves were significant in the QL assessment and, among them, chlorophyll and ammonia contents proved to be useful marker parameters for the objective separation of each QL considered, both on soil and soilless cultivation approach.

The CVS was able to operate without relevant differences on unpackaged and packaged products. The test was done joining all the samples, regardless of the cultivation approach: the results showed a not significant performance loss on packaged leaves (Pearson's linear correlation coefficient of 0.84 for chlorophyll and 0.91 for ammonia) with respect to unpackaged ones (0.86 for chlorophyll and 0.92 for ammonia) (Figure 4).

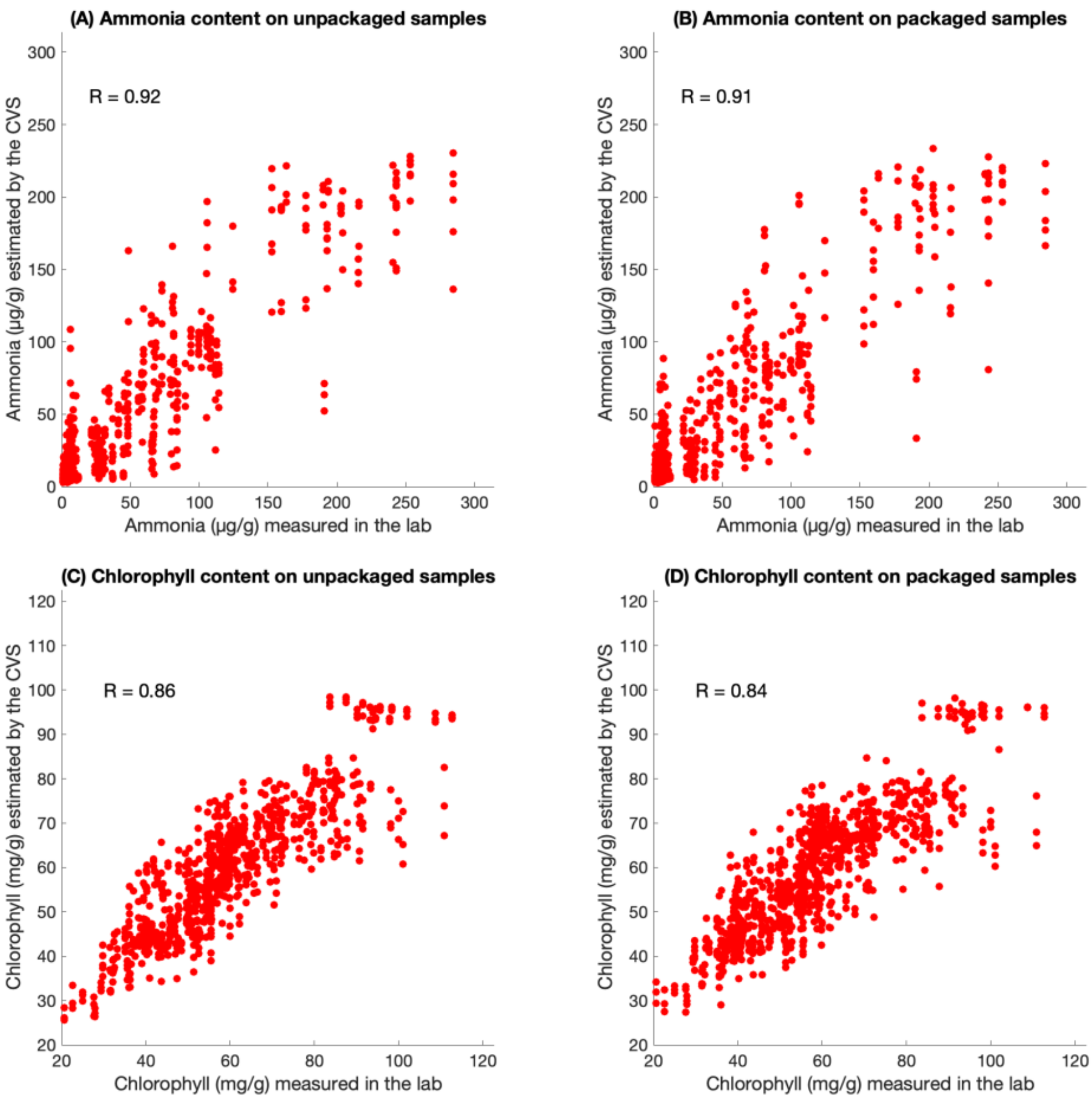


Figure 4. Values estimated by the CVS (abscissa) vs. values measured in the laboratory (ordinate) for ammonia content on unpackaged (A) and packaged (B) rocket leaves and for total chlorophyll content on unpackaged (C) and packaged (D) samples (Palumbo et al., 2022).

Finally, three Partial Least Square (PLS) models were performed to predict the QL using as predictors chlorophyll and ammonia contents obtained by destructive methods (Model I), by CVS on packaged products (Model II) and by CVS on unpackaged ones (Model III) (Table 1). The results showed high performances in terms of R^2 and the model obtained by predictors estimated non-destructively by the CVS (Model II and III) provided better performances in the QL prediction than one obtained by destructive analysis, in both calibration and validation.

Table 1. Root Mean Square Error (RMSE) and the coefficient of determination (R^2) in calibration ($_c$) or validation ($_v$) of the Partial Least Square (PLS) Models predicting visual quality of rocket leaves (Palumbo et al., 2022)

PLS Models	Predictors	$RMSE_c$	R^2_c	$RMSE_v$	R^2_v
I	Total chlorophyll and ammonia obtained by destructive methods	0.45	0.9	0.86	0.70
II	Total chlorophyll and ammonia obtained by CVS on packaged rocket leaves	0.46	0.89	0.75	0.77
III	Total chlorophyll and ammonia obtained by CVS on unpackaged rocket leaves	0.46	0.89	0.7	0.8

2.3. WP3: marketing strategies to support the added value of the products LIP and ND certified

Implementing a marketing strategy, based on often intangible characteristics to consumers such as LIP and ND, it is not an easy task. Low impact practices do not have a highly distinctive impact on product characteristics nor determine unique taste, flavour, or look elements to consumers. However, certifications could be used to signal quality through the application of standards of quality and practices. Whether certifications could be effective in terms of marketing in the case of products LIP and ND, or for signalling quality in general is matter of discussion. Vecchio and Annunziata (2011), for instance, in their work question the possibility of effective understanding of certification by consumers. At this purpose the research team of WP3 decided to implement a different strategy and test it on the market. Visual storytelling certifying LIP and ND has been then hypothesized to better communicate the importance and the impact of those practices on food.

The research activity, therefore, has been organized in three steps: identifying the communication strategy and set-up; testing through focus-groups the opportunity conditions for farms and companies; testing though a survey and an econometric analysis the consumers' preference and their willingness to pay for products with LIP and ND. Therefore, a draft platform has been developed containing basic communication rules in order to highlight sustainability attributes of products through storytelling. Workflow has been established and a simulation has been conducted (Figure 5).

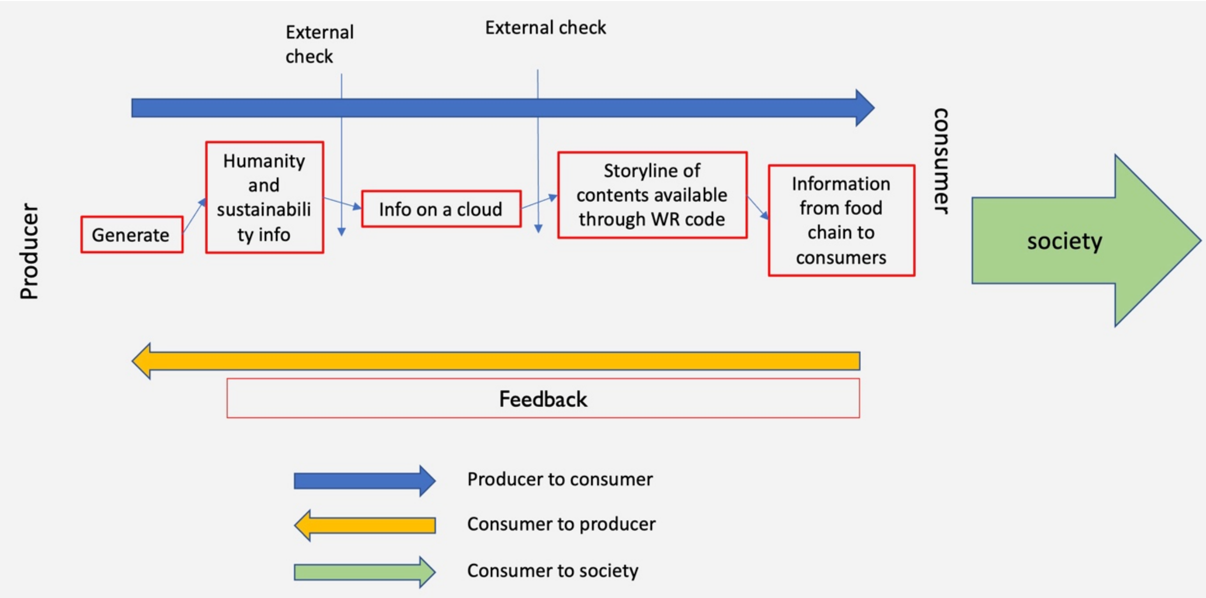


Figure 5. Workflow for products LIP and ND certified platform

Focus group with producers has allowed verifying the general appreciation for the marketing approach and allowed a better set-up of the strategy. Finally, a picture-based simulation has been produced for the final test and the survey to consumers (Figure 6).

Figure 6. Picture based simulation of visual storytelling certification for LIP producs and ND.

As last activity, a questionnaire based survey has been prepared and administered to 467 consumers and an econometric model to estimate willingness to pay and consumers orientation has been set up and then estimated. The whole set of activties within the research project allowed understanding how important is a correct communication of products and how different could be the perception of a product based on how you certify or narrate the production method. Result allow understanding that older conumsers are more aware of sustainability and are more willing to pay for LIP products. Psicological profile such as traditionalism and benevolence identify the consumer that, more than other profiles, would be willing to pay a higher price.

3. CONCLUSIONS

Sensor-based fertigation management applied to rocket leaves and tomato confirmed to be a feasible approach to improve sustainability of soilless cultivation, also in cases where the complete and rapid switch to closed cycle recirculation systems is still impaired by economic, social, and environmental factors such as in Mediterranean area.

The results of this project related to non-destructive discrimination of tomatoes and rocket leaves, according to cultural practices using different levels of inputs (water and fertilizers), indicated the high potentiality of HSI and FT-NIR techniques for the authentication of sustainable growing methods. Moreover, project activities' proved CVS as an effective tool for evaluating the product quality also through the bag, even working only on the regions of the image that provide meaningful colour information about the product's surface. The integration of machine learning modules inside the CVS confirmed to be useful to simplify the design and tuning, done mostly automatically without human intervention. Moreover, the flexibility introduced by machine learning makes the resulting architecture more flexible in adapting to different products and applications.

As regards the marketing approach, consumers resulted willing to pay a higher price for LIP products certified through a visual storytelling platform. In the next future, there could be a good chance that sustainability-oriented practices coupled with a visual storytelling certification style could gain shares on food markets.

REFERENCES

Amodio, M.L., Chaudhry, M.M.A., Colelli, G., 2020. Spectral and hyperspectral technologies as an additional tool to increase information on quality and origin of horticultural crops. *Agronomy*, 10(1), 7.

Amodio, M.L., Capotorto, I., Chaudhry, M.M.A., Colelli, G., 2017a. The use of hyperspectral imaging to predict the distribution of internal constituents and to classify edible fennel heads based on the harvest time. *Computers and Electronics in Agriculture*, 134, 1-10.

Amodio, M.L., Chaudhry, M.M.A., Colelli, G., 2017b. The use of non-destructive techniques to assess the nutritional content of fruits and vegetable. In: Yahia, E.M. (Ed.), *Fruit and Vegetable Phytochemicals: Chemistry and Human Health*, Wiley-Blackwell.

Amodio, M.L., Cabezas-Serrano, A.B., Peri, G., Colelli, G., 2011. Post-cutting quality changes of fresh-cut artichokes treated with different anti-browning agents as evaluated by image analysis. *Postharvest Biology and Technology*, 62, 213–220.

Arendse, E., Fawole, O.A., Magwaza, L.S., Opara, U.L., 2018. Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *Journal of Food Engineering*, 217, 11-23.

Ariana, D.P., Lu, R.F., 2010. Evaluation of internal defect and surface color of whole pickles using hyperspectral imaging. *Journal of Food Engineering*, 96 (4), 583-590.

Arias, R., Lee, T.C., Logendra, L., Janes, H., 2000. Correlation of lycopene measured by HPLC with the L*, a*, b* color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content. *Journal of Agricultural and Food Chemistry*, 48(5), 1697-1702.

Babellahi, F., Paliwal, J., Erkinbaev, C., Amodio, M.L., Chaudhry, M.M.A., Colelli, G., 2020. Early detection of chilling injury in green bell peppers by hyperspectral imaging and chemometrics. *Postharvest Biology and Technology*, 162, 111100.

Bona, E., Marquetti, I., Link, J.V., Makimori, G.Y.F., Da Costa Arca, V., Lemes, A.L.G., Ferreira, J.M.G., Scholz, M.B.S., Valderrama, P., Poppi, R.J., 2017. Support vector machines in tandem with infrared spectroscopy for geographical classification of green arabica coffee. *LWT – Food Science and Technology*, 76, 330-336.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2018. Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material. *Journal of Food Engineering*, 223, 46-52.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2017. Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. *Computers and Electronics in Agriculture*, 140, 303-310.

Cefola, M.; Pace, B., 2015. Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. *Journal of Food Processing and Preservation*, 39, 2523–2532.

Fadda, A., Pace, B., Angioni, A., Barberis, A., Cefola, M., 2016. Suitability for ready-to-eat processing and preservation of six green and red baby leaves cultivars and evaluation of their antioxidant value during storage and after the expiration date. *Journal of Food Processing and Preservation*, 40(3), 550–558.

FAO, 2013. Good Agricultural Practices for greenhouse vegetable crops. In: *Principles for Mediterranean climate areas.* Rome: FAO. E-ISBN 978-92-5-107650-7.

Guo, Z., Huang, W., Chen, L., Zhao, C., Peng, Y., 2013. Geographical classification of apple based on hyperspectral imaging. – In: Moon, S.K., Shu-I, T., Kuanglin, C., (Eds.), *Sensing for Agriculture and Food Quality and Safety* (pp. 103-110). SPIE, 8721.

Kader, A.A.*, 2002.* Methods of gas mixing, sampling and analysis. In. - In: Kader, A.A., (Ed.), *Postharvest Technology of Horticultural Crops* (pp. 145–148), University of California Agriculture and Natural Resources: Oakland, CA, USA.

Kim, J.G., Luo, Y., Tao, Y., Saftner, R.A., Gross, K.C., 2005. Effect of initial oxygen concentration and film oxygen transmission rate on the quality of fresh-cut romaine lettuce. – *Journal of the Science of Food and Agriculture*, 85, 1622–1630.

Khanmohammadi, M., Karami, F., Mir-Marqués, A., Garmarudi, A.B., Garrigues, S., De La Guardia, M., 2014. Classification of persimmon fruit origin by near infrared spectrometry and least squares-support vector machines. *Journal of Food Engineering*, 142, 17-22.

Krystallis, A., Chryssohoidis, G., 2005. Consumers' willingness to pay for organic food: Factors that affect it and variation per organic product type. *British Food Journal*, 107(5), 320-343.

Lara, M.A., Lleó, L., Diezma-Iglesias, B., Roger, J.M., Ruiz-Altisent, M., 2013. Monitoring spinach shelf-life with hyperspectral image through packaging films. *Journal of Food Engineering*, 119(2), 353-361.

Liu, Y., Pu, H., Sun, D.W., 2017. Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications. *Trends in Food Science and Technology*, 69, 25-35.

Massa, D., Magán, J.J., Montesano, F.F., Tzortzakis, N., 2020. Minimizing water and nutrient losses from soilless cropping in south-ern Europe. *Agricultural Water Management*, 241, 106395.

Montesano, F.F., Serio, F., Mininni, C., Signore, A., Parente, A., Santamaria, P., 2015. Tensiometer-based irrigation management of subirrigated soilless tomato: effects of substrate matric potential control on crop performance. *Frontiers in Plant Science*, 6, 1150.

Montesano, F.F., Van Iersel, M.W., Boari, F., Cantore, V., D'amato, G., Parente, A., 2018. Sensorbased irrigation management of soilless basil using a new smart irrigation system: Effects of setpoint on plant physiological responses and crop performance. *Agricultural Water Management*, 203, 20-29.

Montesano, F.F., D'imperio, M., Renna, M., Parente, A., Leoni, B., Serio, F., 2021. Soilless production of rocket (*Diplotaxis tenuifolia* L.) as affected by low and high input of water and fertilizers. Poster presentation at "III International Symposium on Soilless Culture and Hydroponics: Innovation and Advanced Technology for Circular Horticulture", ISHS Hydro2020, Lemesos (Cyprus), 19-22 March 2021.

Pace, B., Cavallo, D.P., Cefola, M., Colella, R., Attolico, G., 2015. Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio. *Innovative Food Science and Emerging Technology*, 32, 200-207.

Pace B., Cefola M., Da Pelo P., Renna F., Attolico G., 2014. Non-destructive evaluation of quality and ammonia content in whole and fresh-cut lettuce by computer vision system. *Food Research International*, 64, 647-655.

Pace, B., Cefola, M., Renna, F., Renna, M., Serio, F., Attolico, G., 2013. Multiple regression models and computer vision systems to predict antioxidant activity and total phenols in pigmented carrots. *Journal of Food Engineering*, 117(1), 74-81.

Pace, B., Cefola, M., Renna, F., Attolico, G., 2011. Relationship between visual appearance and browning as evaluated by image analysis and chemical traits in fresh-cut nectarines. - *Postharvest Biology and Technology*, 61(2), 178-183.

Palumbo, M., Attolico, G., Capozzi, V., Cozzolino, R., Corvino, A., De Chiara, M. L. V., Pace, B., Pelosi, S., Ricci, I., Romaniello, R., Cefola, M., 2022a. Emerging Postharvest Technologies to Enhance the Shelf-Life of Fruit and Vegetables: An Overview. *Foods*, 11(23), 3925.

Palumbo, M., Pace, B., Cefola, M., Montesano, F. F., Colelli, G., Attolico, G., 2022b. Non-destructive and contactless estimation of chlorophyll and ammonia contents in packaged fresh-cut rocket leaves by a Computer Vision System. *Postharvest Biology and Technology*, 189, 111910.

Palumbo, M., D’imperio, M., Tucci, V., Cefola, M., Pace, B., Santamaria, P., Parente, A., Montesano, F.F., 2021a. Sensor-based irrigation reduces water consumption without compromising yield and postharvest quality of soilless green bean. *Agronomy*, 11, 2485.

Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Serio, F., Colelli, G., Attolico, G., 2021b. Self-configuring CVS to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. *Agronomy*, 11 (7), 1353.

Patel, K.K., Kar, A., Jha, S.N., Khan, M.A., 2012. Machine vision system: a tool for quality inspection of food and agricultural products. *Journal of Food Science and Technology*, 49(2), 123-141.

Piazzolla, F., Amodio, M.L., Colelli, G., 2017. Spectra evolution over on-vine holding of Italia table grapes: prediction of maturity and discrimination for harvest times using a Vis-NIR hyperspectral device. *Journal of Agricultural Engineering*, 48(2), 109-116.

Piazzolla, F., Amodio, M.L., Colelli, G., 2013. The use of hyperspectral imaging in the visible and near infrared region to discriminate between table grapes harvested at different times. *Journal of Agricultural Engineering*, 44(2), e7-e7.

Renna, M., D'imperio, M., Maggi, S., Serio, F., 2022. Soilless biofortification, bioaccessibility and bioavailability: signposts on the path to personalized nutrition. *Frontiers in Nutrition*, 9, 966018.

Rinaldi, M., Castrignanò, A., De Benedetto, D., Sollitto, D., Ruggieri, S., Garofalo, P., Santoro, F., Figorito, B., Gualano, S., Tamborrino, R., 2015. Discrimination of tomato plants under different irrigation regimes: analysis of hyperspectral sensor data. *Environmetrics*, 26(2), 77-88.

Rouphael Y., Kyriacou M.C., 2018. Enhancing quality of fresh vegetables through salinity eustress and biofortification applications facilitated by soilless cultivation. *Frontiers in Plant Science*, 9, 1254.

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.

Schiattone, M.I., Candido, V., Cantore, V., Montesano, F.F., Boari, F., 2017. Water use and crop performance of two wild rocket genotypes under salinity conditions. *Agricultural Water Management*, 194, 214-221.

Stanghellini, C., 2014. Horticultural production in greenhouses: Efficient use of water. *Acta Horticulturae*, 1034, 25-32.

Torrellas, M., Antón, A., López, J.C., Baeza, E.J., Parra, J.P., Muñoz, P., Montero, J.I., 2012. LCA of a tomato crop in a multi-tunnel greenhouse in Almeria. *International Journal of Life Cycle Assessment*, 17(7), 863-875.

Tsouvaltzis, P.; Babellahi, F.; Amodio, M. L.; Colelli, G. 2020. Early detection of eggplant fruit stored at chilling temperature using different non-destructive optical techniques and supervised classification algorithms. *Postharvest Biology and Technology*, 159, 111001.

Workman, J., Shenk, J. 2004. Near-Infrared Spectroscopy in Agriculture. *Agronomy Monographs*, 44, 3-10.

Yang, Y.C., Sun, D.W., Pu, H., Wang, N.N., Zhu, Z., 2015. Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imagine coupled with model fusion*.* *Postharvest Biology and Technology*, 103, 55-65.

Vecchio, R., Annunziata, A., 2011. The role of PDO/PGI labelling in Italian consumers' food choices. *Agricultural Economics Research Review*, 12(2), 80-98

Chapter 2

COMPUTER VISION SYSTEM BASED ON CONVENTIONAL IMAGING FOR NON-DESTRUCTIVELY EVALUATING QUALITY ATTRIBUTES IN FRESH AND PACKAGED FRUIT AND VEGETABLES

Michela Palumbo[1,2], Maria Cefola[1], Bernardo Pace[1], Giovanni Attolico[3], Giancarlo Colelli[2]

[1] Institute of Sciences of Food Production, CNR, c/o CS-DAT, Via Michele Protano, 71121 Foggia
[2] Department of Agriculture, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Via Napoli 25, 71122 Foggia.
[3] Institute on Intelligent Industrial Systems and Technologies for Advanced Manufacturing, CNR, Via G. Amendola, 122/O, 70126 Bari

ABSTRACT

Quality assessment of fresh fruit and vegetables is an arduous and complex process which needs very intensive labour of correlation among sensory and subjective determinations and conventional destructive methods. Consumers' requests for fruit and vegetables with high quality in terms of appearance, nutritional value and safety have prompted industries and researchers to develop rapid, precise and low-cost techniques for food analysis. Among innovative techniques, image analysis by computer vision systems based on conventional imagin (CVS-CI) have proved to be effective and suitable for application at industrial level. This review summarizes developments on CVS-CI technology for the evaluation, along the entire distribution chain, of external defects, colour changes and internal chemical-physical attributes of fresh fruit and vegetables, with or without the presence of plastic packaging. The most interesting researches carried out during the last ten years on CVS-CI employments are reported and discussed. The description of each application points out the performances obtained, the hardware components, the image processing techniques used to extract information from the images acquired, the classification/regression models used to grade products and to estimate their quality traits. Finally, future perspectives and possible new applications of CVS-CI in postharvest field are proposed.

Keywords: contactless technology, fresh produce, image analysis, packaging material, quality assessment.

1. INTRODUCTION

The quality of fruit and vegetables is related to their degree of maturation, whose combination of physical and chemical attributes determine the acceptability by consumers (Brasil and Siddiqui, 2018). It is possible to distinguish five different attributes that define the quality of fruit and vegetables: visual quality or appearance (i.e. size, shape, colour, and defects), texture, flavour (taste and aroma), nutritional values, and safety. While appearance, texture and flavour are directly evaluated by consumers through the use of senses, nutritional values and safety cannot be perceived and determined at the time of purchase. According to Kader and Rolle (2004), the quality of fresh fruit and vegetables is initially evaluated by consumers only on the base of external aspects. The consumers' satisfaction of this first experience is then determined by internal quality parameters (such as acidity, sweetness, sugars to acid ratio and texture) related to taste and aroma at the eating moment. So, fresh and fresh-cut fruit and vegetables should have an attractive appearance to induce the first purchase and acceptable flavour and texture, as well as appropriate nutritional values, to convince consumers to continue their purchases.

Generally, the quality of fruit and vegetables is assessed by sensory and subjective determination, using scoring rating scales, whose single quality levels are characterized by a brief description and exemplifying photographs of the product at hand (Amodio et al., 2007). Moreover, conventional destructive methods are used to measure the chemical and physical attributes of fruit and vegetables to support the sensory evaluation: taste (i.e. crunchiness, bitterness and sweetness) and sugar and acid content in chicory and raspberry (François et al., 2008; Stavang et al., 2015; Aaby et al., 2019); appearance and colour traits, flavour and soluble solids content, hedonic liking and volatiles in mango (Salinas-Hernandez et al., 2015; Sung et al., 2019); sensory (i.e. juiciness, mealiness, etc.) and instrumental texture attributes in melon genotypes (Bianchi et al., 2016; Farcuh et al., 2020).

Although still widely used, analytical and destructive techniques are time-consuming and expensive, adversely impact the environment, may require sophisticated equipment and need careful sample preparation. Finally, they are unsuitable for application in industrial lines, where rapid, reliable, non-destructive, less expensive and less polluting methods for grading fruit and vegetables, assessing their quality and detecting defects are required (Narendra and Amithkumar, 2019).

Recently researchers have focused on the use of contactless, non-destructive, rapid and accurate, as well as non-polluting, techniques for fruit and vegetables analysis to objectively assess sensory and compositional quality. They can be considered complementary along the supply chain enabling time and cost saving, continuous and reliable monitoring and reduction of impact on environment (Chaudhry et al., 2020).

Computer vision systems (CVS) represent an innovative and contactless non-destructive technology suitable for in-line grading and quality assessment of fruit and vegetables (Fan et al., 2020).

The most common CVS for the quality inspection of fresh produce is a traditional system based on RGB colour cameras that reproduce the vision of human eyes using three monochromatic filters centred on red (R), green (G) and blue (B) wavelengths at 700, 546 and 435 nm, respectively (Lorente et al., 2012). CVS based on conventional imaging (CVS-CI) mimic human vision, acquiring and analysing images of the visible surface of food to assess its quality (Bhargava and Bansal, 2021). The CVS-CI can measure and detect many external quality traits (i.e. colour, shape, size and texture) and defects. CVS-CI automatically extract from images the most discriminative features related to quality and grade product using properly trained classification or regression models.

Recently, with the advances in hardware, software and high-resolution cameras, multispectral and hyperspectral CVS have been developed as efficient technologies for quality assessment of agricultural products (Baranowski et al., 2012; Chen et al., 2017; Li et al., 2018; Sendin et al., 2018). The spectral imaging data acquired by the hyperspectral and multispectral CVS provide information about internal and external traits that the CVS-CI can difficultly evaluate and analyse (Zhang et al., 2014). Such techniques provide consistent and accurate quality assessment but acquiring multispectral and hyperspectral images produces a large amount of data, requires costly and complex devices for acquisition and processing. Their application often requires specific competences by users and poses relevant constraints on the acquisition environment. These characteristics make them unfeasible for cost effective and pervasive application along the supply-chain to provide a continuous monitoring of the parameters of interest from harvest to final customers. The use of CVS-CI, where feasible, can provide greater flexibility to match cost (they exploit cameras that can be already available on the market at relatively low cost), time constraints and environment requirements at different points of the supply chain up to the final user.

A careful design is required to integrate mechanical and optical components for producing digital images that contain the relevant information for the task at hand (Patel et al., 2012). CVS-CI typically uses high resolution charged coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) digital cameras based on RGB colour for image acquisition. They involve the choice of a proper illumination system, carefully selecting spectral distribution and spatial geometry of the light sources. The position as long as the choice between front or back lighting, the type of lamps (incandescent, fluorescent, lasers, led or infrared lamps), colour quality and uniform distribution of the illumination are all important concerns when designing an efficient and accurate vision system (Zhang et al., 2014). Finally, a personal computer normally collects data acquired by the sensors, performs some basic processing (colour correction, segmentation, feature extraction), accomplishes features classification or parameters estimation by constructing suitable models using statistical methods or machine learning techniques (i.e. decision trees, regression trees, ensemble learning, random forest or convolutional neural network) (Figure 1).

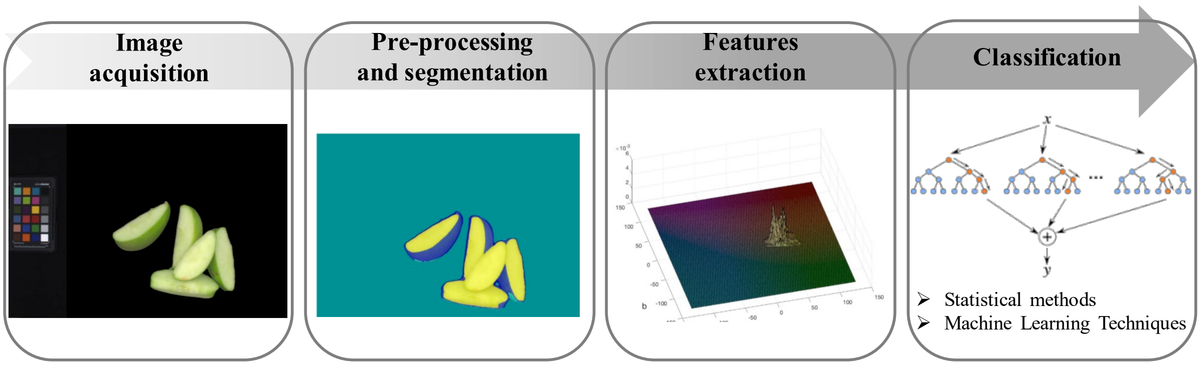


Figure 1. General workflow of an image processing performed by the CVS-CI.

In the last decades, many researchers have studied CVS-CI technology to automatically assess several and relevant attributes of fruit and vegetables in quality control. Classifying and sorting are key factors in quality assessment of fruit and vegetables during postharvest handling. Computer vision is highly effective in grading foods free from defects and with good appearance in terms of colour, texture, shape and weight. Moreover, recent research works have demonstrated the possibility to predict also internal quality parameters of products through elaborated images acquired by a conventional vision system. During postharvest storage, ripening or senescence processes lead to an alteration of inner quality of fruit and vegetables causing changes in the visual appearance (colour or texture) (Xia et al., 2016). The existence of correlations between colour and physical or chemical attributes (i.e. vitamins, antioxidants, total phenols, titratable acidity, soluble solids content or enzymatic activity) allows the estimation of internal fruit properties starting from CVS-CI images.

In addition, a CVS-CI may be very useful also on high-convenience fresh-cut products, where a high level of quality in terms of appearance, sensorial, and nutritional characteristics is required and where the higher convenience often corresponds to a more rapid perishability during the shelf-life (Watada et al., 1996; Barrett et al., 2010; Francis et al., 2012). For these products, another important aspect to consider is the increasing consumers' need to ascertain the real quality level inside the packaging. The use of CVS-CI can be helpful in monitoring fresh-cut fruit quality along the whole supply chain. A relevant challenge to face in ensuring a correct quality assessment through the packaging material by CVS-CI is the separation of opaque and affected regions of bags from the transparent area where the product is visible with acceptable fidelity of visual appearance. This separation step requires robust and powerful segmentation approaches (Cavallo et al., 2018). At the moment, research results have been reported about the application of non-destructive technologies working at higher wavelengths than the visible spectrum to quality assessment of fresh-cut products through packaging material (Lara et al., 2013; Giovenzana et al., 2014). Few researchers have reported interesting results about the use of CVS-CI through packaging. A powerful improvement of applications of CVS-CIs comes from the integration of machine learning techniques in several steps of processing. These methodologies can increase the flexibility of systems while simplifying their tuning and deployment. Moreover, in several fields machine learning paradigms (such as deep-learning architectures) are providing unparalleled performances widening the applicability of CVS-CI to solving relevant problems in food evaluation.

To the best of our knowledge, a detailed summarization of recent developments and applications of CVS-CI to evaluate external and internal quality of fruit and vegetables is not available. The present comprehensive review aims to report the most interesting researches carried out on computer vision technology based on conventional imaging (CVS-CI) for non-invasive analysis of fresh produce. In detail, 37 articles published in the last decade on important international journals and on proceedings of international congresses were reviewed and discussed. The literature research was focussed on image analysis by conventional CVS for the assessment of quality traits of fresh and packaged fruit and vegetables at harvest or during the postharvest storage. Research works from 2011 to date were considered, reporting the methodology adopted and their most important findings on quality prediction in terms of accuracy. Finally, a brief discussion on challenges and future research needs is also mentioned. Each cited work is supported by Table 1, where objectives, classification methods and the main results are reported, and by Table 2, in which the main software and technical aspects of CVS-CI adopted for the image analysis of fruit and vegetables are described. Image analysis and classification can be performed by several kind of software available either as open source or upon paying a fee. In Table 2, the literature is listed according to the software used for the elaboration and classification steps of image analysis. Matlab is the most used one probably because it offers to customers specific image processing applications of easy and rapid insight. Finally, Table S1 shows additional information about methodology, results evaluation and validation tests developed by the authors cited in the review.

2. CVS-CI TO ESTIMATE THE VISUAL APPEARANCE OF FRESH AND PACKAGED FRUIT AND VEGETABLES ON THE BASIS OF EXTERNAL DEFECT AND COLOUR CHANGE

Recent literature reports several works carried out on the visual quality assessment of fruit and vegetables by the use of CVS-CI. Automatic external inspection of fresh produce by the use of innovative CVSs, on the basis of the morphological traits, is widely used in industrial lines for grading and classification (Chopde et al., 2017). In food sorting systems, it has been shown that CVS-CI can judge marketability and edibility of foods on the basis of shape, size and the presence of external defects on their surface, both at harvest and postharvest (Bagri and Johari, 2015; Sa et al., 2016; Jana et al., 2017). Automated grading speeds up the processing time, reducing errors occurring during manual classification. On the contrary, grading fruit and vegetables through other appearance features, such as bruises, colour changes, rottenness or other defects (i.e. stem-end and calyx confusion with other types of defects) is not yet efficient and is not still applied on automatic work lines in industry (Leiva-Valenzuela and Aguilera, 2013).

The colour of both raw and processed fruits is one of the most important and the first perceived quality hint that determines consumers' acceptance. In food technology, colour is traditionally expressed in the CIELab colour space that is device independent and in which distances between colours are closer to the evaluations made by human vision. Measurements are usually performed using colorimeters. Although these instruments are non-destructive, easy to use and to calibrate, they need to touch the samples and they normally measure only very small areas (limited to a few mm^2), being sensitive to the choice of sampling points. For these reasons their use at industrial level and for in-line monitoring is limited (Go˜ni and Salvadori, 2017).

In the last decades, digital image analysis has been largely used for the assessment of food colour. An important difference between a CVS-CI and a traditional colorimeter is the degree of dependency on the environment condition: colorimeters, touching the product, can reduce to a minimum the influence of the light in the environment. CVS-CIs, that operate at variable distances from the observed scenes, need a more careful calibration in order to account for the influence of the light present in the environment on the measurement of colours. Another relevant difference is the amount of colour information that the two instruments can acquire

and analyse. Indeed, a CVS-CI extracts quantitative colour information with the resolution of the pixel of the digital image whose geometrical correspondent region can be controlled by properly choosing sensor resolution and optics. A CVS-CI can examine the whole visible surface of a product and provide distinct colour information for every single pixel. Therefore, it can provide global statistical colour information about the whole product but also semantically separate different regions of the product, describe the spatial distribution of colour and even analyse separately the different objects present in the scene (Wu and Sun, 2013).
Many research applications have demonstrated the speed, accuracy, effectiveness and consistency of CVS-CI in the assessment of colour features and the evaluation of colour change in whole and fresh-cut fruit and vegetables.

3. DETECTION OF EXTERNAL DEFECTS

3.1. Fruit

CVS-CI proved to be a useful tool for the improvement of blueberries quality during the postharvest storage (Leiva-Valenzuela and Aguilera, 2013). A pattern recognition methodology is developed to sort blueberry into four classes and two fruit orientations through the extraction and selection of visible features from images (Table 1). This methodology could also improve common commercial sorting systems whose classification is a grading only on the overall colour of berries, without recognizing specific defects on the fruit surface (Table S1). Among several classifiers applied, the best ones were the linear discriminant analysis (LDA) and the support vector machine (SVM), that successfully distinguish the blueberry's orientation in 96.8 % of the cases and were able to classify berries with good accuracy (Table 1). The proposed statistical pattern recognition methodology is a useful and promising single-step tool for in-line sorting and grading of blueberries with regard to different defects and orientations.
An automated fruit grading system was designed to detect defects on the surface of mango and apples with the aim to achieve greater efficiency for manual classification (Table 1). In detail, for image acquisition, a rotating desk with a 12 V motor was used to acquire images of the entire surface of the products (Table 2). A graphical user interface (GUI) was specifically designed to allow operators to interact with the imaging system and to simplify and speed up the grading operation. The GUI showed the total number and the positions of defects (Table S1) (Ali and Thai, 2017).
It was possible to effectively detect the defects on orange surface by the development of a new segmentation algorithm without additional lightness correction on images acquired by the CVS-CI (Table 2) (Rong et al., 2017). For this purpose, about 1200 images of randomly selected

samples were analysed to classify several types of surface defects (Table 1). The main steps of the image processing consisted of background removal, image binarization with a sliding comparison window, image subtraction, hole filling and identification and removal of stem-end pixels (Table S1). The proposed method was able to correctly detect 97 % of defective oranges and had a performance rate of 91.9 % in the individual defect detection (Table 1).

Quality control plays an important role in apple-based industries. Recently, an apple grading computer vision algorithm was introduced into an ordinary machine vision system (Table 2) (Moallem et al., 2017). This algorithm firstly detected stem-end, by a combination of morphological methods and a classifier based on mahalanobis distance, and calyx area, by applying K-means clustering on the Cb component in YCbCr colour space, separating them from defective regions; then defects segmentation was achieved using a multi-layer perceptron (MLP) neural network. Two types of classification were done analysing colour, textural and geometric features (Table 2): two apple categories were distinguished in the first ranking and 3 categories in the second one (Table 1). A good accuracy was achieved for the calyx detection algorithm both when the stem-end was inside the image (94 %) and when it was outside (81 %), even if it was lower than previous one. Moreover, SVM classifier was the best one for the recognition of 2 categories and 3 quality ranks (Table 1).

A novel methodology, through the construction of three algorithms, was proposed for automatic defect identification and maturity detection of mango by the use of CVS-CI for practical application on industrial lines (Table S1) (Sahu and Potdar, 2017). As results, the fruits were graded into two classes (defected or not-defected and mature or immature) based on the quality ratio and maturity, respectively. If the value of the quality ratio was greater than the threshold value, the fruit was rotten, while if the value was less than the threshold value, the fruit was good. The proposed algorithm efficiently and accurately determined the quality of mango (Table 1) and a similar approach was used also for maturity detection (Table S1).

Finally, as for CVS-CI applications on fruits, a new automatic approach was proposed for manual orange classification (Chen et al., 2018). Oranges were divided into 3 groups (Table 1) for each of the characteristic indexes: colour, size, defect and shape. Finally, the extracted characteristic parameters were studied and classified by using a back propagation (BP) neural network (Table 1). The grading accuracy was very high (Table 1) demonstrating how this classification system, based on image analysis, could realize a real-time automatic oranges grading detection reducing the labour loss and improving the efficiency compared to the traditional method.

3.2. Vegetables and fruit vegetables

As regard the CVS-CI implementation on vegetables, an automatic, accurate and low-cost system was developed for quality inspection of tomatoes in the agroindustry using image processing techniques to classify the fruit vegetables in terms of ripeness and defects (Arakeri and Lakshmana, 2016). Several colour and textural features were extracted from the acquired images and the final classification was carried out through an artificial neural network (ANN) (Table 1). The system was able to separate defective from non-defective and ripe from unripe tomatoes with very good performance (Table 1).

An eggplant grading system was developed to grade healthy or unhealthy samples using an automated CVS-CI, described in Table 2 (Akter and Rahman, 2017). The diseased areas of the eggplants were segmented by Otsu and binary transformation methods (Table 1 and S1). For the final classification, performed by K-nearest neighbour (KNN) method, features on the size, shape, colour and percentage of diseased areas were considered, distinguishing four categories of the product (Table 1). Because of little problems of shadow and lightning created during the image acquisition, the system was able to grade the eggplants with an accuracy of 88 % and further investigation are being studied to improve the final results.

3.3. Fresh-cut fruit and vegetables

Regarding the fresh-cut products, a CVS-CI was used to detect the freshness of fresh-cut spinaches stored in a plastic bag at 4 °C for 12 d (Huang et al., 2019). As they are extremely perishable, for both physiological and microbial degradation, spinaches were divided into 4 grades related to tissue decay, from good to bad (Table 1). KNN, SVM, and back-propagation artificial neural network (BPNN) were applied and compared in predicting spinach freshness (Table S1). BPNN and KNN models achieved the same classification accuracy as reported in Table 1. Furthermore, the study also applied the E-nose technology to obtain odour information of the spinach samples and a multisensory data fusion approach based on machine vision and E-nose data to detect the freshness of spinach during storage. As result, the BPNN model based on this multisensory data fusion widely improved the accuracy of spinach freshness detection (93.75 %).

4. EVALUATION OF COLOUR CHANGES

4.1. Fruit

Banana is a fruit widely addressed by image processing applications. Its peel colour, considered the main quality parameter for traders and consumers, shows relevant changes from harvest to the end of the artificial ripening process. A typical over-ripe process, called senescent spotting, consists of brown spots on the external peel. The banana maturity stage is usually classified by a 7-point scale, from green to yellow and a CVS-CI was implemented to grade the 7 ripening stages of bananas using colour and image texture information, as well as the development of brown spots, extracted by image analysis (Table S1) (Mendoza and Aguilera, 2004). Results showed that a simple classification technique, based on L^*, a^* and b^* parameters, brown area percentage and contrast, was able to identify the 7 ripening stages of bananas with an accuracy of 98 % (Table 1).

A recent research showed that the correct physiological maturity stage of bananas at harvest is as important as the assessment of the ripening stage after harvest to ensure the quality during the storage in the ripening chambers (Prabha and Kumar, 2015). With this objective a CVS-CI, described in Table 2, was implemented to extract colour and size features of fresh banana fruit images to classify 3 main categories before harvest, finding that the mean colour intensity and area features were more significant than other features for the assessment of maturity stages (Table 1). The model developed in this study could be used to identify an automatic detection system for banana maturity assessment directly in the field.

The accuracy of CVS-CI in comparison with the use of traditional colorimeters and sensory tests was demonstrated to evaluate banana peel browning during storage (Cho et al., 2016). The browning degree was estimated using the changes in RGB colour value and the changes in the CIE L^* a^* b^* colour values (Table 2). The methodology demonstrated that the G and CIE L^* a^* b^* values obtained by image analysis showed the highest correlation coefficient with a sensory test, while CIE L^* a^* b^* values measured by the colorimeter also showed high correlation coefficients, but comparatively lower than those detected by image analysis (Table 1).

Additionally, CVS-CI was also used to study the physical properties of persimmon fruits (Mohammadi et al., 2015). The authors developed a sorting algorithm to categorize fruits into 3 commercially maturity stages (Table 1) on the base of external colour able to eliminate the noise and binarize the acquired image, remove the black stains and finally extract 4 standard colour features (Table S1). The best results for grading persimmon fruits were reached by the quadratic discriminant analysis (QDA) classifier with an overall accuracy of 90.24 % (Table 1).

Maturity indices are also important for trade regulation and marketing strategy. In order to provide marketing flexibility and to guarantee the consumers' acceptance of plums, an implementation of image processing techniques, based on JPEG images, was developed to classify fruits into 4 categories related to different maturity stages (Kaur et al., 2018). The multi-attribute decision making (MADM) theory was used for making decisions on the base of colour and textural features to achieve the highest accuracy (Table 1). As relevant result, colour was the dominant factor for grading plums according to maturity stage and the error percentage between the manual and the calculated values of length and width was lower than 2.4 %, as reported in Table 1.

CVS-CI was recently applied to implement a method to classify pomegranate arils on the base of colour and size using image processing and artificial intelligence (Fashi et al., 2019). These features represent the most important elements to consider for grading pomegranate especially when consumed as a ready-to-eat product. In this study, the arils were graded by experts based on size and appearance into 3 categories and the ANN, adaptive neuro fuzzy inference system (ANFIS) and response surface methodology (RSM) classifiers were compared (Table 1 and S1). The ANN model showed a higher accuracy than that obtained for the ANFIS and RSM model (Table 1), representing a valid method to assess the quality of arils through image analysis.

The image processing and machine learning techniques may be combined to reduce the human intervention for the configuration and setting of algorithms with the purpose of increasing their final performances. A simplified CVS-CI was applied to two white table grape cultivars (Victoria and Italia) during cold storage (at 5 and 10 ◦C) for the assessment of 5 different quality levels (QL) (Table 1) (Cavallo et al., 2019). In the proposed approach, a random forest classifier was trained and tested and a large set of potential features were fed into the system, leaving to the machine learning tool the task of selecting how many and which attributes were best suitable for the classification task (Table S1). The system was able to separate the highest QL5 and QL4 from the others with high accuracies on cultivar Italia and Victoria, without significant differences between the compared storage temperatures (Table 1), providing to be suitable for a real context where often temperature regimes are not constant. Moreover, the obtained results demonstrated the possibility to develop classification models specific for single cultivars by the use of CVS-CI for continuous monitoring of the quality along the supply chain.

4.2. Vegetables and fruit vegetables

In a recent work, a low-cost tomato grading machine vision system based on RGB images and machine learning technology was introduced to develop an efficient and accurate calyx detection algorithm and classification approach to sort tomatoes into four different grading categories (Table 1) (Ireri et al., 2019). Colour features, grey level texture features (gray-level co-occurrence matrices, GLCM) and shape features were analysed comparing three recognition models: support vector machine - radial basis function (SVM-RBF), ANN and random forest (Table 1). Calyx detection was performed with an accuracy of 95 % by histogram thresholding based on the mean red and green values of the regions of interest. Defected regions were detected by an SVM-RBF classifier based on $L^*a^*b^*$ colour features, achieving a very high overall accuracy. The SVM-RBF outperformed all the compared models with the highest accuracy in grading 4 tomato categories based on colour and texture features (Table 1). The proposed machine vision system can be used for quality inspection, marketing, and packaging of tomatoes and to detect defects along the supply chain of tomato management.

Recently, an innovative CVS-CI was used to automatically select the relevant features for the classification of rocket leaves with reduced human intervention applying a random forest model (Palumbo et al., 2021). In detail, three colour correction models (white balance, linear correction, and polynomial correction) were evaluated and compared to identify the best solution for consistent colour measurement in terms of classification performance. The proposed CVS-CI was applied to fresh rocket leaves produced with two different cultural practices, with the aim to (i) objectively attribute 5 quality levels (QL) during storage at 10 ◦C and (ii) non-destructively discriminate the cultivation approach, using the colour information extracted from the digital images. The CVS-CI achieved an accuracy of about 95 % in QL attribution and of about 65–70 % in the discrimination of the cultivation approach (Table 1).

Finally, image analysis could represent an effective way to measure colour changes on the external surface of products characterized by a not uniform colour. An image analysis technique was applied to study colour changes and the visual appearance of “Borlotto” beans, commonly used in ready-to-use soup ingredients (Amodio et al., 2011a). These beans are characterized by bright tiny red spots on the surface that, since may undergo severe discolouration depending on the atmosphere composition within the package, represent a good indicator of packaging performance. Given the small and irregular dimensions of these spots, the traditional measurement approach may not be helpful. A new algorithm was defined and was able to objectively measure both spots and background colours in order to study the degradation kinetics over time, reporting positive and negative regressions between the appearance scores evaluated sensorially and the a^* and hue angle values, respectively (Table 1).

4.3. Fresh-cut fruit and vegetables

As for fresh-cut produce, an algorithm for rapid colour detection by image analysis based on RGB scale from the whole surface and the browned areas of fresh-cut artichokes was developed in order to objectively compare the efficacy of anti-browning agents, in addition to a sensory evaluation through a subjective scale (from 5, excellent, to 1, inedible) (Amodio et al., 2011b). For all colour parameters (L^*, a^*, b^*) obtained by CVS-CI a standard regression was performed to find possible correlations with sensory evaluation (Table S1). Significant correlations between all colour features were observed and L^* measured on the whole surfaces showed the highest correlation coefficients, while L^* and b^* measured on the brown areas showed lower, but still significant, correlations (Table 1). On the other hand, the a^* value was the least correlated parameter with an appearance score (showing negative correlation values).

An implementation of the image analysis through CVS-CI was proposed to obtain a global evaluation of colour of fresh-cut nectarine from a standard RGB image acquired by a 3 CCD digital camera (Pace et al., 2011). A novelty was the development of a polynomial colour transformation to correct the images before any further processing (Table S1). The correlation between the visual appearance, scored by a 5–1 sensory scale, and the colour parameters b^* and Chroma measured by the computer vision system was higher than that obtained using a colorimeter (Table 1), demonstrating how CVS-CI was more effective than the standard colorimetric method and how it can be suitable for real-time application in the processing industry.

As additional task, CVS-CI can automatically set parameters which were normally identified by operators and completely automate some of the image processing steps, adapting itself to a specific analysed product (Pace et al., 2015). For this purpose, an innovative CVS-CI, characterized by an automatic procedure for the extraction of colour features, was developed to evaluate the quality levels (QLs) changes occurring on two varieties of fresh-cut radicchio (Correlli and Botticelli) during the cold storage in air or in modified atmosphere packaging (Pace et al., 2015). Brightness and white colour of fresh-cut radicchio are considered freshness indicators and, among these, colour changes after processing and during storage occur on the white parts that tend to become brown. Results of this research proved that the average of the component a^* over the white pixels and the percentage of light white pixels over the whole visible surface allowed the best discrimination of the QL in three quality classes whose true value was verified on the base of sensory evaluation (Table 1).

The CVS-CI, for image acquisition of fresh-cut apples, was used to non-destructively study browning through colour and textural features (Subhashree et al., 2017). Colour and textural features selected by images acquired through a CVS-CI (Table 2) showed good results to detect browning on apple slices by image analysis (Table 1), although further research is needed to better understand the colour kinetics in this fresh-cut fruit.

Moreover, colour changes occurring on fresh-cut potatoes during cold storage were examined by the use of CVS-CI to perform quantitative analysis at different stages of browning (Hongyang et al., 2021). Several methods were built to classify and predict the browning features of fresh-cut potatoes (Table S1). The SVM was able to predict the storage life of the product with an accuracy of 96 %, while the partial least squares regression (PLSR) model showed an accuracy of 86 %, 96 % and 99 % for L^*, a^* and b^*, respectively (Table 1).
Regarding the use of image analysis by CVS-CI for the quality assessment of fresh-cut products through plastic bags, an automatic CVS-CI applied to fresh-cut iceberg lettuce was recently developed (Cavallo et al., 2018). The main purpose was to achieve a careful selection of the bag area that was meaningful to have similar performances of the analysis for both packaged and unpackaged products. In detail, during cold storage, fresh-cut lettuce leaves were placed in polypropylene bags and sensory evaluated according to a 5-point rating scale. Images were then acquired by a CVS-CI for each fresh-cut lettuce sample both with and without the bag. All acquired images were subsequently segmented using a convolutional neural network (CNN), identifying and selecting only pixels belonging to the fresh-cut lettuce and unaffected by artifacts induced by the illumination. The classification was made by a 3-nearest neighbours classifier applied on a two-dimensional colour histogram on the a^* and b^* components in the CIE L*a*b* space (Table 2). The performance loss due to the presence of packaging was not relevant (Table 1) and the proposed model was also successfully tested on commercial bags of lamb's lettuce. The CNN segmentation method was able to separate the opaque and affected regions from the product inside the commercial bag, showing its flexibility for the potential application along the supply chain.

5. CVS-CI TO ESTIMATE CHEMICAL-PHYSICAL ATTRIBUTES IN FRESH AND PACKAGED FRUIT AND VEGETABLES

Besides external attributes, CVSs-CI have already been used also for the assessment of inner quality, providing a valid contribution to the evaluation of fruit and vegetables. The retention of green colour is closely related to the total chlorophyll content, which degrades with the postharvest senescence, inducing a yellowing of the visible surface of the leaves. Besides total chlorophyll, ammonia content is also considered an indicator of freshness, when is detected in low amounts in the vegetal tissues (Tudela et al., 2013) and high concentrations of this component may cause tissue damage with visible effects on the overall quality of the product (darkening and browning of leaves) (Mastrandrea et al., 2016; Amodio et al., 2018). So, the possibility to predict these internal traits by CVS-CI on the base of colour changes occurring during storage allows us to use them as markers and objective parameters for quality level assessment of leafy vegetables.

The development of fruit colour during ripening is considered a maturity index and when its intensity increases, the ripening attributes also improve. Total soluble solids mainly contribute to the sweetness of fruits during ripening. It has been demonstrated that as the fruit colour development enhances, the total soluble solids content increases proportionately (Kaur et al., 2018). On the contrary, the trend of the fruit acidity is reversed: it decreases with the enhancement of the colour of the fruit during ripening.

During the storage period, enzymatic browning generally occurs on the cut surfaces of fresh-cut products for the interaction of phenolic components with polyphenol oxidase and peroxidase, resulting in brown pigments. The notable effect of this physiological disorder on the visual appearance of fruits could allow to predict the enzymatic activity responsible for the browning by the use of image analysis (Cho et al., 2016).

Antioxidant activity and total phenols content are two other chemical parameters closely related to the nutritional quality of fruit and vegetables. It has been demonstrated that total phenols may be the main contributor to the antioxidant activity of many fruits and vegetables (Li et al., 2012). Among the phenolic components, flavonoids are potent antioxidants and include pink, red, blue and purple pigments known as anthocyanins. During the ripening of some fruit and vegetables, a colour change in the superficial tissue from green to purple is due to an alteration of pigment concentration (i.e. accumulation of anthocyanins). This colour modification suggests the use of colour analysis to detect the antioxidant activity and the total phenols content, of these fruit and vegetables (Lou et al., 2012).

Many applications were reported in recent research works about the use of CVS-CI to estimate inner quality parameters of fruit and vegetables.

6. EVALUATION OF CHEMICAL-PHYSICAL ATTRIBUTES

6.1. Fruit

An experimental study was carried out to establish quantitative models between features measured by the CVS-CI and soluble solids content and pH of Kyoto grapes (Xia et al., 2016). Several colour features were extracted from the mean and standard deviation of the pixel values considering each RGB channel (Table 2). The root mean square error of cross-validation (RMSECV) was calculated on the base of all the prediction residuals. Furthermore, the models were also evaluated by the root mean square error of calibration (RMSEC) (Table 1). As result, RMSEC values were 0.097 and 0.773 on average for the SSC and pH prediction, respectively, while the RMSECV values were from 0.099 to 0.100 for pH and from 0.788 to 0.808 for the SSC prediction (Table 1). In a previously reported study, the CVS-CI applied on plums also showed a good correlation between fruit acidity (expressed as total soluble solids content) and the mean intensity of the green colour (R^2 = 0.997) and the R/G ratio (R^2 = 0.846) (Table 1) (Kaur et al., 2018).

The possibility of predicting polyphenol oxidase (PPO) and peroxidase (POD) enzyme activity on the banana peel browning during storage using the image analysis by CVS-CI was recently demonstrated (Nadafzadeh et al., 2018). Several colour features were extracted from the images acquired, as detailed in Table 2, and two equations were obtained by the use of a genetic programming modelling in order to predict PPO and POD activities during the browning process of banana peel. As a result, there were no significant differences between predicted and measured PPO and POD enzymatic activity, demonstrating the high performance of the image analysis and of the derived models (Table 1).

The integration of CVS-CI and colorimetric sensor array (CSA) may be used for a non-destructive methodology for a rapid and accurate evaluation of mango quality, hardness and total soluble solids (TSS) (Huang et al., 2018). In detail, three grades were defined by experts by sensory evaluation while hardness and TSS of mango samples were detected by both conventional and non-destructive methods. All data were elaborated using PCA to reduce dimensionality, support vector classification (SVC) model for qualitative grading of fruits, and support vector regression (SVR) model to elaborate the relationship between the conventional and CVS-CI data (Table S1). The SVC model showed high accuracy for the training and prediction sets, respectively, to classify mango fruits into 3 grades, while the SVR allowed good prediction of hardness and TSS (Table 1).

Additionally, the CVS-CI was proposed to estimate the maturity index of Kinnow mandarins which was strictly related to internal physiology and external peel colour changes (Hadimani and Mittal, 2019). Mandarins are usually graded from homogeneous deep green to typical deep orange/red colour by a numerical value called citrus colour index (CCI). In this work, a polynomial transformation based on camera characterization method was proposed to reduce the complexity of colour space transformation from RGB to $L^*a^*b^*$. Physicochemical parameters of TSS and pH of mandarins were measured and correlated with the changes in peel colour detected by CVS-CI (Table S1). Measured physicochemical properties had a good correlation with the change in peel colour (Table 1) and a high R^2 = 0.97 was achieved with partial least square regression (PLSR), showing confidence in the developed CVS-CI based method.

On oranges, an automatic CVS-CI able to assess the pH value from external visible-range peel images was presented (Sabzi et al., 2020). The proposed system was applied to 3 varieties of orange (Bam, Blood and Thomson): after image acquisition, 6 colour and texture discriminant features for each orange sample were selected among the extracted ones through the use of the hybrid artificial neural network - particle swarm optimization (ANN-PSO) and, finally, estimated orange pH regression values were provided by the multilayer perceptron (MLP) neural network architecture (Table 1). Results showed that the proposed method was able to predict the orange pH value from the external peel images with high accuracy (Table 1). The average time spent to estimate pH with the proposed non-destructive method was 0.42 s, which might be reasonable for in-line industrial applications.

Discrimination of the ripening stage (half-red or red) of strawberries (cv. Candonga) harvested at three different times by CVS-CI image analysis was satisfying (Palumbo et al., 2022a). After several subsequent steps of segmentation (Table S1), the pixel count of red and green area of images was performed to calculate their relative percentage area. Among the chemical indicators of ripening, titratable acidity was well correlated to image analysis (Pearson correlation coefficient of about 1), providing a suitable indicator for fast and non-destructive evaluation of the ripening stage in strawberries (Table 1).

A CVS-CI was developed also to evaluate the application of image processing technique for total soluble solids and pH prediction on strawberries (Basak et al., 2022). The channels of RGB, HSV and HSL colour spaces were used as input variables for developing MLR and support vector machine regression (SVM-R) models (Table 1). The SVM-R model, working on features in the HSV colour space, performed better than MLR model for total soluble solids and pH prediction (Table 1).

Finally, a very interesting application of image analysis provides an innovative and smart technology to predict the shelf life and the quality of kiwifruit during cold storage by acquiring and calculating RGB value extracted from photos taken by a smartphone camera (Table 2) (Li et al., 2022). Results reported that the R to B ratio values (Central R/B) was negatively correlated with titratable acidity and vitamin C contents and firmness and positively correlated with soluble solids content, total soluble sugars and total plate counts (Table 1). The study demonstrated how the extracted the RGB values with mobile phones may provide a rapid and consistent evaluation of postharvest quality of kiwifruit, which can be more suitable for distributors and consumers to rapidly assess the quality and storage time of the product.

6.2. Vegetables and fruit vegetables

The relationships between antioxidant activity (AA) and total phenols (TP) with colour features of a local landrace of pigmented carrot, characterised by different colours (from yellow to purple), were explored by a CVS-CI able to predict the AA or TP contents (Pace et al., 2013). The obtained regression models were able to successfully estimate the AA and the TP contents as their predicted levels showed a very good correlation with the measured AA and TP content when both internal and external parts of carrots were considered. When data included only the internal part (with a not uniform pigmentation) lower determination coefficients were obtained (i.e. $R^2 = 0.93$ for AA and $R^2 = 0.86$ for TP) (Table 1).
Moreover, a new contactless and non-invasive approach was proposed for the prediction of total chlorophyll content of fresh rocket leaves through the use of CVS-CI in order to replace the common non-destructive SPAD-meter, a commercially available device that detects the chlorophyll content in local small areas by touching the leaf, providing the information related to few square millimetres of leaf surface (Cavallo et al., 2017). The classification was performed by random forest regression, a machine learning technique suitable to model the correlation between total chlorophyll content of fresh-cut rocket leaves and colour values in the RGB and $L^*a^*b^*$ colour spaces (Table 2) with a better performance compared to the SPAD-meter (i.e. $R^2 = 0.90$ vs. $R^2 = 0.79$) (Table 1).

6.3. Fresh-cut fruit and vegetables

Recently, the combination of image processing by CVS-CI and the random forests model was proposed to solve the classification problem (visual quality level assessment) and the regression problem (prediction of senescence indicators as chlorophyll and ammonia content) on unpackaged and packaged rocket leaves (Table 2 and S1) (Palumbo et al., 2022b). The experiment showed a not significant performance loss on packaged products with respect to unpackaged ones (Table 1). The same CVS-CI, exploiting the machine learning components, was able to build effective models for both the problems just by changing the training data. Moreover, three partial least square models were built to predict the visual quality level of fresh-cut rocket leaves using as predictors the total chlorophyll and the ammonia contents obtained by destructive methods, CVS-CI through packaging and CVS-CI without packaging. The predictors obtained by CVS on samples with and without package provide better performances compared to the ones estimated by destructive analysis (R^2 in validation of 0.70, 0.77 and 0.80, respectively) (Table 1).

These latter results confirm the ability of the CVS-CI to evaluate the product also through the packaging, even working only on the regions of the image that provide meaningful colour information about the product's surface.

7. CONCLUSIONS AND FUTURE PERSPECTIVES

High demands and attention of modern consumers on the quality of their purchases require detailed information about external and internal quality of fresh products. Research is working on rapid, precise and low-cost techniques for food analysis. Among the innovative techniques qualified for this task, image analysis by computer vision systems (CVS) have proved to be effective and suitable for application at industrial level. Different types of CVS have been developed for quality inspection of fruit and vegetables, based on conventional (visible), multispectral and hyperspectral imaging. The solutions based on conventional imaging (CVS-CI) use RGB colour cameras that are sensible to the visible wavelengths of the electromagnetic spectrum. They are already widely used to measure colour, size, shape, and to detect some external defects. Moreover, their potentiality in determining internal characteristics related to the product nutritional quality have been proved, providing added value to any actor of the supply chain of fruit and vegetables. Unfortunately, as limitation, the only internal quality traits that this technology is able to evaluate are strictly related to external visible changes in terms of physical alteration or colour modification.

This review has presented the most recent and significant developments in computer vision technology for the quality evaluation of fresh fruit and vegetables. A large part of the presented studies used CVS-CI to sort and grade fruit and vegetables on the basis of external defects and colour changes, improving the productivity at industrial level and providing products with high and homogeneous quality to consumers.

Online sorting systems along the whole supply chain could allow the inspection of large quantities of fruit and vegetables in a short time and could provide a good prediction of the external and internal quality of the batch of inspected products. Moreover, the continuous and pervasive monitoring of quality products in addition of their physiological state during postharvest, can enable longer transportation times in order to reduce food losses along the entire supply chain. There are still many open challenges to realize quality assessment in a faster and accurate way in industrial line and during postharvest phase: stem-end calyx recognition, accounting for the distribution of lightness on curve surfaces, whole surface inspection, new and simpler statistical methodologies for algorithms construction, etc. Another big challenge is to weaken the requirements about the environment at acquisition time: current technologies are suitable for industrial lines where the environment can be controlled, but their transfer to less structured environment (logistic, great distribution, customers, …) often still requires custom adaptation that can require significative work. Advanced solutions in which smarter algorithms can cope with the variations induced by environment conditions could simplify a wide spread of these technologies wherever quality control can be helpful. A promising contribution can come from the convergence between computer vision systems and machine learning methodologies, especially in the two most critical steps of computer vision systems: features extraction/ selection and classification/regression on the base of the identified features. Proper learning paradigms can empower and simplify both these tasks: learning from properly built sets of examples both the relevant features and the free parameters of models for classification and regression can reduce the effort for moving a CVS to different products or to the estimation of different parameters. Several researches have proved that a properly designed architecture of CVS-CI can be applied to different products and to the evaluation of different characteristics without changing its architecture but just by changing the samples used for the training phase. Even if machine learning techniques may have significant computational costs and often produce models that are not easily understandable by humans, this convergence promises a strong simplification in the design of applications of CVS-CI to different contexts. It can also decrease the level of expertise required to develop new applications, hopefully up to enabling final users to setup, configure and manage the systems.

Results of the reported literature could have a significant impact on advanced applications of the present vision systems. Portable diagnostic systems can be used also on field, for a complete and non-destructive analysis of the physiological state of the crop and to identify the correct maturity stage at harvest and postharvest during storage and distribution. Integrated systems installable inside supermarkets and household refrigerators can help to reduce food losses and preserving the safety of consumers. Finally, tools at smartphone level could support consumers in the verification of fresh and packaged fruit and vegetables at purchase time.

TABLES

Table 1. Applications of computer vision system based on conventional imaging (CVS-CI) for the evaluation of external defects, colour changes and internal chemical-physical attributes of fresh and packaged fruit and vegetables.

Products	Objective	Classification methods	Main results as Accuracy, Achievements and Model Parameters	Reference
		EVALUATION OF EXTERNAL DEFECTS		
Apple	Two types of classification: - 2 categories (healthy or defected) - 3 quality rank (first rank, second rank and rejected ones)	- Multi-layer perceptron (MLP) - Support vector machine (SVM) - K-nearest neighbor (KNN)	SVM accuracy of: - 92.5 % for the recognition of the 2 categories - 89.2 % for the assessment of 3 quality ranks	Moallem et al. (2017)
Blueberry	Classification in: - 4 classes (good blueberries, shrivelled, decayed and mechanically damaged berries) - 2 fruit orientations (stem-end and calyx-end)	- Linear discriminant analysis (LDA) - support vector machine (SVM)	- Identification of fruit orientation in 96.8 % of the cases Performances of - 97 % for the evaluation of fungally decayed - 93.3 % for shrivelled fruits - 86 % for mechanically damaged berries	Leiva-Valenzuela and Aguilera (2013)
Mango	- Automatic defect identification - Maturity detection	Pre-processing by three algorithms	- Rotten fruit with higher value than the threshold - High fruit quality with lower value than the threshold	Sahu and Potdar (2017)
Mango and apple	Classification of mango and apple into 2 classes (anthracnose or normal mango and bull-eye rot or normal apple)	Graphical user interface (GUI)	- Defect or decay on the surface of fruits identified as dark patches or spots - GUI showed the position of the defect and the total number of defects detected	Ali and Thai (2017)
Orange	Detection of several types of defects (insect injury, wind scarring, thrips scarring, scale infestation, canker spot, dehiscent fruit, copper burn, phytotoxicity)	Processing by a novel algorithm	Performance of: - 93.8 % in stem end classification from defective orange - 91.9 % in detection of individual defects - 97 % in detection of defective samples	Rong et al. (2017)

	Automatic orange classification into 3 groups (grade A, B and C)	Back propagation neural network	- The grading accuracy of the method reached 94.38 % - The grading accuracy of the grade A reached 100 %	Chen et al. (2018)
Eggplant	Classification of 4 categories of healthy or unhealthy eggplant (healthy, partially defected, moderately defected and unhealthy)	K-nearest neighbour	Accuracy of 88 %	Akter et al. (2017)
Tomato	Classification into defective or non-defective and ripe or unripe vegetable fruit	Artificial neural network	Accuracy of - 100 % in defective/non-defective task - 96.47 % ripe/unripe task	Arakeri and Lakshmana (2016)
Fresh-cut spinach	Detection of 4 grades, from good to bad	- K-nearest neighbour (KNN) - Support vector machine (SVM) - Back-propagation artificial neural network (BPNN)	Same results of BPNN and KNN models with a classification accuracy of 85.42 %	Huang et al. (2019)
EVALUATION OF COLOUR CHANGES				
Banana	Classification in 7 ripening stages using colour, development of brown spots and image texture information	- Simple linear correlation - Sequential forward selection	Accuracy of 98 % in the classification based on colour parameters, development of brown spots and image texture information	Mendoza and Aguilera (2004)
	Classification in 3 main categories of physiological maturity stage before harvest (under-mature, mature and over-mature)	- Mean colour intensity algorithm - Area algorithm	To classify under-mature banana, accuracy of - 99.1 % for the mean colour intensity algorithm - 85 % for the area algorithm - unsuccessful accuracy to distinguish mature and over-mature category	Prabha and Kumar (2015)
	Evaluation of peel browning during storage by a 9-point ranging scale (from 1, least browned, to 9, most browned)	- Changes in RGB colour value - Changes in the CIE Lab colour values	Higher correlation coefficient with sensory test (over 0.9) of G and CIE Lab than CIE Lab values measured by the colorimeter	Cho et al. (2016)
Persimmon	Classification into 3 commercially maturity stages (unripe, ripe and overripe)	- Linear discriminant analysis (LDA) - Quadratic discriminant analysis (QDA)	QDA accuracy of 90.24 %	Mohammadi et al. (2015)

Plum	Grading into 4 maturity stages (green mature, colour break, full colour development and over-ripe)	Multi attribute decision making theory	- Colour as the best factor for grading - Lower error percentage (2.4 %) between the manual and the calculated values of length and width	Kaur et al. (2018)
Pomegranate	Classification into 3 categories (good, moderate and weak)	- Performances Artificial neural network (ANN) - Adaptive neuro fuzzy inference system (ANFIS) - Response surface methodology (RSM)	Accuracy of - 98 % for ANN model - 95.5 % for ANFIS model - 75.5 % for RSM model	Fashi et al. (2019)
Table grape	Assessment of 5 quality levels (from QL5 - higher fully marketable qualities to QL1 – waste) during cold storage	Random forest model	Accuracy of - 100 % on cultivar Italia - 92 % on cultivar Victoria in separation of the highest QL5 and QL4 from the others (QL3, QL2 and QL1)	Cavallo et al. (2019)
Borlotto beans	Definition of an algorithm to measure the colour of red spots and seed ground colour	Multiple regression	- Change of seed ground colour as the most effective feature - A positive regression between the sensorial seed's score and a^* value - A negative regression between the sensorial score and the Hue angle	Amodio et al. (2011a)
Rocket leaves	Classification of rocket leaves into 5 quality levels (QL) and discrimination of the cultivation approach	Random forest model	Accuracy of - 95 % in QL assessment - 65–70 % in the discrimination of the cultivation approach	Palumbo et al. (2021)
Tomato	Identification of 4 grading categories and development of a calyx detection algorithm	- Gray-level coccurrence matrices (GLCM) - Support vector machine-radial basis function (SVM-RBF) - Artificial neural network (ANN) - Random forest	SVM-RBF accuracy of - 95 % for the calyx detection - 98 % in the detection of defected regions - 97 % in the identification of 4 categories	Ireri et al. (2019)
Fresh-cut apple	Browning evaluation on 3 fresh-cut apple varieties	Normalization	Significant accuracy in the detection browning	Subhashree et al. (2017)

Fresh-cut artichoke	Quality assessment by colour detection from whole quarter surface and from the browned areas	Standard regression	- Highest correlation coefficients (from 0.90 to 0.92) for L^* measured on the whole quarter surface - Lower correlation coefficients (from 0.77 to 0.91) for L^* and b^* measured on brown areas	Amodio et al. (2011b)
Fresh-cut lettuce	Quality level assessment through plastic bags	- Convolutional neural network - 3-nearest neighbours method	- Irrelevant performance loss due to the presence of packaging - 83 % classification accuracy on packaged lettuce - 86 % accuracy on product without packaging	Cavallo et al. (2018)
fresh-cut nectarine	Classification into 5 quality levels (from 5, excellent, to 1 inedible)	Correlations between visual appearance and colour features extracted	- Higher correlations between visual appearance and colour parameters b^* and Chroma measured by CVS-CI ($R^2 = 0.76$) - Lower correlations obtained using a colorimeter ($R^2 = 0.57$)	Pace et al. (2011)
Fresh-cut potato	Classification and prediction of browning	- Support vector machine (SVM) - Partial least squares regression (PLSR)	- Accuracy of 96 % to predict the storage life of the product by the SVM model PLSR accuracy of - 86 % for L^* - 96 % for a^* - 99 % for b^*	Hongyang et al. (2020)
Fresh-cut radicchio	Evaluation of 5 quality levels (QL) (from 5, very good, to 1, very poor) during the cold storage	Dendrogram based on the Euclidean distance	Best discrimination of the QLs in 3 quality classes (high, middle and poor) by colour parameter a^* over the W and the percentage of W2	Pace et al. (2015)
		EVALUATION OF CHEMICAL-PHYSICAL ATTRIBUTES		
Banana	Evaluation of peel browning during storage by laboratory measurement of polyphenol oxidase (PPO) and peroxidase (POD) enzyme activity	Non-linear mathematical models genetic programming	No significant differences of correlation coefficients between predicted values (0.98) and measured ones (0.97) of PPO and POD enzymatic activity	Nadafzadeh et al. (2018)

Grapes	Estimation of - soluble solids content (SSC) - pH	- Root mean square error of cross-validation (RMSECV) - Root mean square error of calibration (RMSEC)	RMSEC values of - 0.09 for SSC prediction - 0.77 for pH prediction RMSECV values from - 0.09 to 0.10 for pH prediction - 0.79 to 0.80 for the SSC prediction	Xia et al. (2016)
Kiwifruit	Estimation of - titratable acidity (TA) - vitamin C from photos taken by a smartphone camera	- R to B (Central R/B) ratio - B to G (Central B/G) ratio of the central site of kiwifruit	- Negative correlation between central R/B value and TA, vitamin C contents and firmness - Positive correlation among central R/B value with soluble solids content, total soluble sugars and total plate counts	Li et al. (2022)
Mandarin	Estimation of - Total soluble solids (TSS) - pH	Polynomial transformation to convert RGB to $L^* a^* b^*$ colour space - Linear regression - Partial least square regression (PLSR) models	- Good correlations (0.88 for TSS and 0.45 for pH) between measured physicochemical properties and the changing peel colour of the fruits - High accuracy ($R^2 = 0.97$) with PLSR	Hadimani and Mittal (2019)
Mango	- Classification into 3 grades (green ripe, fully ripe and over ripe) - Prediction of hardness and total soluble solids (TSS)	- Support vector classification (SVC) - Support vector regression (SVR)	To classify mango fruits into 3 grades, SVC accuracy rates of - 98.75 for the training set - 97.5 % for the prediction set High SVR correlation coefficients for - hardness (0.9051 in training sets and 0.8897 in prediction sets) - TSS (0.9515 and 0.924 in training and prediction sets, respectively)	Huang et al. (2018)
Orange	pH estimation of 3 orange varieties	- Hybrid artificial neural network-particle swarm optimization (ANN-PSO) - Multilayer perceptron (MLP) neural network	High accuracy of MLP ($R^2 = 0.95$) in orange pH value prediction as mean value of the 3 orange varieties	Sabzi et al. (2020)
Plum	Prediction of total soluble solids content	Multi attribute decision making theory	Strong associations between total soluble solids content with mean intensity of green colour ($R^2 = 0.997$) and R/G ratio ($R^2 = 0.846$)	Kaur et al. (2018)

Strawberry	Prediction of titratable acidity (TA)	- Threshold method for the segmentation of strawberries - Correlation matrices based on the Pearson correlation coefficient - Hierarchical clustering procedure by Euclidean distance and Ward's method	Significant correlation between TA and the image data (Pearson correlation coeffient is about 1)	Palumbo et al. (2022a)
	Estimation of - total soluble solids (TSS) - pH	- Multiple linear regression (MLR) - Support vector machine regression (SVM-R) models	SVM-R accuracy of - 84.1 % and 79.2 % for TSS in training and testing test - 78.8 % and 72.6 % for pH in training and testing test	Basak et al. (2022)
Fresh rocket leaves	Total chlorophyll content prediction	Random forest regression	- Accuracy of the random forest regression of $R^2 = 0.90$ - Accuracy of a common SPAD-meter of $R^2 = 0.79$	Cavallo et al. (2017)
Pigmented carrot	Prediction of - Antioxidant activity (AA) - Total phenols (TP)	Multivariate model	- Good correlation between AA and TP predicted levels and the real AA and TP measurements ($R^2 = 0.97$ and $R^2 = 0.94$, respectively) for both internal and external parts of carrots - Lower determination coefficients ($R^2 = 0.93$ for AA and $R^2 = 0.86$ for TP) for the internal part	Pace et al. (2013)
Fresh-cut rocket leaves	Prediction of - Total chlorophyll content - Ammonia content through the packaging	- Random forest regression - Partial Least Square (PLS) models	Pearson's linear correlation coefficient of - 0.84 and 0.91 to predict chlorophyll and ammonia content on packaged products - 0.86 for chlorophyll and 0.92 for ammonia on unpackaged products PLS models accuracy of - 0.70 for destructive methods - 0.77 for CVS-CI through packaging - 0.80 for CVS-CI without packaging	Palumbo et al. (2022b)

Table 2. Technical aspects of computer vision system based on conventional imaging (CVS-CI) applied to fruit and vegetables by digital cameras using as software Matlab, Microsoft Visual Basic, Phyton, R, or C++

Background	Lights	Features extracted	Reference
		MATLAB	
Black	4 fluorescent lamps (60 cm long) at 35 cm above the sample and at an angle of 45 ° with the sample	Colour: - brown spots as a percentage of the total area (% BSA) - number of brown spots per cm2 of surface (NBS/cm2)	Mendoza and Aguilera (2004)
	Ambient light	Colour (*a** and hue angle)	Amodio et al. (2011a)
	4 Fluorescent 15W lamps	- colour (lightness, redness, and yellowness component) - browned area (L_{brown}, a_{brown}, b_{brown}) - browning percentage	Amodio et al. (2011b)
	8 halogen lamps (divided along two rows placed at the two sides of the imaged area) at a 45° angle	Colour (L, a, b, Chroma, hue angle)	Pace et al. (2011)
	Eight halogen lamps (placed in two lines)	Colour features (L, a, b, hue angle)	Pace et al. (2013)
	8 halogen lamps (divided along two rows placed at the two sides of the imaged area) at a 45° angle	Colour features: - white (W) and red (R) pixels, and their subdivision into two components (W1 or R1, dark) and (W2 or R2, light) - average value of L^*, a^* and b^* - percentage of L^*, a^* and b^* with respect to the product's surface	Pace et al. (2015)
	Darkness	Colour features (R, G, B, L, a and b)	Cho et al. (2016)
	Eight halogen lamps (placed in two lines)	Colour features	Cavallo et al. (2017)
	Four lamps that contained two fluorescent tubes	Computation of: - image moment - shoulder fullness	Sahu and Potdar (2017)
	Eight halogen lamps (divided along two rows placed at the two sides of the imaged area)	Colour features (complete histogram in the a* b* plane)	Cavallo et al. (2018)

	Natural light	- Colour - texture - size	Kaur et al. (2018)
	8 halogen lamps (divided along two rows placed at the two sides of the imaged area) at a 45° angle	2 set of features: - by statistical measures evaluated over the whole foreground on the channels in the CIELAB colour space - by a centroid-based colour segmentation algorithm	Cavallo et al. (2019)
	4 white light LED strips	Colour features (from L, a and b values citrus colour index, CCI)	Hadimani and Mittal (2019)
	Strip LED	Colour features (gray mean and standard deviation of each channel R, G and B)	Hongyang et al. (2020)
	LED, Fluorescent and Tungsten lamps	5 different types of features: - 336 colour - 80 texture - 6 histogram - 10 moments - 20 shape	Sabzi et al. (2020)
	8 halogen lamps, placed along two rows at the two sides of the imaged area, at 45 °	Colour features (complete histogram in the a* b* plane of the foreground pixels)	Palumbo et al. (2021)
	2 led ramps with 120 lamps	Colour features	Palumbo et al. (2022a)
	Eight halogen lamps (placed in two lines)	Colour features (complete histogram of the foreground pixels, expressing the number of occurrences of each colour in the a* b* plane)	Palumbo et al. (2022b)
Dark	Ambient light	- 8 colour features (mean and standard deviation of RGB and hue (H) component) - 5 textural features (average of gray level cooccurrence matrices, contrast, correlation, energy, homogeneity and entropy) - 3 geometric features (defect ratio, defect perimeter and defect medial axes length)	Moallem et al. (2017)
	Ambient light	- Colour (mean, standard deviation, and range) - texture (contrast, correlation, energy, homogeneity and entropy) - shape (centroid)	Ireri et al. (2019)

White	4 fluorescent lamps at 40 cm above the samples with 45° angle	- Colour features (mean, variance, skewness, kurtosis) - size features (area, perimiter, lenght and width)	Prabha and Kumar (2015)
	Ambient light	Textural features	Ali and Thai (2017)
	Two types of halogen lamps 50W	17 colour features (R, G, B, r, g, b, H, S, I, and 8 equal areas of hue range (0-360°, each of these areas was 45°)	Nadafzadeh et al. (2018)
	40 LED lights	15 features (minimum, mean, maximum, standard deviation, mode and correlation coefficient, range of variations, skewness, elongation, entropy, variance, median, harmonic mean, covariance, and contrast)	Fashi et al. (2019)
	Ambient light	- 6 colour features from RGB space - 28 colour features from other colour spaces (HIS, NTSC, YCbCr, HSV and CMY) - 18 colour features from arithmetically calculated images (the ratio and normalized operations between the red, green, and blue channel images)	Xia et al. (2016)
Blue	LED lights	- 18 colour features (R, G, B, H, S, V, L, a, b, R_δ, G_δ, B_δ, H_δ, S_δ, V_δ, L_δ, a_δ, b_δ)	Huang et al. (2019)
Not defined	Ambient light	- 6 standard features (the mean, standard deviation and mean first and second derivative along the boundaries of the region of interest) - 16 Fourier descriptors - 413 local binary pattern features - 469 Gabor features (frequency and orientation of textures for the differentiation of classes) - 27 features from the grayscale images - 7 Hu moments - 4 Flusser and Suk moments	Leiva-Valenzuela and Aguilera (2013)
	4 LED and 2 fluorescent lamps	4 standard colour features (mean, maximum, minimum, and standard deviation)	Mohammadi et al. (2015)
	Fluorescent lights	- colour features (colour mean, standard deviation and skewness) - texture features (contrast, correlation, energy and homogeneity)	Arakeri and Lakshmana (2016)
	Fluorescent lamps	- 4 colour features (normalized L* values, hue angle, browning index, and total colour change). - 3 textural features (entropy, contrast and homogeneity)	Subhashree et al. (2017)

	Incandescent lamps	4 features (fruit surface colour, size, surface defect and shape)	Chen et al. (2018)
	Light system	- 6 color features (hue, intensity, saturation, L, a, and b) - 4 texture features (consistency, angular second moment, entropy, and correlation)	Huang et al. (2018)
		MICROSOFT VISUAL BASIC	
Dark	Ambient light	- Size - shape - colour - percentage of diseased area	Akter et al. (2017)
		PYTHON	
Black	Two strips of light-emitting diodes (LEDs)	Colour features (RGB, HSV and HSL colour spaces)	Basak et al. (2022)
		R	
Not defined	Cold white light	Colour features	Li et al. (2022)
		C++	
Blue	Six white fluorescent tubes	Colour features	Rong et al. (2017)

Table S1. Additional information about methodology, results evaluation and validation tests adopted by authors cited in the review

Methodology	Results evaluation	Validation test	Reference
Median filtering used to remove noise from the RGB image. The diseased effected areas of the eggplant were segmented using Otsu and binary transformation methods	Calculation of the pixel volume of defected area and the total size of the eggplant to grade an eggplant according to its amount of defected area. The percentage of defected area determined	The collected eggplant images divided into two groups: training samples and testing samples. When the feature came from the testing sample, it was compared with the feature values of database to make decision about classification and grading	Akter and Rahman (2017)
Conversion of colored RGB image into grayscale intensity image and then into binary image. Canny edge detection method for the detection of the edge of the fruits image to extract their boundary. GUI interface to fill and show dark patches or holes in in the binary image	Association of any dark spots upon the analyzed image to a defect or decay on the surface of fruits. The graphical user interface (GUI) showed the position of the defect with a red circle and identified it using a unique number.	/	Ali and Thai (2017)
Conversion RGB into HSV space, double segmentation, conversion of RGB into Lab values, classification	Excel data	A calibration and a validation set of data containing 126 images each (representing every treatment and storage duration), generated across the two experiments	Amodio et al. (2011a)
First segmentation, second segmentation to obtain brown area from the samples, conversion from RGB into Lab values	/	A calibration and a validation set of data containing 126 images each	Amodio et al. (2011b)
/	Loading of the image into the graphical user interface developed for easy usage of the system. Tomato analyzed to recognize the defects upon clicking the button “check for defective”. The conveyor belt moved tomato to the respective bins after classifying it as ripe or unripe and defective or no-defective.	The leave-one-out method used to train and test the classifier. The average accuracy of n iterations used to estimate the accuracy of the classifier.	Arakeri and Lakshmana (2016)

Conversion of each RGB image to HSV (hue, saturation and value) and HSL (hue, saturation and lightness) colour spaces. Data pre-processing techniques performed to measured data. Application of a rank correlation (Pearson correlation technique). A different range of variables in a same dataset used as an input parameter for developing machine learning models. Application of the Z-score data normalization technique before the classification step.	/	The study utilized 80 % data during the training stage and 20 % data during the testing stage	Basak et al. (2022)
Image segmentation, features extraction, dataset creation and classification	/	Stratified k-fold cross validation (and in particular 10-fold cross validation) performed to assess the performance of machine learning techniques	Cavallo et al. (2017)
Correction of colours, segmentation, selection of the pixels belonging to the samples (packaging artefacts, reflections elimination), features extraction and classification	Classification accuracy was used to evaluate the predictive model	Estimation of performance using a Leave-1-Out Validation method	Cavallo et al. (2018)
Authomatic features extraction by the use of a random forest classifier. The machine learning tool selected authomatically how many and which attributes were better to accomplish the classification task	/	A nested cross-validation used: the internal cross-validation used for tuning split the training data used by the outer cross-validation	Cavallo et al. (2019)
The image preprocessing divided into four parts: RGB three-component selection, image smoothing, image sharpening and image segmentation.	Results obtained by the traditional artificial method of Support vector machine compared to the new classifier BP neural network.	80 orange images with four outputs as training samples	Chen et al. (2018)
Firstly, Changes in RGB colour value of browning degree of banana peel analysed for each channel (R, G, and B). Secondly, changes in the CIE Lab colour values analysed using algorithms associated with colour space conversion based on RGB colour space	/	/	Cho et al. (2016)

Convertion of each image into different color spaces (RGB, CMY, Gray, BW, HSV, I1I2I3, Lab, NNGNb, YCbCr, YCrCgCb, YIQ, and YUV); then, conversion of images into monochrome environments of various channels and their histograms drawn. When the best threshold was obtained, the background was separated from the fruit using an algorithm.	/	30 % of the data used for testing and the rest used for training.	Fashi et al. (2019)
Image segmentation, features extraction, and classification	/	The dataset of 271 Kinnow fruits randomly divided into training set of 180 fruits and test set of 91 fruits.	Hadimani and Mittal (2019)
Image prediction (to reduce the noise), background segmentation, channel segmentation and conversion and features extraction	/	The proportions of calibration set and cross-validation set for modeling were 3:2 and 2:1, respectively.	Hongyang et al. (2021)
Image segmentation, features extraction, and classification	/	For every grade, a random sample partition method to classify the 2/3 model samples into the training set, and the remaining ones classified into the prediction set	Huang et al. (2018)
Extraction of leaf area of the spinach as the region of interest. In the segmentation of the intact spinach, extraction of the back 2/3 of the segmented region to reduce the area to processed. Morphology and regional difference set operation used to realize the complete segmentation of the leaf area	/	Training model contemplated the separation of samples' data into two parts: the training set and the prediction set. At a ratio of 2 to 1, selection of 96 out of 144 samples into training set and putting the 48 remaining samples into the prediction set	Huang et al. (2019)
Segmentation, calyx and stalk scar detection and defect segmentation, extraction features, and classification	/	The image acquisition system captured a total of 2000 images. In each grading category, 70 % of the dataset used as a model training dataset while the 30 % used as the testing dataset	Ireri et al. (2019)

Extraction of R, G, and B channels from the RGB images.The image noises and marginal lines removed. The image holes filled to complete the carrot shape. Selection of efficient features of carrot shapes using the quadratic discriminant analysis	/	An algorithm programmed in MATLAB 2012a software using cross-validation method based on quadratic discriminant analysis.	Jahanbakhshi and Kheiralipour (2020)
Segmentation, features extraction and classification	Graphical user interface was designed	25 images used for calibration purpose and 10 images for validation purpose	Kaur et al. (2018)
Two steps image segmentation: 1) recognition of single berries by cropping the original image in pre-defined regions; 2) building a binary mask separating the fruit from the background using a threshold. After segmentation, color images decomposed into RGB channels and converted to grayscale and in the CIELab color space, producing seven intensity images (1640 images measuring 100 x 100 pixels analyzed using pattern recognition algorithms).	The performance of the classifiers measured as the ratio of correctly classified images with respect to the total number of tested images.	The validation performed using a 10-fold stratified cross-validation technique, yielding an average estimate of classifier performance with 95 % confidence intervals for the classification pool. In the cross-validation, 90 % of the samples used for training and 10 % used for 10 validation replications.	Leiva-Valenzuela and Aguilera (2013)
RGB values of different fruit parts (head, central and mesocarp) captured using the ColorPicker app	/	/	Li et al. (2022)
Pre-processing of the digital images using a linear Gaussian low pass filter to pre-smooth the noisy images and improve their quality. After segmentation, determination of a linear transformation between RGB signals and a device independent system such as CIEXYZ. The calibrated color data in CIEXYZ converted to CIELab color space.	/	/	Mendoza and Aguilera (2004)
Multi-layer perceptron (MLP) neural network used for segmentation (background removal, stem end and calyx detection and primary defects segmentation)	/	8 first rank, 8 second rank and 8 rejected apple images as test set of apple grading into three quality categories	Moallem et al. (2017)

The whole image analysis process included filtering, binaryzation, reversing, erosion and removing black spots	/	/	Mohammadi et al. (2015)
Pre-processing operations (equalizing of the image histogram and the removal of noises); fruit area extracted by performing the method of Otsu and color features extracion	/	/	Nadafzadeh et al. (2018)
Polynomial colour transformation to correct the images before any processing. Conversion of the colours of the to the CIELab space. Building of a two dimensional histogram. Each cell of this histogram associated with a colour (identified by its coordinates a* and b*) and a mass expressing the number of pixels with that colour in the image	/	/	Pace et al. (2011)
Polynomial colour transformation to correct the images before any processing. Conversion of the colours of the to the CIELab space. Building a two dimensional histogram. Each cell of this histogram associated with a colour (identified by its coordinates a* and b*) and a mass expressing the number of pixels with that colour in the image	/	60 carrots were used to obtain the prediction models and 24 carrots for their subsequent validation	Pace et al. (2013)
Authomatic segmentation: color-chart detection, foreground extraction and color segmentation for features extraction and selection (first and second segmentation were performed)	/	/	Pace et al. (2015)
Colour chart and foreground detection, colour correction comparing three colour transformations, features extraction and classification	/	A 10-fold cross validation approach used. According to the 10-folds validation strategy, the training was done 10 times.	Palumbo et al. (2021)

The segmentation of the strawberries carried out by a threshold method. Application of a morphological filter on binary images, to erode the strawberry edge. A secondary segmentation performed to individuate the green and red areas. An enhanced image obtained by subtracting the G image to R image. The R image thresholded and a secondary mask obtained to show the red area. The green area obtained by subtracting the secondary to the primary mask. The pixel count of each area performed to calculate the red and green percentage area.	The pixel count of each area performed to calculate the red and green percentage area and correlated to chemical parameters related to maturity stage	/	Palumbo et al. (2022a)
Colour chart and foreground detection, colour correction by a linear transformation, selection of the pixels belonging to the samples (packaging artefacts, reflections elimination), features extraction and classification	Mean square error and Pearson's correlation coefficient between the estimated and the true values measured in the lab used to quantify the performance of regression	To assess the performances, in terms of R^2, of predictors estimated , calibration and validation tests (30 % of the data) performed	Palumbo et al. (2022b)
Image segmentation, colour and size features extraction, data analysis and classification	The calibration image dataset used to fix threshold values for developing an algorithm. Evaluation of the accuracy of the algorithms using both calibration images and validation images.	A total of 120 images obtained for analysis: 60 images selected as calibration images which were used to develop the algorithm and remaining 60 selected as validation images	Prabha and Kumar (2015)
The image process consisted of several main steps: background removal, image binarization with a sliding comparison window, image subtraction, hole filling and identification and removal stem-end pixels	Application of an algorithm on a set of 1191 images	The performance result based on a personal computer was 0.108 s per image	Rong et al. (2017)
Image segmentation, features extraction, suboptimal feature selection and classification	/	70 % training set, 15 % validation set and remainder 15 % test set	Sabzi et al. (2020)

Developing of an algorithm-1 for pre-processing of the database; conversion of a colour image into a grayscale and binary image. Algorithm-2 proposed for identification of defects by calculation quality ratio b=a/(x*y), where b was the area of defected region and (x*y) pixel value. Detection of the maturity of the harvested mango by Algorithm-3 based on the value of matrix difference maturity of mango	The proposed algorithm tested using 28 mango images and it efficiently and accurately determined the quality of mango: if the value of the quality ratio was greater than the threshold value, the fruit was rotten, while if the value was less than the threshold value, the fruit was good.	/	Sahu and Potdar (2017)
Calibration of CVS, image acquisition, segmentation, cropping and extracting features	/	80 % of images acquired used for training, 10 % for testing, 10 % for validation	Subhashree et al. (2017)
Image segmentation, features extraction, calibration algorithms and model evaluation standard	/	Execution of the calibration model by a cross-validation process. Within each process of cross-validation, few samples left-out while the remaining samples used to establish the calibration model. The predicted values of the left-out samples used for the calculation of prediction residuals.	Xia et al. (2016)

REFERENCES

Aaby, K., Skaret, J., Røen, D., Sønsteby, A., 2019. Sensory and instrumental analysis of eight genotypes of red raspberry (*Rubus idaeus* L.) fruits. *Journal of Berry Research*, 9(3), 483-498.

Akter, Y. A., Rahman, M. O., 2017. Development of a computer vision based eggplant grading system. In IEEE (Ed.), 4th International Conference on Advances in Electrical Engineering (ICAEE). IEEE, Dhaka, Bangladesh, pp. 285-290.

Ali, M.A., Thai, K.W., 2017. Automated fruit grading system. IEEE 3rd International Symposium in Robotics and Manufacturing Automation (ROMA), 1-6.

Amodio, M.L., Cornacchia, R., Colantuono, F., Colelli, G., 2011a. Color Degradation Kinetics of Rehydrated 'Borlotto' Beans Stored in Different Gas Atmospheres as Measured by Image Analysis. *Journal of Agricultural Engineering*, 42(4), 33-40.

Amodio, M.L., Cabezas-Serrano, A.B., Peri, G., Colelli, G., 2011b. Post-cutting quality changes of fresh-cut artichokes treated with different anti-browning agents as evaluated by image analysis. *Postharvest Biology and Technology*, 62(2), 213-220.

Amodio M.L., Cabezas-Serrano A.B., Rinaldi R., Colelli G., 2007. Implementation of rating scales for visual quality evaluation of various vegetable crops. In: Kader, A.A., Cantwell, M. (Eds), *Produce Quality Rating Scales and Color Charts*. Postharvest horticulture series n. 23, Univ of California, Davis CA-USA.

Arakeri, M.P., Lakshmana, 2016. Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry. *Procedia Computer Science*, 79, 426-433.

Bagri, N., Johari, P.K., 2015. A comparative study on feature extraction using texture and shape for content based image retrieval. *International journal of advances in science and technology,* 80(4), 41-52.

Baranowski, P., Mazurek, W., Wozniak, J., Majewska, U., 2012. Detection of early bruises in apples using hyperspectral data and thermal imaging. *Journal of Food Engineering*, 110(3), 345-355.

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.

Basak, J.K., Madhavi, B.G.K., Paudel, B., Kim, N.E., Kim, H.T., 2022. Prediction of total soluble solids and pH of strawberry fruits using RGB, HSV and HSL colour spaces and machine learning models. *Foods* 11 (14), 2086.

Bhargava, A., Bansal, A., 2021. Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University-Computer and Information Sciences,* 33(3), 243-257.

Bianchi, T., Guerrero, L., Gratacós-Cubarsí, M., Claret, A., Argyris, J., Garcia-Mas, J., Hortós, M., 2016. Textural properties of different melon (*Cucumis melo* L.) fruit types: Sensory and physical-chemical evaluation. *Scientia Horticolturae*, 201, 46-56.

Brasil, I.M., Siddiqui, M.W., 2018. Postharvest quality of fruits and vegetables: An overview. *Preharvest modulation of postharvest fruit and vegetable quality*, 1-40.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2019. Non-destructive and contactless quality evaluation of table grapes by a computer vision system. *Computers and Electronics in Agriculture*, 156, 558-564.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2018. Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material. *Journal of Food Engineering*, 223, 46-52.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2017. Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. *Computers and Electronics in Agriculture*, 140, 303-310.

Chaudhry, M.M.A., Babellahi, F., Amodio, M.L., Colelli, G., Sahar, A., 2020. Image analysis. In: Khan, M.K.I. (Ed.), *Advances in Noninvasive Food Analysis* (pp. 200-201), CRC Press.

Chen, X., Pang, T., Tao, H., Lin, M., Yang, H., 2017. Prediction of apple internal qualities using hyperspectral imaging techniques. In: IEEE (Ed.), *International Conference on Advanced Mechatronic Systems* (ICAMechS) (pp. 450-455). IEEE, Xiamen, China.

Chen, Y., Wu, J., Cui, M., 2018. Automatic classification and detection of oranges based on computer vision. In: IEEE (Ed.), *4th International Conference on Computer and Communications* (ICCC) (pp. 1551-1556). IEEE, Chengdu, China.

Cho, J.S., Lee, H.J., Park, J.H., Sung, J.H., Choi, J.Y., Moon, K.D., 2016. Image analysis to evaluate the browning degree of banana (Musa spp.) peel. *Food Chemistry*, 194, 1028-1033.

Chopde, S., Patil, M., Shaikh, A., Chavhan, B., Deshmukh, M., 2017. Developments in computer vision system, focusing on its applications in quality inspection of fruits and vegetables-A review. *Agricultural Reviews*, 38(2), 94-102.

Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C., Huang, W., 2020. On line detection of defective apples using computer vision system combined with deep learning methods. *Journal of Food Engineering*, 286, 110102.

Farcuh, M., Copes, B., Le-Navenec, G., Marroquin, J., Jaunet, T., Chi-Ham, C., Cantu, D., Bradford, K.J., Van Deynze, A., 2020. Texture diversity in melon (*Cucumis melo* L.): Sensory and physical assessments. *Postharvest Biology and Technology*, 159, 111024.

Fashi, M., Naderloo, L., Javadikia, H., 2019. The relationship between the appearance of pomegranate fruit and color and size of arils based on image processing. *Postharvest Biology and Technology*, 154, 52-57.

Francis, G.A., Gallone, A., Nychas, G.J., Sofos, J. N., Colelli, G., Amodio, M. L., Spano, G. (2012). Factors affecting quality and safety of fresh-cut produce. *Critical Reviews in Food Science and Nutrition*, 52(7), 595-610.

François, I.M., Wins, H., Buysens, S., Godts, C., Van Pee, E., Nicolaï, B., De Proft, M., 2008. Predicting sensory attributes of different chicory hybrids using physico-chemical measurements and visible/near infrared spectroscopy. *Postharvest Biology and Technology*, 49(3), 366-373.

Giovenzana, V., Beghi, R., Buratti, S., Civelli, R., Guidetti, R., 2014. Monitoring of fresh-cut Valerianella locusta Laterr. shelf life by electronic nose and VIS–NIR spectroscopy. *Talanta*, 120, 368-375.

Goñi, S.M., Salvadori, V.O., 2017. Color measurement: comparison of colorimeter vs. computer vision system. *Journal of Food Measurement and Characterization*, 11(2), 538-547.

Hadimani, L., Mittal, N., 2019. Development of a computer vision system to estimate the colour indices of Kinnow mandarins. *Journal of Food Science and Technology*, 56(4), 2305-2311.

Hongyang, T., Daming, H., Xingyi, H., Aheto, J.H., Yi, R., Yu, W., Ji, L., Shuai, N., Mengqi, X., 2021. Detection of browning of fresh-cut potato chips based on machine vision and electronic nose. *Journal of Food Process Engineering*, 44(3), e13631.

Huang, X., Yu, S., Xu, H., Aheto, J.H., Bonah, E., Ma, M., Wu, M., Zhang, X., 2019. Rapid and nondestructive detection of freshness quality of postharvest spinaches based on machine vision and electronic nose. *Journal of Food Safety*, 39(6), e12708.

Huang, X., Lv, R., Wang, S., Aheto, J.H., Dai, C., 2018. Integration of computer vision and colorimetric sensor array for nondestructive detection of mango quality. *Journal of Food Process Engineering,* 41(8), e12873.

Ireri, D., Belal, E., Okinda, C., Makange, N., Ji, C., 2019. A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. *Artificial Intelligence in Agriculture,* 2, 28-37.

Jana, S., Basak, S., Parekh, R., 2017. Automatic fruit recognition from natural images using color and texture features. In: IEEE (Ed.) *Devices for Integrated Circuit* (DevIC), (pp. 620-624). IEEE, Kalyani, India.

Kader, A.A., Rolle, R.S., 2004. The role of post-harvest management in assuring the quality and safety of horticultural produce. *Food and Agriculture Organization of the United Nations*, vol. 152.

Kaur, H., Sawhney, B.K., Jawandha, S.K., 2018. Evaluation of plum fruit maturity by image processing techniques. *Journal of Food Science and Technology*, 55(8), 3008-3015.

Lara, M.A., Lleó, L., Diezma-Iglesias, B., Roger, J.M., Ruiz-Altisent, M., 2013. Monitoring spinach shelf-life with hyperspectral image through packaging films. *Journal of Food Engineering*, 119(2), 353-361.

Leiva-Valenzuela, G.A., Aguilera, J.M., 2013. Automatic detection of orientation and diseases in blueberries using image analysis to improve their postharvest storage quality. *Food Control*, 33(1), 166-173.

Li, H., Lv, S., Feng, L., Peng, P., Hu, L., Liu, Z., Hati, S., Bimal, C., Mo, H., 2022. Smartphone-Based Image Analysis for Rapid Evaluation of Kiwifruit Quality during Cold Storage. *Foods* 11 (14), 2113.

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.

Li, H., Deng, Z., Zhu, H., Hu, C., Liu, R., Young, J.C., Tsao, R., 2012. Highly pigmented vegetables: Anthocyanin compositions and their role in antioxidant activities. *Food Research International,* 46(1), 250-259.

Lorente, D., Aleixos, N., Gómez-Sanchis, J.U.A.N., Cubero, S., García-Navarrete, O.L., Blasco, J., 2012. Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology*, 5(4), 1121-1142.

Lou, H., Hu, Y., Zhang, L., Sun, P., Lu, H., 2012. Nondestructive evaluation of the changes of total flavonoid, total phenols, ABTS and DPPH radical scavenging activities, and sugars during mulberry (Morus alba L.) fruits development by chlorophyll fluorescence and RGB intensity values. *LWT-Food Science and Technology*, 47(1), 19-24.

Mendoza, F., Aguilera, J.M., 2004. Application of image analysis for classification of ripening bananas. *Journal of Food Science*, 69(9), E471-E477.

Moallem, P., Serajoddin, A., Pourghassem, H., 2017. Computer vision-based apple grading for golden delicious apples based on surface features. *Information Processing in Agriculture,* 4(1), 33-40.

Mohammadi, V., Kheiralipour, K., Ghasemi-Varnamkhasti, M., 2015. Detecting maturity of persimmon fruit based on image processing technique. *Scientia Horticulture*, 184, 123-128. https://doi.org/10.1016/j.scienta.2014.12.037

Nadafzadeh, M., Mehdizadeh, S.A., Soltanikazemi, M., 2018. Development of computer vision system to predict peroxidase and polyphenol oxidase enzymes to evaluate the process of banana peel browning using genetic programming modeling. *Scientia Horticulture,* 231, 201-209.

Narendra, V.G., Amithkumar, V.G., 2019. An intelligent computer vision system for vegetables and fruits quality inspection using soft computing techniques. *Agricultural Engineering International: CIGR Journal*, 21(3), 171-178.

Pace, B., Cavallo, D.P., Cefola, M., Colella, R., Attolico, G., 2015. Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio. *Innovative Food Science and Emerging Technologies,* 32, 200-207.

Pace, B., Cefola, M., Renna, F., Renna, M., Serio, F., Attolico, G., 2013. Multiple regression models and computer vision systems to predict antioxidant activity and total phenols in pigmented carrots. *Journal of Food Engineering*, 117(1), 74-81.

Pace, B., Cefola, M., Renna, F., Attolico, G., 2011. Relationship between visual appearance and browning as evaluated by image analysis and chemical traits in fresh-cut nectarines. *Postharvest Biology and Technology*, 61(2-3), 178-183.

Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Colelli, G., Attolico, G., 2022b. Non-destructive and contactless estimation of chlorophyll and ammonia contents in packaged fresh-cut rocket leaves by a Computer Vision System. *Postharvest Biology and Technology*, 189, 111910.

Palumbo, M., Cozzolino, R., Laurino, C., Malorni, L., Picariello, G., Siano, F., Stocchero, M., Cefola, M., Corvino, A., Romaniello, R., Pace, B., 2022a. Rapid and Non-Destructive Techniques for the Discrimination of Ripening Stages in Candonga Strawberries. *Foods*, 11(11), 1534.

Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Serio, F., Colelli, G., Attolico, G., 2021. Self-configuring CVS to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. *Agronomy*, 11(7), 1353.

Patel, K.K., Kar, A., Jha, S.N., Khan, M.A., 2012. Machine vision system: a tool for quality inspection of food and agricultural products. *Journal of Food Science and Technology*, 49(2), 123-141.

Prabha, D.S., Kumar, J.S., 2015. Assessment of banana fruit maturity by image processing technique. *Journal of Food Science and Technology*, 52(3), 1316-1327.

Rong, D., Rao, X., Ying, Y., 2017. Computer vision detection of surface defect on oranges by means of a sliding comparison window local segmentation algorithm. *Computers and Electronics in Agriculture*, 137, 59-68.

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deepfruits: A fruit detection system using deep neural networks. *Sensors*, 16(8), 1222.

Sabzi, S., Javadikia, H., Arribas, J.I., 2020. A three-variety automatic and non-intrusive computer vision system for the estimation of orange fruit pH value. *Measurement*, 152, 107298.

Sahu, D., Potdar, R.M., 2017. Defect identification and maturity detection of mango fruits using image analysis. *American Journal of Artificial Intelligence*, 1(1), 5-14.

Salinas-Hernández, R.M., González-Aguilar, G.A., Tiznado-Hernández, M.E., 2015. Utilization of physicochemical variables developed from changes in sensory attributes and consumer acceptability to predict the shelf life of fresh-cut mango fruit. *Journal of Food Science and Technology,* 52(1), 63-77.

Sendin, K., Manley, M., Williams, P. J. (2018). Classification of white maize defects with multispectral imaging. *Food Chemistry*, 243, 311-318.

Stavang, J.A., Freitag, S., Foito, A., Verrall, S., Heide, O.M., Stewart, D., Sønsteby, A., 2015. Raspberry fruit quality changes during ripening and storage as assessed by colour, sensory evaluation and chemical analyses. *Scientia Horticulturae*, 195, 216–225.

Subhashree, S.N., Sunoj, S., Xue, J., Bora, G.C., 2017. Quantification of browning in apples using colour and textural features by image analysis. *Food Quality & Safety*, 1(3), 221-226.

Sung, J., Suh, J.H., Chambers, A.H., Crane, J., Wang, Y., 2019. Relationship between sensory attributes and chemical composition of different mango cultivars. *Journal of Agricultural and Food Chemistry,* 67(18), 5177-5188.

Tudela, J.A., Marín, A., Garrido, Y., Cantwell, M., Medina-Martínez, M. S., Gil, M.I., 2013. Off-odour development in modified atmosphere packaged baby spinach is an unresolved problem. *Postharvest Biology and Technology*, 75, 75-85.

Watada, A.E., Ko, N.P., Minott, D. A. (1996). Factors affecting quality of fresh-cut horticultural products. *Postharvest Biology and Technology,* 9(2), 115-125.

Wu, D., Sun, D.W., 2013. Colour measurements by computer vision for food quality control– A review. *Trends in Food Science and Technology*, 29(1), 5-20.

Xia, Z., Wu, D., Nie, P., He, Y., 2016. Non-invasive measurement of soluble solid content and pH in Kyoho grapes using a computer vision technique. *Analytical Methods*, 8(15), 3242-3248.

Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., Liu, C., 2014. Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International,* 62, 326-343.

PART II: EXPERIMENTAL APPLICATIONS

OBJECTIVES

The major objective of this doctoral thesis is to develop and validate predictive models based on the use of CVS to determine the quality level and the main quality parameters of fresh rocket leaves (*Diplotaxis tenuifolia* L.). The proposed vision system is able to automatically select, without the intervention of an operator, the relevant quality traits of the product for the classification using the Random Forest as a machine learning model.

During first experiments, CVS was applied to fresh-cut rocket leaves, obtained by low-impact agricultural practices, to objectively assess its quality levels (QL) during the storage at 10 °C according to a 5 to 1 rating scale and to discriminate the fertilization levels and irrigation managements applied during the cultivation.

Subsequently, five experiments were conducted to validate the CVS in estimating internal quality traits (chlorophyll and ammonia content) related to the shelf-life loss of rocket leaves, even through the package. The performances of the CVS data on packed and unpacked samples were compared to verify its applicability along the whole supply chain, regardless the presence of packaging.

Finally, a clustering approach to identify relevant and representative colour traits and to construct simpler algorithms to predict marker parameters of the quality of rocket leaves was studied and developed.

Chapter 3

SELF-CONFIGURING CVS TO DISCRIMINATE ROCKET LEAVES ACCORDING TO CULTIVATION PRACTICES AND TO CORRECTLY ATTRIBUTE VISUAL QUALITY LEVEL

Michela Palumbo[1,2], Bernardo Pace[1], Maria Cefola[1], Francesco Fabiano Montesano[3], Francesco Serio[3], Giancarlo Colelli[2], Giovanni Attolico[4]

[1] Institute of Sciences of Food Production, CNR, c/o CS-DAT, Via Michele Protano, 71121 Foggia
[2] Department of Agriculture, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Via Napoli 25, 71122 Foggia
[3] Institute of Sciences of Food Production, CNR, Via G. Amendola, 122/O, 70126 Bari,
[4] Institute on Intelligent Industrial Systems and Technologies for Advanced Manufacturing, CNR, Via G. Amendola, 122/O, 70126 Bari

ABSTRACT

Computer Vision Systems (CVS) represent a contactless and non-destructive tool to evaluate and monitor the quality of fruits and vegetables. This research paper proposes an innovative CVS, using a Random Forest model to automatically select the relevant features for classification, thereby avoiding their choice through a cumbersome and error-prone work of human designers. Moreover, three color correction techniques were evaluated and compared, in terms of classification performance to identify the best solution to provide consistent color measurements. The proposed CVS was applied to fresh-cut rocket, produced under greenhouse soilless cultivation conditions differing for the irrigation management strategy and the fertilization level. The first aim of this study was to objectively estimate the quality levels (QL) occurring during storage. The second aim was to non-destructively, and in a contactless manner, identify the cultivation approach using the digital images of the obtained product. The proposed CVS achieved an accuracy of about 95 % in QL assessment and about 65–70 % in the discrimination of the cultivation approach.

Keywords: *Diplotaxis tenuifolia* L.; automatic configuration of the CVS; color correction models; non-destructive contactless quality evaluation; fertilization and irrigation recognition fromdigital images.

1. INTRODUCTION

Recently, there has been growing interest in contactless, non-destructive, rapid and accurate techniques for the evaluation of the quality of fruits and vegetables to replace the traditional sensory and conventional destructive methods. These methods are generally time-consuming, expensive, polluting and are not suitable for the application in an industrial line (Bhargava and Bansal, 2018; Narendra et al., 2019). Moreover, it has been observed that more than in other agri-food sectors, consumers are particularly attentive to the sustainability of the vegetable production process as an important issue influencing their perception of quality (Verain et al., 2016). Furthermore, the increasing sensibility of modern consumers toward the environmental impact of production processes has been the impetus for many researchers to develop non-destructive tools for the discrimination of production origin and agricultural practices, in order to better support the added value of the products.

Nowadays, the emerging non-destructive methods in food technology, include near infrared spectroscopy (NIR), hyperspectral imaging (HSI) and computer vision system (CVS). In relation to vegetables, most of the research have applied hyperspectral or multispectral techniques (Amodio et al., 2017; Chaudhry et al., 2018; Chaudhry et al., 2020; Løkke et al., 2013; Sánchez et al., 2013). The complexity of spectroscopy and hyperspectral imaging, both in terms of time and costs required for the acquisition and for the following processing, makes the application of these techniques more difficult in a pervasive way along the supply chain to enable a continuous monitoring of the parameters of interest. On the contrary, CVS is simpler and can hopefully exploit cameras that are already available along the path from harvest to final consumers.

Increasing interest has been observed in the last few years in CVSs to automatically evaluate several properties of different products: They involve optical instrumentation, electromagnetic sensing, and digital image processing technologies (Patel et al., 2012). This technology mimics human visual evaluation of quality, acquiring images of the whole visible surface of products. These digital images are analyzed by extracting the most discriminative colors among the large set of possible visual characteristics (such as shape, color, and defects) and processing the data through suitable regression or classification models and algorithms (Bhargava and Bansal, 2018).

Normally, human designers exploit previous experiences and use a trial-and-error process to select the features used by the classification/regression methods or a vocabulary of features out of which algorithms can extract the most effective subset. In many recent researches, CVS have been used to evaluate the quality level (QL) of fresh and fresh-cut fruits, such as table grape

(Cavallo et al., 2019), fresh-cut nectarines (Pace et al., 2011) and apples (Arivu et al., 2012). As reported by many authors, CVS have also been used to evaluate the QL, chlorophyll and ammonium content of leafy vegetables. The authors in Pace et al. (2014) demonstrated that two color features detected by the CVS were able to evaluate the QL and the ammonium content (considered an indicator of senescence) in iceberg lettuce. Moreover, an innovative and automatic procedure, applied for the quality evaluation of fresh-cut radicchio allows a self-configuration of the CVS by optimizing its performance and limiting the subjective human intervention, was reported by Pace et al. (2015). The authors in Cavallo et al. (2017) proposed a procedure to predict total chlorophyll content of rocket leaves using CVS and a machine learning model (Random Forest Regression) applied to manually selected features, obtaining higher performance (R^2 = 0.90) than the SPAD-meter (R^2 = 0.79). This work supports the relevance of the color information. The consistency of color information must be enforced using color correction methods based on the color reference provided by a color-chart inserted in the scene.

In machine learning, random forests represent an ensemble (a set) of tree predictors that can be used for both classification and regression. They exploit the principle that a group of weak learners can globally provide better results than a strong learner (Breiman, 2001) and can reduce the risk of overfitting. Therefore, several instances of the selected models (trees) are trained, and the final predictions are made by combining the outputs of the models by voting (classification) or mean (regression). Specifically, random forest consists of an extension of bagging (bootstrap aggregating) ensemble (Breiman, 1996), whereby each model is trained on a different set of training examples randomly sampled with repetition from the available data. Moreover, this method builds each tree using a randomly selected subset of the available features. This makes possible to use a quite large vocabulary of features without seriously impact on the efficiency of the method and without requiring the critical and often subjective choice of the most relevant features. The final performance of the random forest depends on the strength of the individual classifiers and on their independence from each other (Breiman, 2001).

To the best of our knowledge, there is no general agreement regarding the best method to correct the colors and make them consistent among different acquisitions: This paper compares three different color correction methods, with different power and complexity. Their performance was measured, in terms of their effects on the classification accuracy. The simplest method (white balance) provided poor performance. The two other methods (linear correction and polynomial correction) provided similar performance with the second having a greater computational complexity. The linear correction is proposed as the best trade-off between efficacy and efficiency.

Moreover, the paper proposes the complete color histogram in the CIE$L^*a^*b^*$ color space as the vocabulary of features for the machine learning Random Forest model. This represents a relevant simplification of the CVS design that do not require the designer to select the features through a cumbersome and error-prone trial and error process.

Finally, the paper proposes to apply the same innovative approach to CVS design to different tasks: to obtain a non-destructive contactless and objective evaluation of the QL of rocket leaves during storage and to identify different fertilization levels (sustainable or not) using two irrigation management approaches applied during the cultivation. To the best of our knowledge, there are no previous application of CVS to the latter task. It is relevant that the same framework can be used to solve these two different tasks without changes in the architecture of the CVS: The only difference is in the final Random Forest classification: this final phase uses the same model for the two tasks but learns proper parameters for each of them by providing different expected values as input data.

2. MATERIALS AND METHODS

2.1. Plant Material, Growing System, Water and Fertilizers Use Efficiency

Rocket (*Diplotaxis tenuifolia* L. cv Dallas) was cultivated under soilless cultivation growing system in the autumn-winter (2019–2020) period in an unheated greenhouse at the experimental farm 'La Noria' of the Institute of Sciences of Food Production (CNR-ISPA), located in Mola di Bari (Puglia, South of Italy). A randomized blocks experimental design was adopted with 3 replications; each block consisted of 4 sub-blocks, each one hosting one of the four cultivation treatments under comparison. Plastic pots, 20 per each sub-block, were filled with a 3:1 (v:v) peat (Brill 3 Special, Brill Substrate GmbH & Co., Georgsdorf, Germany): perlite (Agrilit 3, Perlite Italiana, Corsico, MI, Italy) mixture as a substrate.

Two irrigation management strategies (Timer and Sensor) and two fertilization levels (FL_1 and FL_2) were applied, following a factorial combination resulting in four agronomic treatments (Timer–FL_1; Timer–FL_2; Sensor–FL_1; Sensor–FL_2). In detail, in Timer the irrigation was empirically managed with a timer providing a fixed irrigation schedule, periodically adjusted on the basis of the amount of the drainage fraction (about 35 % according to the common practice). Whereas, in Sensor the irrigation was automatically applied through dielectric sensors (GS3, Decagon Devices, Pullman, WA, USA) based on real time measurement of the substrate volumetric water content variations, thus reflecting plant water consumption and needs and resulting in a more sustainable use of irrigation water. A 0.35 m^3 m^{-3} volumetric water content irrigation set-point was adopted, corresponding to a moisture level slightly lower than substrate maximum water holding capacity. The sensor-controlled automatic irrigation system, composed by a CR1000 datalogger and a SDM16AC/DC relay driver (Campbell Scientific, Logan, UT, USA), turned on irrigation valves based on real-time sensor readings and maintained substrate volumetric water content close to the irrigation set-point. Tensiometers (one per experimental unit) were used to monitor the substrate matric potential, which showed a mean value of -25 hPa over the growing cycle with similar values in all the experimental units. In relation to fertilization, a mix of Osmocote Exact and Osmocote CalMag, (ICL Specialty Fertilizers, Treviso, Italy) was used in the substrate in a dose of 3.75 and 1 g L^{-1}, respectively, for FL_1, while a 40 % reduced dose was provided in FL_2.

The doses of fertilizers (FL_1 and FL_2) were selected according to the standard recommendations provided in the label of the fertilizer products used in the experiment, reporting indications for "high dosage" or "low dosage", respectively. Water Use Efficiency (WUE) was calculated at crop level as yield (expressed as grams of product marketable fresh weight) per liter of applied irrigation water (De Pascale et al., 2011). Similarly, Fertilizers Use Efficiency (FUE) was calculated as grams of product fresh weight per grams of applied fertilizer.

Three harvests were carried out at 62 (H1), 104 (H2) and 132 (H3) days after sowing, respectively.

After each harvest time the fresh-cut rocket leaves were immediately transported in refrigerate conditions to the Postharvest laboratory.

2.2. Sensory Classification of Rocket Leaves Visual Quality Level during Storage

Rocket leaves at each harvest time, separated for each treatment, were selected in order to avoid damaged samples and put in 50 × 30 cm open polyethylene bags (Orved, Musile di Piave (VE), Italy) containing each one about 350 g of product. In total, 12 bags (3 replicates × 4 agronomic treatments) were prepared after each harvest and stored at 10 °C (as commonly occur in the market) for 12 days for the H1 and for 18 days for the H2 and the H3. The length of storage was defined by the number of days required to reach the lowest QL, as reported in Cavallo et al. (2019). Therefore, at a proper time during storage, the amount (about 70 g) of sample to analyze was taken from each bag and subjected to a sensory evaluation by a group of 6 panelists using the following 5 to 1 QL scale (Figure 1): 5 = very good (very fresh, no signs of yellowing, bright, dark and uniform green, no defects), 4 = good (fresh, slight signs of yellowing, light green, slight loss of texture), 3 = fair (slight wilting, moderate signs of yellowing, slight discoloration, minor defects, loss of texture), 2 = poor (wilting, evident yellowing, discoloration, severe loss of texture), 1 = very poor (unacceptable quality due to decay, severe wilting and yellowing, complete loss of texture and other evident defects). A score of 3 was considered to be the limit of marketability, while a score of 2 represented the limit of edibility. Images of rocket leaves at each QL were acquired and processed by CVS and the same samples were subjected to postharvest quality evaluation.

Figure 1. The figure shows the quality level (QL) scale used for the sensory evaluation of rocket leaves.

2.3. Computer Vision System Color Analysis

For the H1 and the H2, images of the samples of products were taken at 0, 4, 7 and 12 days, corresponding to QL from 5 to 2. For the H3, images of the samples were taken at 0, 4, 7, 12 and 18, corresponding to QLs from 5 to 1. At each acquisition, a sample of about 60 g of product was taken from each of the 12 bags prepared for that harvest (3 replications for each agronomic treatment). The 12 samples were analyzed by the CVS. Two images were acquired for each sample, by stacking randomly the leaves before each acquisition to maximize the surface seen by the CVS as reported in Figure 2. Therefore, 24 images were available at each time (2 images for each of the 3 replications for each of the four agronomic treatments).

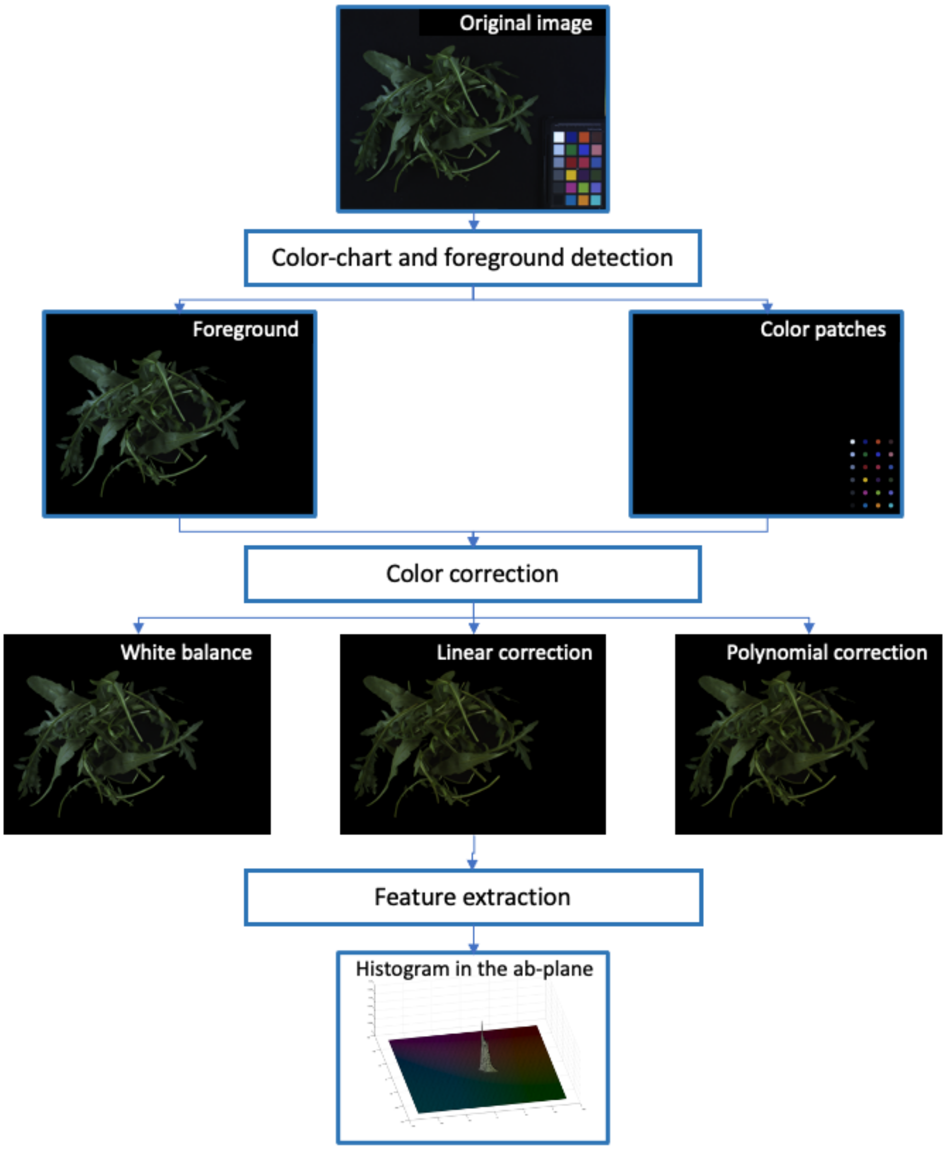


Figure 2. The figure shows the flow chart of the processing done on the images. It is possible to appreciate the effects of each step on the input image and the results provided to the following steps. Data extracted from the patches of the color chart have been used to evaluate the parameters of the three different correction models compared. For each of them, a histogram was evaluated, such as the one shown for the linear model, which provides the best trade-off between efficacy and computational complexity.

The final dataset was composed by 96 images for each of the H1 and H2, and by 120 images for the H3. The complete collection was composed by 312 images. We did not distinguish the images coming from different harvests. Therefore, the final image dataset was composed by 72 images for each quality from 5 to 2 and 24 images for quality 1. In relation to the irrigation and fertilization management, the image dataset was composed by 78 images for each combination of IS and FL. These data are reported in the Table 1.

Table 1. Composition of the training set of images with respect to harvests (H1, H2 and H3), irrigation management strategies (Timer, Sensor) and doses of fertilizers (FL_1, FL_2).

Harvest	Replication	Number of images					Total
		VQ5	VQ4	VQ3	VQ2	VQ1	
H1							
Timer-FL_1	3	6	6	6	6		***24***
Timer-FL_2	3	6	6	6	6		***24***
Sensor-FL_1	3	6	6	6	6		***24***
Sensor-FL_2	3	6	6	6	6		***24***
Total H1	***12***	***24***	***24***	***24***	***24***		***96***
H2							
Timer-FL_1	3	6	6	6	6		***24***
Timer-FL_2	3	6	6	6	6		***24***
Sensor-FL_1	3	6	6	6	6		***24***
Sensor-FL_2	3	6	6	6	6		***24***
Total H2	***12***	***24***	***24***	***24***	***24***		***96***
H3							
Timer-FL_1	3	6	6	6	6	6	***30***
Timer-FL_2	3	6	6	6	6	6	***30***
Sensor-FL_1	3	6	6	6	6	6	***30***
Sensor-FL_2	3	6	6	6	6	6	***30***
Total H3	***12***	***24***	***24***	***24***	***24***	***24***	***120***
Total H1+H2+H3	***36***	***72***	***72***	***72***	***72***	***24***	***312***

The following paragraphs will describe all the processing steps used by the CVS. All the software was developed using Matlab 2019a (Mathworks Inc., Natick, Massachusetts, United States). A flowchart of these processing steps, along with examples of their effects, is shown in Figure 2.

2.3.1. Acquisition of Calibrated Color Images

To acquire calibrated color images, color changes due to environment conditions (lighting, geometry, sensor instability) were evaluated and reduced to the minimum. Images were acquired using the set-up reported in Cavallo et al. (2017, 2018), Pace et al. (2015, 2017), using a 3CCD (with a dedicated Charged Coupled Device for each color channel) digital camera (JAI CV-M9GE) having a resolution of 1024 × 768 pixels. The imaged area is about 32 × 24 cm. A 3CCD sensor has been used to avoid the artifacts introduced by the demosaicing methods required to record color information using a single CCD. The optical axis of the LinosMeVis 12 mm lens system was perpendicular to the black background. Two DC power suppliers delivered current to eight halogen lamps, placed along two rows at the two sides of the imaged area and oriented at a 45 angle with respect to the optical axis. All the images were saved using the uncompressed TIFF format to avoid the artifacts introduced by compression algorithms.

2.3.2. Color Chart Processing and Foreground Segmentation

A small X-Rite color-chart with 24 patches of known colors was placed into the scene to measure color variations due to environmental conditions and sensor characteristics by comparing the expected numerical values released by the manufacturer with the ones acquired by the camera. The color-chart was automatically detected regardless of its position and orientation (Cavallo et al., 2017). Its white patch was used by the white-balance algorithm. All the colors in the color-chart were used to estimate the linear and polynomial transformations used for color correction.

Image processing worked only on the part of each image belonging to the product at hand (foreground). The background was discarded. The CVS automatically separated foreground and background without any human intervention: Two thresholds were derived from the analysis of the whole image in the HSV color space, as described in Cavallo et al. (2017). The segmentation was identified the region belonging to the product as a whole and did not separate its different parts, and neither discarded any region of the leaves. It was designed to be conservative, that is to discard all the background pixels even at the cost of removing some marginal borders of the product. It removed also background area inside the stack of leaves as long as part of the leaves are too dark (for example for self-shadowing of the product) to provide meaningful color information.

2.3.3. Color Correction

Color correction needs to be effective (to provide consistent color measurements) and efficient (computationally simple enough to be suitable for real applications along the supply chain). Three different color correction models, with increasing level of complexity, were compared to compensate the change in color rendering due to acquisition environment. Let it be $\left[r_e^i\ g_e^i\ b_e^i\right]^T$ and $\left[r_m^i\ g_m^i\ b_m^i\right]^T$ the expected and the measured RGB values respectively for the i-th patch i = 1,…,24. Let it be $[r_{we}\ g_{we}\ b_{we}]^T$ and $[r_{wm}\ g_{wm}\ b_{wm}]^T$ the expected and measured whites respectively. The simplest model was white balance (WB). Using the white patch in the color chart, a different correction coefficient was evaluated for each channel, as reported below (1):

$$c_r = \frac{r_{we}}{r_{wm}} \quad c_g = \frac{g_{we}}{g_{wm}} \quad c_b = \frac{b_{we}}{b_{wm}} \tag{1}$$

The three correction coefficients were used to correct the corresponding channel by multiplying the corresponding color component of each foreground pixel. A linear correction (LC) (a 3 × 3 matrix) was evaluated to reduce the distance between the expected and the measured values on the color chart (2):

$$\begin{bmatrix} r_c \\ g_c \\ b_c \end{bmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{bmatrix} r_m \\ g_m \\ b_m \end{bmatrix} \tag{2}$$

where $\begin{bmatrix} r_c \\ g_c \\ b_c \end{bmatrix}$ are the colors corrected using the matrix whose elements were evaluated using a least-square approach applied on all the patches of the color-chart. The same matrix was therefore used to correct all the foreground pixels of the image.

The last transformation was a polynomial correction (PC) (with degree 2) where all the linear and quadratic elements were considered (*r, g, b, rg, rb, gb,* r^2, g^2, b^2). The coefficients of such transformation were again evaluated using a least-square approach (3).

$$\begin{bmatrix} r_c \\ g_c \\ b_c \end{bmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} & m_{16} & m_{17} & m_{18} & m_{19} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} & m_{26} & m_{27} & m_{28} & m_{29} \\ m_{31} & m_{32} & m_{33} & m_{34} & m_{35} & m_{36} & m_{37} & m_{38} & m_{39} \end{pmatrix} \begin{bmatrix} r_m \\ g_m \\ b_m \\ r_m g_m \\ r_m b_m \\ g_m b_m \\ r_m^2 \\ g_m^2 \\ b_m^2 \end{bmatrix} \tag{3}$$

All the foreground pixels were corrected using the same matrix.

The transformations provided by the three methods were different for each image (they were evaluated from the color-chart appearance in each specific image) to adapt to the specific conditions of each acquisition.

The time required by the three-color correction methods is different. Using the MATLAB code used in the experiments, without specific optimization or the use of special hardware, the application of the white balance to an image takes 70 ms. The linear correction requires 73 ms while the polynomial correction requires 89 ms. In an industrial application of the system, the difference between linear and polynomial corrections can negatively affect the maximum speed achievable by the production. Therefore, it is important to evaluate if the performance gain justifies the loss of productivity.

2.3.4. Features Extraction

On the base of previous experiences, the device independent and perceptually uniform CIE $L^*a^*b^*$ color space was chosen to accomplish color analysis. Given that the L^* component is fragile, being too sensible to not uniform illumination levels across the scene, the complete histogram in the a^* b^* plane of the foreground pixels was used as feature set for the classification. The color histogram represents the number of occurrences of each color, that is of each (a^*, b^*) pair, in all the foreground pixels. It represents the property of the whole observed product. The continuous (a^*, b^*) plane has been discretized using 30 bins for each axis (a^* and b^*): therefore, the complete histogram was a matrix with 900 elements. This representation is more detailed than statistical measures, such as mean, median or standard deviation: it describes completely the palette of colors present in the scene and their relative relevance. The hypotheses were that such information was able to represent the appearance of new colors due to senescence as far as the effects of the cultivation management on product appearance, if any. To achieve the goal of avoiding any human intervention in the identification of proper color features, the complete matrix containing all the values of the bins of this histogram was reshaped as a vector and passed to the classification phase. The use of a quite large vector (900 elements in our case) was feasible as the ensemble method used for classification can sample for each tree a subset of features from even a quite large set. This approach automatically identifies their best use, while keeping reasonable the computational complexity. Even if it is not possible to identify few specific colors suitable to discriminate product quality or cultivation management, the ensemble of trees exploits a quite large subset of the provided features, that is (a^*, b^*) pairs, which globally achieve the desired classification.

2.3.5. Classification

Random Forest models were trained to assign the QL to the product and to identify the treatment used. The values of the cells of the histogram in the a^* b^* plane (of the CIE$L^*a^*b^*$ color space) of each image provided the vocabulary of features used for training the models. The approach for training each tree involved randomly sampling the available training data (to select the training examples) and then randomly selecting a set of features (in this case randomly selecting which values of the histogram to use to build the tree at hand). Each tree of the forest allows a maximum of 10 branches. Due to the limited number of samples, a 10-fold cross validation approach was used. The available data were divided into 10 groups (folds), each having approximately the same number of elements. The partition was made with stratification. Therefore, each group approximated the same distribution of classes of the whole training set. According to the 10-folds validation strategy, the training was done 10 times. At each round, a different fold was separated for testing the results while the other nine folds were used for training. The average of the results obtained in the ten rounds estimated the performance of the method. Accuracy is used as a quick indication of the performance of classification in the results' section but, to provide a complete description of the obtained results, the confusion matrices are provided. In fact, they provide all the information needed to describe the behavior of the method. To increase the robustness of performance measures, the 10-fold cross-validation process was repeated 20 times. The confusion matrices and accuracy values represent the average of the values over these 20 different repetitions. At each repetition, a new stratified partition of training data into 10 folds was randomly generated. That increases the significance of the obtained results by making less relevant the effects of chance in sampling training data and features. In spite of the significative number of trees in the resulting forest (200 trees were allowed for each forest) the increase in accuracy provided by their combination does not require high computational costs. The code, written in Matlab without any specific optimization, requires about 25 s for building the Random Forest model and about 0.13 s to apply the model to a new sample and to classify it.

2.4. Postharvest Quality Parameters

2.4.1. Color Analysis by Colorimeter

Color parameters ($L*$, $a*$ and $b*$) were measured, for each replicate, on 3 random points on the surface of 10 rocket leaves using a colorimeter (CR400, Konica Minolta, Osaka, Japan) in the reflectance mode and in the CIE$L*a*b*$ color scale. Colorimeter was calibrated with a standard reference having values of $L*$, $a*$ and $b*$ corresponding to 97.44, 0.10 and 2.04, respectively. To measure color variations on each sensory evaluation, ΔE* was calculated according to the following equation (4) (Martínez-Sánchez et al., 2011):

$$\Delta E* = \sqrt{(L_0^* - L^*)^2 + (a_0^* - a^*)^2 + (b_0^* - b^*)^2} \quad (4)$$

where L^*_0, a^*_0 and b^*_0 represents color parameters detected on fresh samples. Yellowness index (YI) was calculated from primary L^*, a^* and b^* readings, while the degreening index (DI) was obtained by the Hunter L a b values (obtained converting the CIEL^* a^* b^* readings), according to the following equations (5,6) (Pathare et al., 2013; Jiménez-Cuesta et al., 1981):

$$\mathrm{YI} = \frac{(142.86 \times b*)}{L*} \tag{5}$$

$$\mathrm{DI} = \frac{(1000 \times a)}{(L \times b)} \tag{6}$$

2.4.2. Respiration Rate, Electrolyte Leakage and Total Chlorophyll Content

The respiration rate of rocket leaves was determined at 10 °C initially and at each sampling time using a closed system as reported by (Kader, 2002). In particular, about 50 g of product for each replicate were put into 3.6 L sealed plastic jar (one jar for each replicate) where CO_2 was allowed to accumulate up to 0.1 % as the concentration of the CO_2 standard. The time taken to reach this value was detected by taking CO_2 measurements at regular time intervals. The CO_2 analysis was conducted by taking 1 mL of gas sample from the head space of the plastic jars through a rubber septum and injecting it into a gas chromatograph (p200 micro GC-Agilent, Santa Clara, CA, USA) equipped with dual columns and a thermal conductivity detector. Carbon dioxide (CO_2) was analyzed with a retention time of 16 s and a total run time of 120 s on a 10-m porous polymer (PPU) column (Agilent, Santa Clara, CA, USA) at a constant temperature of 70 °C. The respiration rate was expressed as µmol CO_2 kg^{-1} s^{-1}.

To determine electrolyte leakage, the method reported by Kim et al. (2005) was used with slight modifications. About 2.5 g of disks obtained using a cork borer (ø 8 mm) were placed in plastic tubes and immersed in 25 mL of distilled water. After 30 min of storage at 10 °C, the conductivity of the solution was measured using a conductivity meter (Cond. 51+-XS Instruments, Carpi, Italy). Then, the tubes with samples and solution were frozen at -20 °C and, after 48 h, the conductivity was detected after thawing and considered as total conductivity. Electrolyte leakage was calculated as the percentage ratio of initial over total conductivity.

The total chlorophyll content was detected according to the spectrophotometric method reported by Cefola and Pace (2015). In detail, 5 g of chopped rocket leaves was extracted in acetone/water (80:20 v/v) with a homogenizer (T-25 digital ULTRA-TURRAX®-IKA, Staufen, Germany) and then centrifuged at 15,000 rpm for 5 min. To remove all pigments, the extraction was repeated 5 times and extracts were combined. The absorbance was read immediately after the extraction procedure on extracts proper diluted using a spectrophotometer (UV-1800, Shimadzu, Kyoto, Japan) at three wavelengths, at 663.2 nm, 646.8 nm, and 470 nm. The total chlorophyll content was expressed as mg per 100 g of fresh weight using the equation reported by Wellburn (1994).

2.5. Statistical Analysis

The relationship among QLs and the postharvest quality parameters (color, respiration rate, electrolyte leakage and total chlorophyll content of rocket leaves) was tested by performing a one-way ANOVA. Then, a multifactor ANOVA was performed with the aim to evaluate the effects of fertilization levels (FL_1 or FL_2) and irrigation management approach (Timer or Sensor) on WUE, FUE, visual quality, color parameters, respiration rate, electrolyte leakage and total chlorophyll content.

The mean values were separated using the Student-Newman-Keuls (SNK) test and Statgraphics Centurion (version 18.1.12, Warrenton, Virginia, USA) was used for statistical analyses.

3. RESULTS AND DISCUSSION

3.1. Effects of Agronomic Treatments on Water and Fertilizers Use Efficiency and Postharvest Quality Parameters

Treatments resulted in a substantial differentiation of the sustainability of the production process expressed in terms of resources (water and fertilizers) use efficiency. Mean values of WUE were 21.4 and 34.4 g L^{-1}, on average, in Timer-based and Sensor-based irrigation treatments respectively, with no effects of the fertilization level. On the other hand, treatments showed a significant interaction on FUE, as reported in Figure 3.

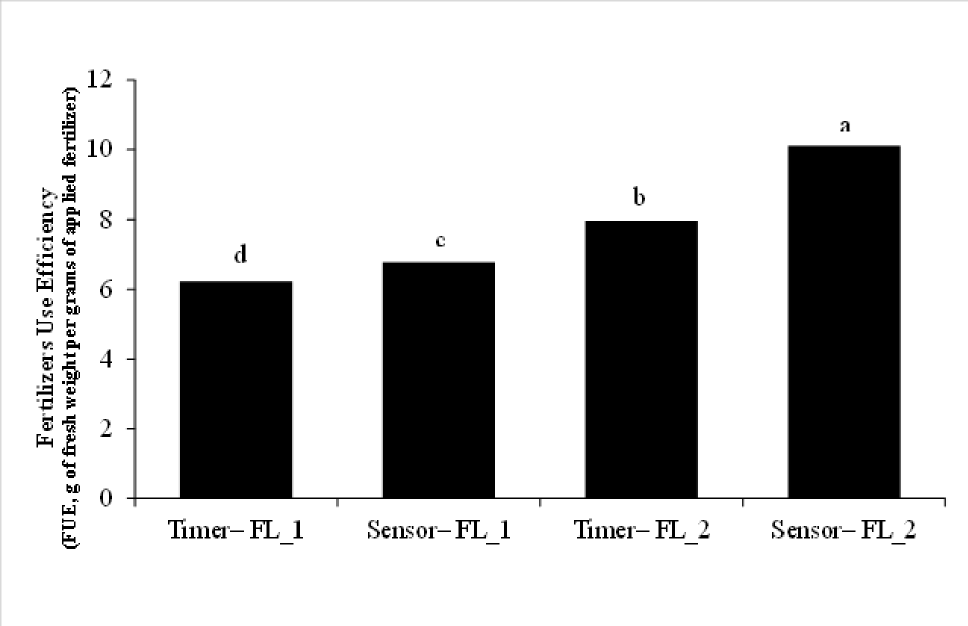


Figure 3. Fertilizer use efficiency (FUE), of greenhouse soilless rocket (*Diplotaxis tenuifolia* L.) subjected to different irrigation strategies (timer-based or sensor-based) and two fertilization levels (FL_1, high fertilization level; FL_2 low fertilization level). "Irrigation management strategy × Fertilization level" interaction significant at $p < 0.001$. Different letters above the columns indicate significant difference between the treatments ($p < 0.05$, means separation performed with SNK test).

Greenhouse soilless production can boost intensive cropping systems with impressive efficiency on water and nutrients use, and very high product yield and quality (Massa et al., 2020). Both sensor-based irrigation management (Montesano et al., 2015, 2016, 2018) and the rational application of fertilizers (Montesano et al., 2010; Santamaria et al., 2002) have been identified as promising approaches in combining high product quality with sustainable use of resources in greenhouse soilless vegetables production.

In relation to the effects of fertilization levels (FL_1 or FL_2) using the Timer or Sensor irrigation management approach, results obtained from the multifactor ANOVA showed that all factors (irrigation management strategies, fertilization levels and their interaction) did not influence the visual quality, the respiration rate, the color parameters and the total chlorophyll content of rocket leaves While, the electrolyte leakage was affected only by the irrigation management strategies (Table 2). In detail, fresh-cut rocket irrigated with the Sensor approach showed a mean value slightly higher (23.3 ± 5.2 %, on average) than that reported in samples irrigated with the Timer strategy (19.7 ± 6.2 %, on average), probably as a result of the lower water availability (Kirnak et al., 2003).

Table 2. Effects of irrigation management strategies (Timer or Sensor), fertilization levels (FL_1 or FL_2) and their interaction on visual quality, physical and chemical parameters of rocket leaves stored at 10 °C.

Parameters	VQ	*Physical Parameters*					*Chemical Parameters*
		Respiration Rate	ΔE^*	Yellowness Index	Degreening Index	Electrolyte Leakage	Total Chlorophyll Content
	(5-1)	*(μmol CO_2 kg^{-1} s^{-1})*				*%*	*(mg 100 g^{-1})*
Irrigation management strategies (A)	ns	ns	ns	ns	ns	****	ns
Timer	3.25	30.43	7.19	80.19	−20.69	19.66 b	46.96
Sensor	3.27	29.68	5.78	78.02	−21.24	23.35 a	48.52
Fertilization levels (B)	ns	ns	ns	ns	ns	ns	ns
FL_1	3.25	29.81	6.90	79.61	−20.69	22.16	47.80
FL_2	3.26	30.29	6.07	78.60	−21.24	20.85	47.67
A × B	ns	ns	ns	ns	ns	ns	ns

ns: not significant; **** significant for $p \leq 0.0001$.

A 5 to 1 rating scale was used for visual quality, where 5 = very good (very fresh, no signs of yellowing, bright, dark and uniform green, no defects), 4 = good (fresh, slight signs of yellowing, light green, slight loss of texture), 3 = fair, limit of marketability (slight wilting, moderate signs of yellowing, slight discoloration, minor defects, loss of texture), 2 = poor, limit of edibility (wilting, evident yellowing, discoloration, severe loss of texture), 1 = very poor (unacceptable quality due to decay, severe wilting and yellowing, complete loss of texture and other evident defects). The results are provided as the mean values of 6 samples for irrigation management strategies and fertilization levels (3 replicates × 2 irrigation management strategies or 2 fertilization levels). The mean values followed by different letters (a, b) are significantly different ($p \leq 0.05$), according to Student-Newman-Keuls test.

3.2. Relationship among Rocket Visual Quality Levels and Postharvest Quality Parameters

The color parameters (YI and DI), obtained by the colorimeter, were able to discriminate four QL: leaves very good (QL5) and good (QL4) from fair (QL3), poor (QL2) and very poor (QL1) (Table 3). As for YI, that indicates the degree of yellowness, rocket leaves on QL1 showed values 31 % higher (YI = 94.1 ± 6.5) than samples on QL5 (YI = 71.9 ± 9.7) and the same statistical differences between levels were observed for DI. In QL1 samples DI parameter resulted about 40% higher (DI = −14.7 ± 2.4) than rocket leaves belonging to QL5 (DI = −24.2 ± 2.0), indicating a gradual decrease of green color from QL5 to QL1. In the case of *ΔE**, three class were separated, QL 5-4-3 (mean value 1.6 ± 1.4) from QL2 (11.3 ± 5.1) and QL1 (20.1 ± 8.1). Similar results were reported by Pace te al. (2014), in which *ΔE** discriminated the 80 % of the QLs in fresh-cut lettuce, separating the QL5 from QL4-3, QL2 and QL1.

In the present study, the respiration rate of rocket leaves at harvest (QL5) was 35.52 ± 5.7 µmol CO_2 kg^{-1} s^{-1} and it remained rather low in all QLs, showing a slight decrease from QL5 to QL2 and increasing in QL1 rocket leaves, discriminating only the QL5 from the other levels (Table 3). The authors in Martínez-Sánchez et al. (2006) reported values of respiration rate in *Diplotaxis tenuifolia* L. stored at 4 °C of about 8.5 µmol CO_2 kg^{-1} s^{-1}, while in Kenigsbuch et al. (2014) the respiration rate in wild rocket stored at 17 °C was about 32.8 µmol CO_2 kg^{-1} s^{-1}. According to Luca et al. (2017), the differences on this parameter in rocket are related to the storage temperature and to the maturity of the leaves at harvest: they reported that young wild rocket leaves at 10 °C had higher (79.8 µmol CO_2 kg^{-1} s^{-1}) respiration rate than the old ones (47.7 µmol CO_2 kg^{-1} s^{-1}).

In the present research, electrolyte leakage was able to discriminate the marketable samples (QL5 and QL4) from the QL3 and the waste (QL2 or QL1) ones. Furthermore, these two classes of waste were well discriminated by electrolyte leakage (Table 3).

The same discrimination was observed in the case of total chlorophyll content, that showed a decrease of about 37 % from the QL5 to QL1 (Table 3). In particular, this parameter well separated the marketable samples (QL5 and QL4) from the QL3 and QL2; moreover, the waste samples (QL1) were well discriminated from the edible ones. The authors in Koukounaras et al. (2006) reported that the chlorophyll degradation, which causes yellowing leaves, is related to the quality loss of the product. Indeed, the total chlorophyll content is considered a good objective parameter for the QL assessment.

Table 3. Respiration rate, color parameters, electrolyte leakage and total chlorophyll content in rocket leaves stored at 10 °C, at each quality level (QL).

Parameters	***QL***					
	5	**4**	**3**	**2**	**1**	***p*-Value**
Respiration rate (µmol CO_2 kg^{-1} s^{-1})	35.5 a	30.2 bc	27.0 bc	25.9 c	31.5 b	****
*ΔE**	0 d	2.0 cd	2.8 c	11.3 b	20.1 a	****
Yellowness Index	71.9 d	70.0 d	77.4 c	86.6 b	94.1 a	****
Degreening Index	−24.2 d	−24.0 d	−22.2 c	−17.7 b	−14.7 a	****
Electrolyte leakage (%)	19.3 c	18.4 c	22.2 b	22.5 b	25.9 a	****
Total chlorophyll content (mg 100 g^{-1})	54.1 a	53.1 a	48.4 b	45.9 b	34.0 c	****

For each parameter the mean values followed by different letters (a, b, c, d) are significantly different (p-value < 0.05) according to Student-Newman-Keuls (SNK) test. Significance: **** = significant at p-value ≤ 0.0001.

3.3. Application of Self-Configuring CVS to Objectively Attribute the Visual Quality Level of Rocket Leaves and to Discriminate Them according to Preharvest Practices

The goal of the CVS was to reproduce the QL sensory evaluation of rocket leaves and to identify agronomic treatments in an objective, non-destructive and contactless way by simply imaging the product in proper conditions. Since color is the key information aspect used by the CVS, it was necessary to make its measurement as consistent as possible.

The color-chart, introduced in the scene, provided a reference that was used to measure, and then minimize the effects of any uncontrolled change in the acquisition environment. This was carried out by correlating the 24 expected color values provided by the manufacturer with the values measured in each image. This correlation was used to determine the parameters of the model that was used for correcting all the colors of the image. The three previously presented color correction models were applied and quantitatively compared to point out the best model for such kind of application. Two metrics to measure the effectiveness of color corrections models were considered. The first one evaluated their ability to reduce the distance between expected and measured color on the 24 patches of the color chart. The second one measured their effects on final classification accuracy, keeping unchanged all the following processing on images. The former method evaluated the correction on the same data used to estimate the parameters of the model: this made the response weaker and less reliable.

This paper proposes the latter method to achieve a better evaluation using the accuracy of the classification process applied to the images corrected using the three different color correction models. In this way, the data (the colors of the product) on which the models are compared are different from the ones used during the model construction. Moreover, the effectiveness of color correction is evaluated on the task of interest. The experiments pointed out that the two metrics do not provide always the same answer. As shown in Figure 4, the global distance between the expected and the measured values on the color chart patches was still large for WB, much smaller for LC and minimum for the PC. This was a natural result of the higher degrees of freedom of the PC that made easier to correct the 24 colors of the color-chart. When applied to the evaluation of QL of rocket leaves and to the identification of the Timer and Sensor approaches, the differences between LC and PC were small and could be considered negligible.

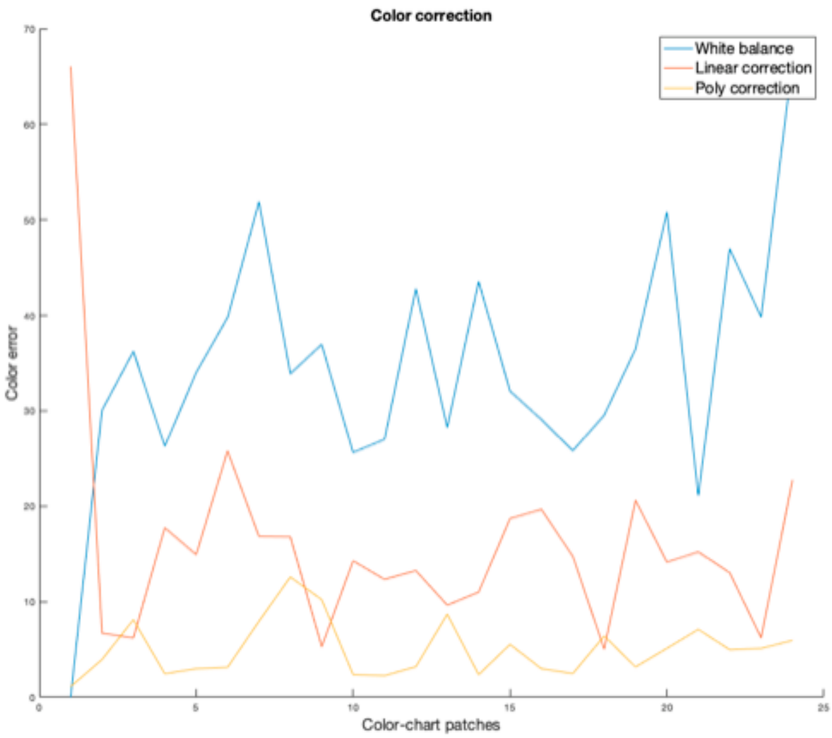


Figure 4. Average difference (Euclidean distance) between the expected colors and the measured values over the dataset of images corrected using White Balance (blue), Linear Correction (orange), Polynomial Correction (yellow). In abscissa there are the different color patches of the color chart.

In the experiments, all the images were corrected using all the three exposed models. Three different image datasets were generated, each associated to a different color correction model. The same subsequent processing and classification process, which was based on a Random Forest approach, was applied to the histograms associated with the three datasets. The corresponding performances were then measured. The classification was accomplished using two different resolutions of practical relevance. In the first case, the product completely marketable (QL5-QL4-QL3) was separated from the product just below the marketable limit (QL2) and from the not edible items. This may be useful because the leaves belonging to QL2 might be reusable to reduce waste. In the second case, the marketable product (QL5-QL4-QL3) was separated from unmarketable leaves (QL2-QL1): this might be a valid solution for commercial applications where it is important only to recognize the unmarketable product to remove it from the shelves.

Moreover, the classification was tried on the task of recognizing leaves from the Timer vs Sensor and those from FL_1 vs FL_2 treatments. The Table 4 shows the Confusion Matrices obtained by the Random Forest applied to the problem of distinguishing the marketable product

(QLs 5-4-3) from the edible one (QL 2) and from the waste (QL 1). The accuracy obtained by applying WB was 93 %, by applying LC was 96 %, by applying the PC was 95%. In this case the PC behaved slightly worse than LC even if on the color-chart the results were opposite.

Table 4. Confusion Matrices obtained by the classification working on the datasets coming from the three different color correction models and with the task of separating marketable product (QL 5-4-3) from the edible (QL 2) and from the waste (QL 1) ones.

	White Balance			**Linear Correction**			**Polynomial Correction**		
	Predicted QL			***Predicted QL***			***Predicted QL***		
Real QL	**1**	**2**	**3-4-5**	**1**	**2**	**3-4-5**	**1**	**2**	**3-4-5**
1	24	0	0	23	1	0	22	2	0
2	0	58	14	0	62	10	3	60	9
3-4-5	0	8	208	0	2	214	0	1	215

The Table 5 shows the Confusion Matrices obtained by applying the classification to the task of separating marketable product (QL5-4-3) from non-marketable product (QL 2-1). The accuracy obtained by applying WB was 93 %, by applying LC was 96 %, by applying the PC was 97 %.

Table 5. Confusion Matrices obtained by the classification working on the datasets coming from the three different color correction models and with the task of separating marketable product (QL 5-4-3) from non-marketable product (QL 2-1).

	White Balance		**Linear Correction**		**Polynomial Correction**	
	Predicted QL		***Predicted QL***		***Predicted QL***	
Real QL	**1-2**	**3–4-5**	**1-2**	**3-4-5**	**1-2**	**3-4-5**
1-2	80	16	87	9	87	9
3-4-5	7	209	2	214	1	215

There was a light improvement in the accuracy of LC and PC while the difference between them was still negligible. In this case, PC slightly outperformed LC in accord with the results on the color-chart but with a much smaller difference. The experiments pointed out that the two models produce mostly the same effects on the task. Therefore, is natural to use the LC model, which exhibits a lower computational complexity. The proposed self-configuring CVS used for the QL classification of rocket leaves allowed an objective system to be obtained, that can reproduce the human sensory evaluation, that consider a set of descriptors (such as color, defects and texture) as reference (Figure 1). Therefore, the proposed system, based on the extraction of color features, classified rocket leaves miming the end users involved in the visual QL assessment (Bhargava and Bansal, 2018).

Tables 6 and 7 show the Confusion Matrices associated to the tasks of recognizing the fertilization levels (FL_1 vs FL_2) using the irrigation management approach Timer and Sensor, respectively. The accuracies were quite low; approximately 70 % using Timer and 66 % using Sensor. LC and PC behaved similarly on irrigation management approach while LC outperformed PC on FL.

Table 6. The table shows the Confusion Matrices obtained by the classification working on distinguishing the different fertilization levels (FL_1 vs FL_2) using the timer-based irrigation management from datasets provided by the three different color correction methods.

	White Balance		**Linear Correction**		**Polynomial Correction**	
	Predicted FL		***Predicted FL***		***Predicted FL***	
Real FL	**FL_1**	**FL_2**	**FL_1**	**FL_2**	**FL_1**	**FL_2**
FL_1	107	49	109	47	100	56
FL_2	54	102	46	110	44	112

Table 7. The table shows the Confusion Matrices obtained by the classification working on distinguishing the different fertilization levels (FL_1 vs FL_2) using the sensor-based irrigation management from datasets provided by the three different color correction methods.

	White Balance		**Linear Correction**		**Polynomial Correction**	
	Predicted FL		***Predicted FL***		***Predicted FL***	
Real FL	**FL_1**	**FL_2**	**FL_1**	**FL_2**	**FL_1**	**FL_2**
FL_1	102	54	105	51	106	50
FL_2	46	110	52	104	50	106

The results obtained by recognizing the differences in the fertilization treatments (sustainable or conventional) were weaker. Nonetheless, these results were in accord with the indications provided by the statistical analysis of the measures supplied by the colorimeter and by the destructive tests made in laboratory. The performance is due to the small differences between products obtained by different treatments. This substantially uniform product quality confirmed that reducing water and fertilizer supply to exactly match real plant needs, without excesses, provides adequate growing conditions.

4. CONCLUSIONS

The experiments proved that the adopted agronomic treatments significantly improved the sustainability of the production process. This is demonstrated by the high values of WUE and FUE, obtained using sensors and reducing fertilizer inputs, while guaranteeing high product quality in all treatment conditions.

The proposed CVS was based on calibrated color images: Linear color correction proved to represent the best trade-off between efficacy and efficiency in making consistent color measurements. The proposed new form of integration of the Random Forest model in the color analysis was able to define and select color features suitable for classification without any human intervention. This new CVS achieved a high accuracy (about 95 %) in evaluating the rocket quality levels during storage. The same system was used to recognize traits related to the sustainability of the cultivation management with specific reference to water and nutrients use. In this second task, performance was lower and not relevant for practical application. However, it was fully in accord with the results provided by the standard methods currently used (colorimeter and destructive analytical tests in laboratory). Therefore, the different cultivation approaches did not significantly affect the characteristics of the product. For this last task, further investigations are needed.

The proposed computer vision system is cheap, fast and can be easily moved to an industrial production line. Given that the system is non-destructive and contactless, it enables an extended monitoring of products along the whole supply chain, thereby providing the opportunity for timely detection quality change and a reduction in economic losses and production waste.

REFERENCES

Amodio, M.L., Ceglie, F., Chaudhry, M.M.A., Piazzolla, F., Colelli, G., 2017. Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. *Postharvest Biology and Technology,* 125, 112–121.

Arivu, C.V.G., Prakash, G., Sarma, A.S.S., 2012. Online image capturing and processing using vision box hardware: Apple grading. *International Journal of Engineering Research*, 2, 639–643.

Bhargava, A., Bansal, A., 2018. Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University Computer and Information Sciences,* in press.

Breiman, L., 1996. Bagging predictors. *Machine Learning,* 24, 123–140.

Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5–32.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2019. Non-destructive and contactless quality evaluation of table grapes by a computer vision system. *Computers and Electronics in Agriculture,* 156, 558–564.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2017. Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. *Computers and Electronics in Agriculture,* 140, 303–310.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2018. Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material. *Journal of Food Engineering,* 223, 46–52.

Cefola, M., Pace, B., 2015. Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. *Journal of Food Processing and Preservation*, 39, 2523–2532.

Chaudhry, M.M., Amodio, M.L., Babellahi, F., de Chiara, M.L., Rubio, J.M.A., Colelli, G., 2018. Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves. *Journal of Food Engineering,* 238, 122–133.

Chaudhry, M.M.A., Amodio, M.L., Amigo Rubio, J.M., de Chiara, M.L., Babellahi, F., Colelli, G., 2020. Feasibility study for the surface prediction and mapping of phytonutrients in minimally processed rocket leaves (Diplotaxis tenuifolia) during storage by hyperspectral imaging. *Computers and Electronics in Agriculture*, 175, 105575.

De Pascale, S., Dalla Costa, L., Vallone, S., Barbieri, G., Maggio, A., 2011. Increasing water use efficiency in vegetable crop production: From plant to irrigation systems efficiency. *HortTechnolog,* 21, 301–308.

Jiménez-Cuesta, M., Cuquerella, J., Martinez-Javaga, J.M., 1981. Determination of a color index for citrus fruit degreening. In: Matsumoto, K., (Ed.). *Proceedings of the International Society of Citriculture/International Citrus Congress*, Tokyo, Japan, 9–12 November 1981.

Kader, A.A., 2002. Methods of gas mixing, sampling and analysis. In: Kader, A.A., (Ed.). *Postharvest Technology of Horticultural Crops (*pp. 145–148). University of California Agriculture and Natural Resources: Oakland, CA, USA.

Kenigsbuch, D., Ovadia, A., Shahar-Ivanova, Y., Chalupowicz, D., Maurer, D., 2014. “Rock-Ad”, a new wild rocket (Diplotaxis tenuifolia) mutant with late flowering and delayed postharvest senescence. *Scientia Horticulturae,* 174, 17–23.

Kim, J.G., Luo, Y., Tao, Y., Saftner, R.A., Gross, K.C., 2005. Effect of initial oxygen concentration and film oxygen transmission rate on the quality of fresh-cut romaine lettuce. *Journal of the Science of Food and Agriculture,* 85, 1622–1630.

Kirnak, H., Kaya, C., Higgs, D., Tas, I., 2003. Responses of drip irrigated bell pepper to water stress and different nitrogen levels with or without mulch cover. *Journal of Plant Nutrion,* 26, 263–277.

Koukounaras, A., Siomos, A.S., Sfakiotakis, E., 2006. 1-Methylcyclopropene prevents ethylene induced yellowing of rocket leaves. *Postharvest Biology and Technology,* 41, 109–111.

Løkke, M.M., Seefeldt, H.F., Skov, T., Edelenbos, M., 2013. Color and textural quality of packaged wild rocket measured by multispectral imaging. *Postharvest Biology and Technoogy,* 75, 86–95.

Luca, A., Kjær, A., Edelenbos, M., 2017. Volatile organic compounds as markers of quality changes during the storage of wild rocket. *Food Chemistry,* 232, 579–586.

Martínez-Sánchez, A., Marín, A., Llorach, R., Ferreres, F., Gil, M.I., 2006. Controlled atmosphere preserves quality and phytonutrients in wild rocket (Diplotaxis tenuifolia). *Postharvest Biology and Technology,* 40, 26–33.

Martínez-Sánchez, A., Tudela, J.A., Luna, C., Allende, A., Gil, M.I., 2011. Low oxygen levels and light exposure affect quality of fresh-cut Romaine lettuce. *Postharvest Biology and Technology,* 59, 34–42.

Massa, D., Magán, J.J., Montesano, F.F., Tzortzakis, N., 2020. Minimizing water and nutrient losses from soilless cropping in southern Europe. *Agricultural Water Management*, 241, 106395.

Montesano, F., Parente, A., Santamaria, P., 2010. Closed cycle subirrigation with low concentration nutrient solution can be used for soilless tomato production in saline conditions. *Scientia Horticulturae,* 124, 338–344.

Montesano, F.F., Serio, F., Mininni, C., Signore, A., Parente, A., Santamaria, P., 2015. Tensiometer-based irrigation management of subirrigated soilless tomato: Effects of substrate matric potential control on crop performance. *Frontiers in Plant Science,* 6, 1150.

Montesano, F.F., Van Iersel, M.W., Boari, F., Cantore, V., D'Amato, G., Parente, A., 2018. Sensor-based irrigation management of soilless basil using a new smart irrigation system: Effects of set-point on plant physiological responses and crop performance. *Agricultural Water Management,* 203, 20–29.

Montesano, F.F., Van Iersel, M.W., Parente, A., 2016. Timer versus moisture sensor-based irrigation control of soilless lettuce: Effects on yield, quality and water use efficiency. *HortScience,* 43, 67–75.

Narendra, V.G., Amithkumar, V.G., 2019. An intelligent computer vision system for vegetables and fruits quality inspection using soft computing techniques. *Agricultural Engineering International: CIGR Journal*, 21, 171–178.

Pace, B., Cavallo, D.P., Cefola, M., Attolico, G., 2017. Automatic identification of relevant colors in non-destructive quality evaluation of fresh salad vegetables. *International Journal of Food Processing Technology,* 4, 1–5.

Pace, B., Cavallo, D.P., Cefola, M., Colella, R., Attolico, G., 2015. Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio. *Innovative Food Science and Emerging Technologies*, 32, 200–207.

Pace, B., Cefola, M., Da Pelo, P., Renna, F., Attolico, G., 2014. Non-destructive evaluation of quality and ammonia content in whole and fresh-cut lettuce by computer vision system. *Food Research International,* 64, 647–655.

Pace, B., Cefola, M., Renna, F., Attolico, G., 2011. Relationship between visual appearance and browning as evaluated by image analysis and chemical traits in fresh-cut nectarines. *Postharvest Biology and Technology,* 61, 178–183.

Patel, K.K., Kar, A., Jha, S.N., Khan, M.A., 2012. Machine vision system: A tool for quality inspection of food and agricultural products. *Journal of Food Science and Technology,* 49, 123–141.

Pathare, P.B., Opara, U.L., Al-Said, F.A.J., 2013. Colour measurement and analysis in fresh and processed foods: A review. *Food and Bioprocess Technology,* 6, 36–60.

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 Technol*ogy, 85, 116–123.

Santamaria, P., Elia, A., Serio, F., 2002. Effect of solution nitrogen concentration on yield, leaf element content, and water and nitrogen use efficiency of three hydroponically-grown rocket salad genotypes. *Journal of Plant Nutrition*, 25, 245–258.

Verain, M.C., Sijtsema, S.J., Antonides, G., 2016. Consumer segmentation based on food-category attribute importance: The relation with healthiness and sustainability perceptions. *Food Quality and Preference*, 48, 99–106.

Wellburn, A.R., 1994. The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology,* 144, 307–313.

Chapter 4

NON-DESTRUCTIVE AND CONTACTLESS ESTIMATION OF CHLOROPHYLL AND AMMONIA CONTENTS IN PACKAGED FRESH-CUT ROCKET LEAVES BY A COMPUTER VISION SYSTEM

Michela Palumbo[1,2], Bernardo Pace[1], Maria Cefola[1], Francesco Fabiano Montesano[3], Giancarlo Colelli[2], Giovanni Attolico[4],

[1] Institute of Sciences of Food Production, CNR, c/o CS-DAT, Via Michele Protano, 71121 Foggia
[2] Department of Agriculture, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Via Napoli 25, 71122 Foggia.
[3] Institute of Sciences of Food Production, CNR, Via G. Amendola, 122/O, 70126 Bari.
[4] Institute on Intelligent Industrial Systems and Technologies for Advanced Manufacturing, CNR, Via G. Amendola, 122/O, 70126 Bari

ABSTRACT

Computer Vision Systems (CVS) offer a non-destructive and contactless tool to assign visual quality level to fruit and vegetables and to estimate some of their internal characteristics. The innovative CVS described in this paper exploits the combination of image processing techniques and machine learning models (Random Forests) to assess the visual quality and predict the internal traits on unpackaged and packaged rocket leaves. Its performance did not depend on the cultivation system (traditional soil or soilless). The same CVS, exploiting its machine learning components, was able to build effective models for either the classification problem (visual quality level assignment) and the regression problems (estimation of senescence indicators such as chlorophyll and ammonia contents) just by changing the training data. The experiments showed a negligible performance loss on packaged products (Pearson's linear correlation coefficient of 0.84 for chlorophyll and 0.91 for ammonia) with respect to unpackaged ones (0.86 for chlorophyll and 0.92 for ammonia). Thus, the non-destructive and contactless CVS represents a valid alternative to destructive, expensive and time-consuming analyses in the lab and can be effectively and extensively used along the whole supply chain, even on packaged products that cannot be analyzed using traditional tools.

Keywords: contactless quality level assessment, *Diplotaxis tenuifolia* L., image analysis, packaged vegetables, senescence indicators prediction.

1. INTRODUCTION

Rocket is a green leafy vegetable usually marketed and consumed as fresh-cut salad, alone or mixed to other leafy vegetables. It is well known and appreciated for its pleasant bitter taste and for its high content in bioactive compounds, such as vitamins, minerals and antioxidants. The two species commonly sold on the market are *Eruca sativa* and *Diplotaxis tenuifolia* or wild rocket that is known to have longer shelf life (Mastrandrea et al., 2016).

The quality loss during the postharvest storage is mainly due to senescence, strictly related to chlorophyll degradation that, therefore, is the most common index used to evaluate quality and freshness of this product (Limantara et al., 2015; Cavallo et al., 2017). Generally, as reported by Pace et al. (2019), a 30 % loss of total chlorophyll content is considered the shelf life limit in rocket leaves stored for about 16 days at temperature between 5 and 20 °C. Another important indicator of leaves senescence in fresh-cut rocket is ammonia accumulation in plant tissues. It is reported that ammonia is a product of protein catabolism, thus it is considered an indicator of freshness, when is detected in low amount in the vegetal tissues (Chandra et al., 2006; Cefola et al., 2010). Moreover, since chlorophyll degradation, responsible for rocket yellowness during storage, causes protein catabolism, it may contribute to ammonia accumulation and a relationship with discoloration or yellowing process may be expected (Amodio et al., 2018). High concentrations of this component may cause tissue damages with visible effects that impact the overall quality of the product (darkening and browning of detached leaves) (Chibnall, 1939; Mastrandrea et al., 2016; Amodio et al., 2018).

Traditional approaches for chlorophyll and ammonia content measurements in leafy vegetables include destructive methods, based on spectrophotometric assays. If these approaches have been considered the standard used methods for these determinations for a long time, they require specific laboratory equipment and destructive sampling and they are expensive and time consuming. While for the ammonia analysis the destructive method is widely applied, for chlorophyll evaluation, modern, handheld sensors have received a considerable attention in the last decades because of their high accuracy and real time measurement in a non-destructive way directly on field or on minimally processed products. So, many researchers developed various types of chlorophyll metres (Novichonok et al., 2016), e.g. multispectral and hyperspectral sensors (Chen et al., 2010; Li et al., 2014) that measures the spectral reflectance of leaves to assess the total chlorophyll content. Many of these techniques are costly and complex and require the presence of specialised personnel. Most widely used for chlorophyll content measurement are fluorimeters (Ferrante and Maggiore, 2007) and SPAD metre or similar devices (e.g. the atLEAF Chlorophyll metre, FT Green LLC, Wilmington, DE, USA (Ling et al., 2011; Liu et al., 2012; Yue et al., 2015; Yuan et al., 2016; Baresel et al., 2017). Although such instruments are simpler, faster and cheaper than chemical analysis, they need to touch the leaf to measure the chlorophyll content of a limited area of the leaf surface. Therefore, their use in industrial lines is limited, also because the estimation of the chlorophyll content depends on the quality of sampling (Cavallo et al., 2017).

Recently, image analysis based on common digital RGB cameras has proved to be a promising approach for the assessment of chlorophyll content of leafy vegetables in smart agriculture (Mohan and Gupta, 2019) and postharvest quality assessment (Pace et al., 2014, 2015; Cavallo et al., 2017).

Imaging systems proved to be more robust than area-based instruments as they work at pixel level considering the entire visible surface of the product. Meanwhile, many studies evaluated the use of digital images to analyse the total chlorophyll content of leaves of rice (Wang et al., 2013, 2014), soybean (Rigon et al., 2016), corn (Vesali et al., 2017), spinach (Agarwal and Gupta, 2018) and rocket leaves (Cavallo et al., 2017; Pace et al., 2019), both during production and postharvest. Currently, they also exploit the use of common smartphone cameras that are often equipped with high-speed processor (Mohan and Gupta, 2019). The success of Computer Vision Systems (CVS) is due to the possibility of establishing relationships between spectral reflectance indices and chlorophyll absorbance, and RGB (red, green and blue) components of an image (Santos do Amaral et al., 2019). Recently, Cavallo et al. (2018) indicated that to assess the quality of fresh-cut iceberg lettuce by CVS was possible also on packaged samples with minimal performance loss with respect to unpackaged samples. These authors recorded a performance loss of only about 3 % due to the presence of packaging (accuracy of 83 % on packed product instead of 86 % on unpacked one), showing the power of image analysis in monitoring the quality of fresh-cut vegetables. The Convolutional Neural Network (CNN) segmentation method was able to separate the graphical elements and the regions affected by lighting artefacts from the product inside the commercial bag. This enable a pervasive use of the system along the whole supply chain, regardless the presence of the packaging. Further investigations are needed to confirm and implement this emerging technology for a continuous check of the quality of fresh-cut products along the whole supply chain.

Few applications regarding the use of CVS for the detection of ammonia content in leafy vegetables, are reported. Pace et al. (2014), applied CVS on whole and fresh-cut lettuce for the non-destructive evaluation of this parameter often used as a senescence indicator in leafy vegetables (Tudela et al., 2013; Cefola and Pace, 2015). Starting from these considerations, the aims of the present investigation was to verify and assess the capability of the non-invasive contactless CVS in assessing the visual quality changes during postharvest storage of packaged fresh-cut rocket leaves and in estimating some of their internal characteristics (chlorophyll and ammonia contents): experiments have been made on samples coming from soil and soilless growing systems.

2. MATERIALS AND METHODS

2.1. *Plant material and experimental setup*

Rocket leaves (*Diplotaxis tenuifolia* L. 'Dallas') were cultivated at the same time in soil or soilless cultivation systems at the CNR-ISPA experimental farm La Noria (CNR-ISPA) located in Mola di Bari (in the South of Italy). Harvests were performed at 55, 70 and 110 d and at 60, 110 and 145 d after sowing for soilless and soil system, respectively. At each harvest time, fresh-cut rocket leaves, separated per cultivation system were provided to the laboratory for image analysis by CVS and postharvest quality determinations. Then, fresh-cut leaves were selected to avoid samples with defects and mechanical damages and packed in open PP bags (dimensions 50 ×30 cm, Orved, Musile di Piave, Italy) of about 600 g each one. In detail, 13 bags (replicates) were prepared for samples cultivated on soil system, while 10 bags for rocket leaves cultivated on soilless system. Then, all samples were stored at 10 °C for 16 and 18 d for soilless and soil system, respectively.

2.2. *Sensory visual quality attribution during cold storage of rocket leaves*

During storage, samples were taken and observed to attribute the visual quality level (QL) according to the scale reported by Palumbo et al. (2021). In detail, at each sampling day, an amount of sample was taken from each PP bag and evaluated by a group of 5 researchers using the 5–1 rating scale cited above, where 5 = very good (very fresh, no signs of yellowing, bright, dark and uniform green, no defects), 4 = good (fresh, slight signs of yellowing, light green, slight loss of texture), 3 = fair (slight wilting, moderate signs of yellowing, slight discoloration, minor defects, loss of texture), 2 = poor (wilting, evident yellowing, discoloration, severe loss of texture), 1 = very poor (unacceptable quality due to decay, severe wilting and yellowing, complete loss of texture and other evident defects). A score of 3 was considered to be the limit of marketability, while a score of 2 represented the limit of edibility. Then, images of packaged and unpackaged fresh-cut rocket leaves were acquired by CVS and the quality of the same samples was evaluated through destructive conventional methods as detailed below.

2.3. ***Colour analysis by colorimeter, total chlorophyll content, ammonia content, and electrolyte leakage***

The CIELAB colour parameters ($L*$, $a*$ and $b*$) were detected, for each replicate, on 3 random points on the surface of 10 rocket leaves using a colorimeter (CR400, Konica Minolta, Osaka, Japan). The instrument was calibrated with a standard reference having values of $L*$, $a*$ and $b*$ corresponding to 97.44, 0.10 and 2.04, respectively. Then, the colour parameter of Hue angle (h°) was calculated from primary $L*$, $a*$ and $b*$ readings according the equation below:

$$h° = \tan^{-1}\frac{b*}{a*} \qquad (1)$$

The total chlorophyll content was measured according to the spectrophotometric method reported by Cefola and Pace (2015). Five grams of rocket leaves were chopped and extracted in acetone/water (80:20 v/v) with a homogenizer (T-25 digital ULTRA-TURRAX® - IKA, Staufen, Germany) and then centrifuged at 6440 g for 5 min (C2500-R Prism R, Labnet, Edison, US). To remove all pigments, the extraction was repeated 5 times (10 mL per times) and extracts were combined. The absorbance was read immediately after the extraction procedure on extracts proper diluted using a spectrophotometer (UV-1800, Shimadzu, Kyoto, Japan) at three wavelengths, at 663.2 nm, 646.8 nm, and 470 nm. Total chlorophyll content was expressed as mg g^{-1} of fresh mass using the equation reported by Wellburn (1994).

Ammonia content was evaluated according to Fadda et al. (2016). Chopped rocket leaves (5 g) were homogenized for 2 min in 20 mL of distilled water on an ice bath, and then centrifuged for 5 min at 6440 g at 4 °C. Then, the supernatant (0.5 mL) was mixed with 5 mL of nitroprusside reagent (phenol and hypochlorite in alkali reaction mixture) and heated at 37 °C for 20 min. The colour development after incubation, was determined with the spectrophotometer (reading the absorbance at 635 nm). The content of NH_4^+ was expressed as µg NH_4^+ g^{-1} of fresh mass, using ammonium sulfate as standard (0–10 µg mL^{-1}, $R^2 = 0.99$).

The electrolyte leakage was determined following the method reported by Palumbo et al. (2021). In detail, about 2.5 g of rocket leaves disks obtained using a cork borer (ø 8 mm) were placed in plastic tubes and immersed in 25 mL of distilled water. After 30 min of storage at 10 °C, the conductivity of the solution was measured using a conductivity meter (Cond. 51+ - XS Instruments, Carpi, Italy). Then, the tubes with samples and solution were frozen at – 20 °C and, after 48 h, the conductivity was detected after thawing and considered as total conductivity. Electrolyte leakage was calculated as the percentage ratio of initial over total conductivity.

2.4. *Image analysis by computer vision system*

At each sampling day, a sample of about 60 g of product was taken from each replicate and was placed inside a 20 × 25 cm polypropylene (PP) bags (Carton Pack S.p.A., Rutigliano, Italy). Three images of the packaged product were acquired by randomly shuffling the rocket in the bag before each acquisition: this procedure generated three different images from each packaged sample to increase the amount of observed surface and the variability of visual appearance considered by the CVS. Then, the same product was extracted from the bag and three images were acquired by randomly shuffling and stacking the leaves before each acquisition. In this way, for each replicate the CVS acquired six images: three of the packaged product and three of the unpackaged product. Therefore, 78 images were available at each time for samples coming from soil system (6 images for each of the 13 replicates) and 60 images for samples coming from the soilless system (6 images for each of the 10 replicates). The final dataset was composed by 429 and 450 image from soil and soilless cultivation respectively after all three harvest dates (Table 1). Therefore, the final image dataset, including images from all the harvests, was composed by 207 images for the quality level 5 and by 168 images for each quality level from 4 to 1 of packaged and unpackaged rockets (Table 1).

Table 1. Composition of the dataset of images with respect to harvests (H1, H2 and H3), and quality levels (QL).

	Soil							***Soilless***						
	Replicates	**QL**					**Total**	**Replicates**	**QL**					**Total**
		5	**4**	**3**	**2**	**1**			**5**	**4**	**3**	**2**	**1**	
		15	15	15	15	15	**75**		15	15	15	15	15	**75**
H 1	**13**	15	15	15	15	15	**75**	**10**	15	15	15	15	15	**75**
		9	9	9	9	9	**45**		-	-	-	-	-	**-**
		15	15	15	15	15	**75**		15	15	15	15	15	**75**
H 2	**13**	15	15	15	15	15	**75**	**10**	15	15	15	15	15	**75**
		9	9	9	9	9	**45**		-	-	-	-	-	**-**
		15	-	-	-	-	**15**		15	15	15	15	15	**75**
H 3	**13**	15	-	-	-	-	**15**	**10**	15	15	15	15	15	**75**
		9	-	-	-	-	**9**		-	-	-	-	-	**-**
		117	**78**	**78**	**78**	**78**	***429***		**90**	**90**	**90**	**90**	**90**	***450***

H1, H2 and H3 were performed at 55, 70 and 110 d and at 60, 110 and 145 d after sowing for soilless and soil system, respectively.

2.5. *Image processing steps by Computer Vision System*

2.5.1. Acquisition of calibrated colour images

A flowchart of the processing steps (whose sequence is slightly different for packaged and unpackaged products) along with their effects on the images, is shown in Figure 1. All the software was developed using Matlab 2019a (Mathworks Inc., Massachussets, USA). To acquire calibrated colour images, colour changes due to environment conditions (lighting, geometry, sensor instability) were evaluated and reduced to the minimum. Images were acquired using the set-up reported earlier (Pace et al., 2015, 2017; Cavallo et al., 2017, 2018) using a 3CCD (with a dedicated Charged Coupled Device for each colour channel) digital camera M9GE (Jai Ltd., Yokohama, Japan) having a resolution of 1024 × 768 pixels. The imaged area is about 32 × 24 cm. A 3CCD sensor has been used to avoid the artefacts introduced by the demosaicing methods required to record colour information using a single CCD. The optical axis of the Linos MeVis (Linos Photonics Ltd, Edinburgh, UK) 12 mm lens system was perpendicular to the black background. Two DC power suppliers delivered current to eight halogen lamps, placed along two rows at the two sides of the imaged area and oriented at a 45° angle with respect to the optical axis. All the images were saved using the uncompressed TIFF format to avoid the artifacts introduced by compression algorithms.

2.5.2. Colour-chart processing and foreground segmentation

A small X-Rite colour-chart (X-rite Italy srl, Prato, Italy) with 24 patches of known colours was placed into the scene to measure colour variations due to environmental conditions and sensor characteristics by comparing the expected numerical values released by the manufacturer with the ones acquired by the camera. The colour-chart was automatically detected regardless of its position and orientation (Cavallo et al., 2017). All the colours in the colour-chart were used to estimate the linear transformation used for colour correction. Image processing worked only on the part of each image belonging to the product at hand (foreground), while the background was discarded. The CVS automatically separated foreground and background without any human intervention: two thresholds were derived from the analysis of the whole image in the HSV colour space as described in Cavallo et al. (2017). This segmentation identified the region belonging to the product as a whole and did not separate its different parts neither discarded any region of leaves. It was designed to be conservative, that is to discard all the background pixels even at the cost of removing some marginal borders of the product. It removed also background area inside the stack of leaves as long as part of leaves too dark (for example for self-shadowing of the product) to provide meaningful colour information.

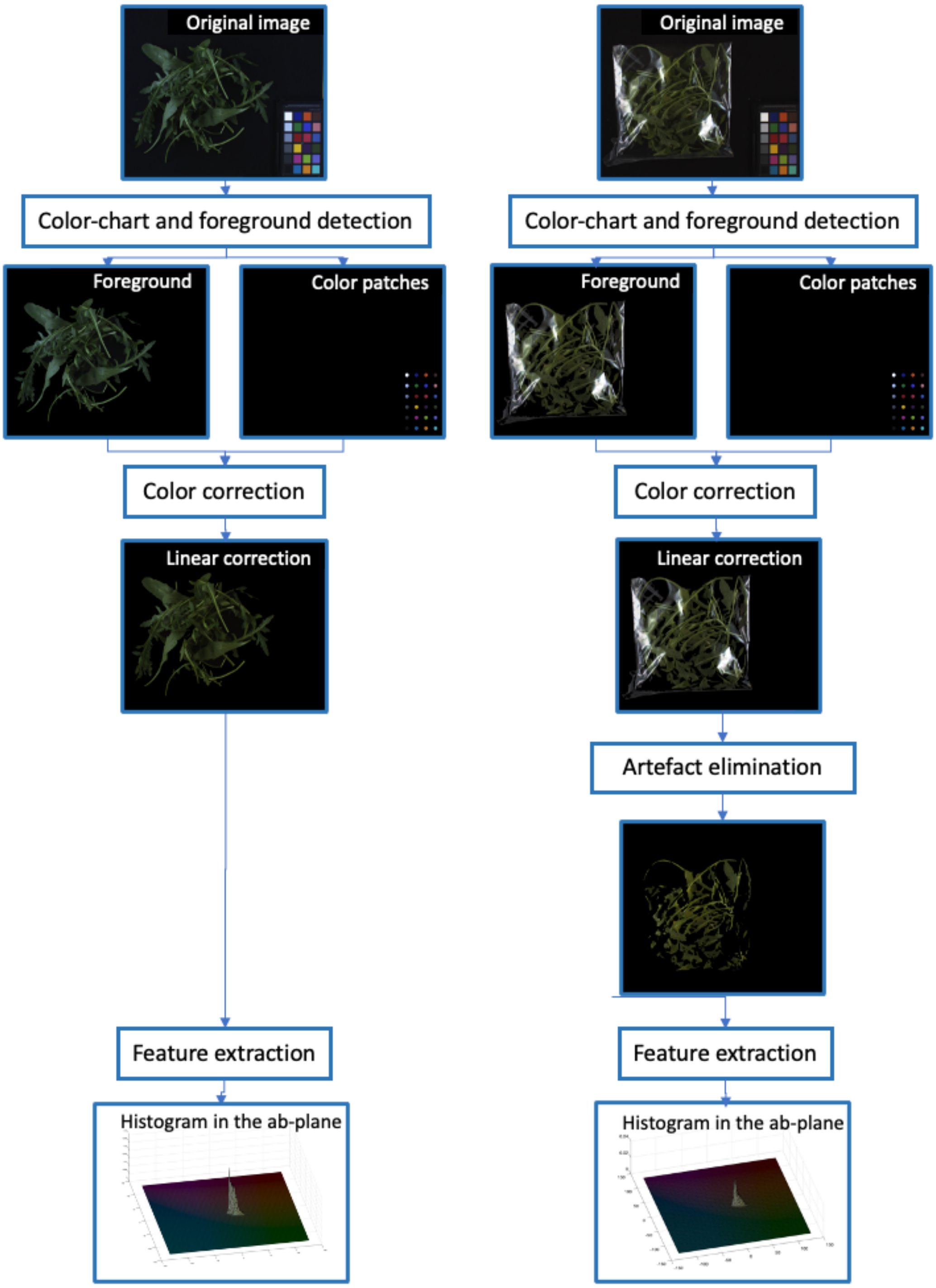


Figure 1. The flowcharts of the processing done on the images of unpackaged products (on the left) and on packaged products (on the right). They differ for the artefact elimination step applied to packaged products that selects the pixels where the camera measures meaningful colours.

2.5.3. Colour correction

The linear correction model proved to be the best trade-off between effectiveness and computational complexity, and it was used in the experiments (Palumbo et al. 2021). Let it be $[r_e^i\ g_e^i\ b_e^i]^T$ and $[r_m^i\ g_m^i\ b_m^i]^T$ the expected and the measured RGB values respectively for the i-th patch i=1, …, 24. A linear correction (LC) model, a 3 x 3 matrix, was evaluated to reduce the distance between the expected and the measured values on the colour chart:

$$\begin{bmatrix} r_c \\ g_c \\ b_c \end{bmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{bmatrix} r_m \\ g_m \\ b_m \end{bmatrix}$$

where $\begin{bmatrix} r_c \\ g_c \\ b_c \end{bmatrix}$ are the colours corrected using the matrix whose elements were evaluated using a least-square approach and all the patches of the colour-chart. The same matrix was therefore used to correct all the foreground pixels of the image. The linear transformation was different for each image (it was evaluated from the colour-chart appearance in each specific image) to adapt to the specific conditions of each acquisition. The linear correction requires 73 ms.

2.5.4. Artefact elimination from packaging

The unpredictable orientation bag's surface with respect to light can generate artefacts such as reflections. In those regions, the camera cannot measure meaningful colours from the product at hand. Before feature extraction, those areas must be removed from each image. The flowchart in the Figure 1 shows the placement of the artefact elimination step inside the processing chain and its effects on the image. Each image was converted in the HSV (Hue-Saturation-Value) colour space. Artefacts that are colourless and much brighter than the product were removed using two data driven thresholds automatically extracted on the Hue ($thresh_h$) and on the Value ($thresh_v$) components of the image using the Otsu algorithm. Pixels with hue greater than $thresh_h$ and value lower than $thresh_v$ were considered as product and maintained in the following processing. This conservative choice significantly reduced the risk of keeping saturated pixels in the image while leaving a resulting area large enough to accomplish the tasks of interest.

2.5.5. Features extraction

The device independent and perceptually uniform CIE $L*a*b*$ colour space was chosen to accomplish colour analysis. The $L*$ component was discarded being too sensible to not uniform illumination levels across the scene. The complete histogram of the foreground pixels (expressing the number of occurrences of each colour in the a^* b^* plane) was used as feature set for both classification and regression. The continuous (a^*, b^*) plane was discretized using 30 bins for each axis. The histogram was reshaped as a vector and passed to the supervised learning module. The ensemble method used for classification and regression can sample for each tree a subset of this quite large set of features and figure out their best use, keeping acceptable the computational complexity.

2.5.6. Image analysis

To assign to a sample a value out of a finite set of quality levels is a classification problem. To estimate the values of chlorophyll and ammonia contents is a regression problem. In both cases, the same model was used, and its parameters were set applying a supervised learning approach to the available samples. Each sample was composed by the features (the histogram described in the previous paragraph), an integer value for the expected visual quality and two real values for chlorophyll and ammonia contents measured in the lab. The same Random Forest approach was used to accomplish both classification and regression (Breiman, 1996, 2001). In case of classification, the model was trained to assign the QL to the product. In case of regression, two different models were trained to estimate the chlorophyll content and the ammonia content of the product. All the models shared the same architecture whose free parameters were fixed for each specific task. The values of the cells of the histogram in the a^* b^* plane (of the CIE $L^*a^*b^*$ colour space) of each image provided the vocabulary of features used for training the models. The approach, for training each tree, randomly selected the training examples from the available training data and randomly selected which values of the histogram to use as features to build the tree at hand. Each tree of the forest was allowed to have a maximum of 10 branches. Due to the limited number of samples, a 10-fold cross validation approach was used: that means to divide the available data into 10 groups (folds), each having approximately the same number of elements. The partitions were made with stratification; therefore, each fold approximated the same distribution of the whole training set. According to the 10-folds validation strategy, the training was done 10 times. At each round, a different fold was separated for testing the results while the other nine folds were used for training. The average of the results obtained in the ten rounds estimated the performance of the method for a single iteration. To increase the stability of results, 10 iterations were run for each task (classification for visual quality, regression for chlorophyll estimation, regression for ammonia estimation): the best value and the average of their results were stored. The best value was very similar to the average, proving that the randomness of the choice of training samples and of features for each tree did not affect significantly the performance of the resulting model.

2.6. Statistical analysis

A one-way ANOVA was performed to study the relationship between the most important quality parameters (total chlorophyll content, ammonia content, hue angle and electrolyte leakage) and the QLs during the rocket leaves cold storage (10 °C) with the aim of identifying the physical and chemical parameters able to classify in an objective and consistent way the QLs of rocket leaves.

The mean values were separated using the Student-Newman-Keuls (SNK) test and Statgraphics Centurion (version 18.1.12, Warrenton, Virginia, USA) was used for statistical analyses.

Partial least squares regression (PLSR) was run using the software The Unscrambler X (CAMO AS, Oslo, Norway).

3. RESULTS AND DISCUSSION

3.1. Changes in quality parameters during storage of fresh-cut rocket leaves

During storage the change in the sensory QL was mainly due to the reduction of green colour of rocket leaves from the score 5 to 1 in both cultivation system, as showed in Figure 2. Considering the quality parameters determined during the rocket leaves storage, a significant separation of the visual QL was obtained by colour analysis (Table 2).

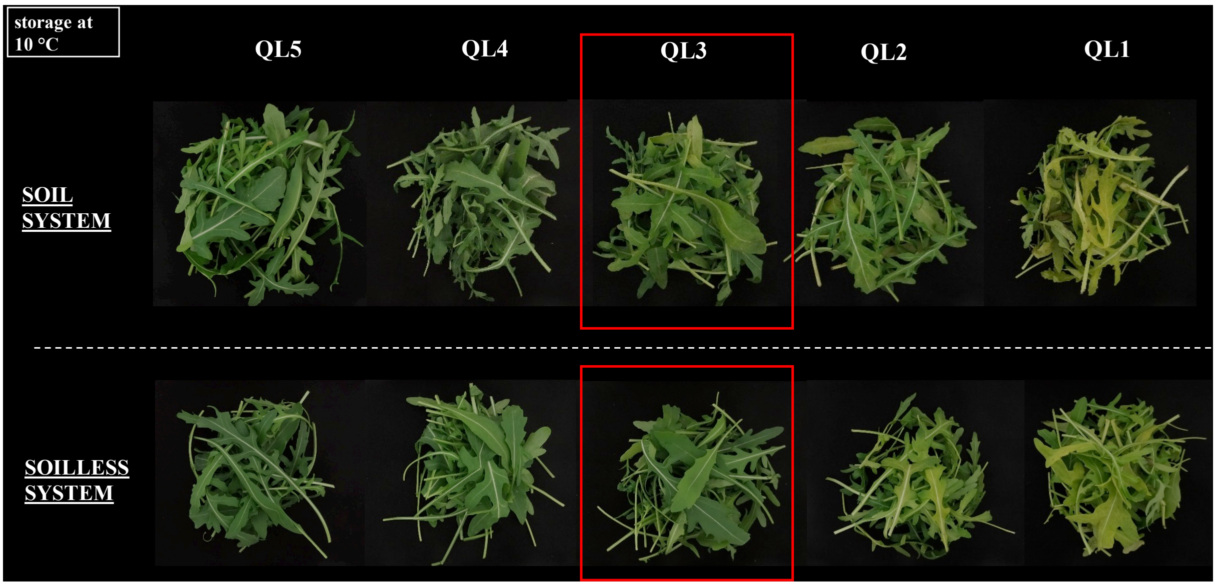


Figure 2. Changes in the sensory quality level (QL) of fresh-cut rocket leaves cultivated in soil or soilless system during the storage at 10 °C. QL5 = very good; QL4 = good; QL3 = fair; QL2 = poor; QL1 = very poor.

Table 2. Hue angle, total chlorophyll content, ammonia content and electrolyte leakage in fresh-cut rocket leaves cultivated on different systems (soil, soilless or all samples) and stored at 10 °C at each quality level.

Parameter	***Quality Level***										***P*-value**
	5 *very good*		**4** *good*		**3*** *fair*		**2** *poor*		**1** *very poor*		
Soil system											
Hue angle ($h°$)	125.6	a	124.7	a	121.3	b	118.1	c	114.4	d	****
Total chlorophyll content (mg g^{-1})	0.9	a	0.7	b	0.6	c	0.5	d	0.4	e	****
Ammonia content (µg NH_4^+ g^{-1})	7.3	c	3.1	c	6.8	c	89.7	b	184.3	a	****
Electrolyte leakage (%)	12.7	e	21.0	d	26.2	c	30.8	b	40.7	a	****
Soilless system											
Hue angle ($h°$)	126.8	a	125.5	b	123	c	119.6	d	116.8	e	****
Total chlorophyll content (mg g^{-1})	0.7	a	0.7	a	0.5	b	0.4	c	0.4	c	****
Ammonia content (µg NH_4^+ g^{-1})	4.3	c	2.9	c	9.9	c	38.5	b	80.1	a	****
Electrolyte leakage (%)	18.9	c	20.3	c	23.5	b	27.2	a	28.0	a	****
All samples (soil and soilless)											
Hue angle ($h°$)	126.1	a	125.2	b	122.2	c	118.8	d	115.6	e	****
Total chlorophyll content (mg g^{-1})	0.8	a	0.7	b	0.6	c	0.5	d	0.4	e	****
Ammonia content (µg NH_4^+ g^{-1})	3.0	c	6.0	c	8.4	c	67.8	b	133.7	a	****
Electrolyte leakage (%)	15.9	e	20.6	d	24.6	c	28.8	b	34.1	a	****

*: limit of marketability.

For each parameter the mean values followed by different letters are significantly different (P-value < 0.05) according to Student-Newman-Keuls (SNK) test.

Significance: **** = significant at P-value ≤ 0.0001.

In detail, for samples cultivated on soil system the $h°$ separated the leaves very good (QL5) and good (QL4) from fair (QL3), poor (QL2) and very poor (QL1) ones, showing a decrease of about 8.9 % from QL5 to QL1. In soilless cultivation system, all the QLs of rocket leaves were well separated by $h°$. Its value had a decrease of 7.8 % from QL5 (very good leaves) to QL1 (waste samples). The same QL separation was obtained by $h°$ value when all samples (soilless and soil) were considered (Table 2). Similar results were reported by Mastrandrea et al. (2016) in which $h°$ values decreased of 7.2 % in rocket leaves stored at 10 °C. Since, higher $h°$ represents a greener product (Pathare et al. 2013), in the present study the reduction in $h°$ values during the storage (from QL5 to QL1), means a gradual degreening of rocket leaves.

The electrolyte leakage was closely related to the quality and shelf life of fresh-cut produce and it is a physical parameter commonly used to measure the intensity of oxidative damages to cell membranes due to reactive oxygen species development in fresh-cut tissues (Kou et al. 2014). So, higher values in electrolytic leakage indicate higher physiological stress of leaves tissues. In this research, electrolyte leakage significantly increased in rocket leaves obtained by soil system well discriminating all the five QLs (Table 2). The same QL separation was observed considering all samples, with a significant increase of electrolyte leakage from the QL5 to QL1. On the contrary, this parameter in rocket leaves cultivated on soilless system proved to significantly discriminate QL5) and QL4 from QL3, recording an increase of 24.3 %. Moreover, in this cultivation system, the QL3 samples were well separated from QL2 and QL1 (Table 2). Similar results were reported by Palumbo et al. (2021) on rocket leaves cultivated under soilless cultivation system and stored at 10 °C, in which this parameter was able to discriminate QL5 and QL4 from the QL3 and the QL2 and QL1. In the present research work, the percentage increase of electrolyte leakage along the different QLs in rocket leaves grown on soil system was higher (220.5 %) than that detected on samples cultivated under soilless system (48.1 %) (Table 2), pointing out that the last cultivation system was probably more efficient in terms of reduction of induced oxidative stresses on cell membranes (Bonasia et al. 2017).

Moreover, a significant relationship was also found between decreasing visual QL and total chlorophyll and ammonia content, which can be considered objective markers of quality loss for rocket leaves (Table 2). The chlorophyll content allowed to have the same QL discrimination when the samples from soil and soilless system were joined (Table 2). For samples cultivated on soilless system, the chlorophyll content at harvest (QL5) was 24.6 % lower than that measured in samples cultivated on soil system, showing a reduction of the green colour intensity of leaves. Many research works proved the influence of pre-harvest factors on the postharvest quality of vegetables and the cultivation system is one of them (Elia and Colelli, 2009; Frezza et al. 2010). On the contrary, at the end of storage (QL1), the samples cultivated on the two different systems, showed similar values (about 0.4 mg g^{-1}) with a reduction of about 50 % for the samples from soil system. As for the samples grown on soilless system, the chlorophyll content proved to discriminate the very good (QL5) and good (QL4) rocket leaves from fair (QL3), in which a decrease of 16.3 % was recorded. Furthermore, marketable samples (QL3) were also well discriminated from edible (QL2) and waste (QL1) ones. The chlorophyll degradation during the postharvest storage is strictly related to the senescence of the product. At harvest, rocket leaves are dark or bright green in colour, but during senescence changes into yellow, with a general loss of visual quality (Watkins, 2006; Cefola et al. 2010). This process involves many enzymatic reactions in chloroplasts and vacuoles and, particularly, leaves yellowing is strictly related to the activity of chlorophyllase (Matile et al. 1999; Shi et al. 2016; Li et al. 2017). Chlorophyll breakdown take place also when a physiological stress occurs on the tissues, such as the mechanical stress induced by cutting (Toivonen and Brummel, 2008; Torales et al. 2020). Many authors suggested that ethylene production in damaged tissues (such as in fresh-cut products) is responsible for the chlorophyll loss because it causes an increase of the chlorophyllase activity (Yamauchi and Watada, 1991). In the present study, the chlorophyll decrease, resulting in yellowing of leaves at the end of storage at 10 °C, was probably due to the activity of the enzymes related to the chlorophyll degradation during postharvest, in accordance with the results observed by Torales et al. (2020) on fresh-cut rocket leaves during cold stored.

The same QL discrimination was observed in the case of ammonia content for both cultivation systems and for all samples, recording a rapid increase from QL3 to QL1 (Table 2). In particular, this parameter well separated the marketable samples (QL5, QL4 and QL3) from the non-marketable ones (QL2 and QL1); moreover, even the waste samples (QL1) were well discriminated from the edible ones (QL2). At QL1 samples cultivated on soil system showed 56.3 % higher ammonia content than that grown on soilless system. Similar results were reported by Pace et al. (2014) who identify in the ammonia content a good classifier for whole and fresh-cut iceberg lettuce, separating the acceptable product (ranged from QL5 to QL3) from the edible (QL2) and waste (QL1) ones. Ammonia accumulation in plant tissues as a consequence of protein catabolism is another aspect associated with the leaf senescence in leafy vegetables. High concentrations of this compound may cause tissue damages with visible senescence effects, that impact on the overall quality evaluation of the product Chibnall (1939) first reported that ammonia accumulation was the cause of darkening and browning of detached leaves during postharvest, also later demonstrated by Cantwell et al. (2010) on spinach leaves and Mastrandrea et al. (2016) on rocket leaves. Moreover, postharvest chlorophyll degradation may contribute to ammonia accumulation: chlorophyll catabolism consists of the protein-pigment complexes breakage that release the chlorophyll; this causes the degradation of apoproteins by protease and remobilization of the nitrogen of the chlorophyll apoproteins

(Amodio et al. 2018). Mastrandrea et al. (2016) reported high correlations ($R^2 > 0.98$) between changes in ammonia content and hue angle variations (that correspond to yellowing) in rocket leaves stored at 10 °C. Generally, leafy vegetables under stressful condition showed a reduction in the glutamine synthetase activity, an enzyme that leads to the ammonia reintegration during protein synthesis (Chandra et al. 2006). In addition, ammonia accumulation occurs very often in closed systems, such as a package, where may reach high levels. Indeed, in minimally processed products, deteriorative processes like proteolysis are enhanced by injuries occurred during handling steps, especially in highly active products (such as green leafy vegetables) (Wang et al. 2004; Cefola et al. 2010), and by the activity of PAL, that cause lignification of tissues, releasing ammonia (Joy, 1988). Moreover, Yang et al. 1982 demonstrated that also ethylene biosynthesis process from methionine produces little amount of ammonia.

These results support the use of total chlorophyll and ammonia content as objective quality parameters for the assessment of QLs of fresh-cut rocket leaves (Pace et al., 2014; Cavallo et al., 2017). As consequence, their prediction by the CVS, though the package, may represent a valid tool for reducing subjectivity and time cost of manual operations along the whole supply chain, from industrial production lines to the final consumer.

3.2. Non-destructive quality evaluation of packed and unpacked rocket leaves by CVS

The Computer Vision System was applied to estimate the visual QL of packaged fresh-cut rocket and to predict the total chlorophyll and the ammonia content in packaged and unpackaged rocket leaves.

Regarding the first task, table 3, 4, 5 and Figure S1 show the results obtained considering three different cases: i) to separate non-marketable product (QL 1 and 2) from marketable product (QL from 3 to 5); ii) to increase the resolution to separate edible but not marketable product (QL 2) from waste (QL 1); iii) to further increase the resolution to separate also the limit of marketable product (QL 3). In all these cases, the CVS was able to operate through the packaging with negligible loss of accuracy. Moreover, the increase in class separation (from Table 3 to Table 5 and Figure S1) showed a reduction in accuracy on soilless samples (and therefore when all the samples were considered). Instead, the accuracy remained at high level for samples coming from the soil system.

Table 3. Confusion matrix obtained separating 2 class (QL 1-2, QL 3-4-5).

	Quality Level (QL)	***Unpacked*** QL **1-2**	***Unpacked*** QL **3-4-5**	***Unpacked*** *Accuracy* (r)	***Packed*** QL **1-2**	***Packed*** QL **3-4-5**	***Packed*** *Accuracy* (r)
All samples	**1-2**	273	63	0.899	281	55	0.907
	3-4-5	26	517		27	516	
Soil	**1-2**	152	4	0.981	141	15	0.947
	3-4-5	3	270		7	266	
Soilless	**1-2**	147	33	0.893	146	34	0.891
	3-4-5	15	255		15	255	

*Quality level (QL): 5=very good; 4= good; 3=fair; 2=poor; 1=very poor.

*QL3: limit of marketability

Table 4. Confusion matrix obtained separating 3 class (QL 1, QL2, QL 3-4-5).

	Quality Level (QL)	***Unpacked*** QL **1**	***Unpacked*** QL **2**	***Unpacked*** QL **3-4-5**	***Unpacked*** *Accuracy* (r)	***Packed*** QL **1**	***Packed*** QL **2**	***Packed*** QL **3-4-5**	***Packed*** *Accuracy* (r)
All samples	**1**	116	34	18	0.831	122	37	9	0.826
	2	22	86	60		28	79	61	
	3-4-5	1	13	528		0	18	525	
Soil	**1**	72	6	0	0.954	69	9	0	0.905
	2	6	68	4		9	53	16	
	3-4-5	0	4	269		0	6	267	
Soilless	**1**	74	13	3	0.829	69	18	3	0.826
	2	14	40	36		10	43	37	
	3-4-5	0	11	259		0	10	260	

*Quality level (QL): 5=very good; 4= good; 3=fair; 2=poor; 1=very poor.

*QL3: limit of marketability.

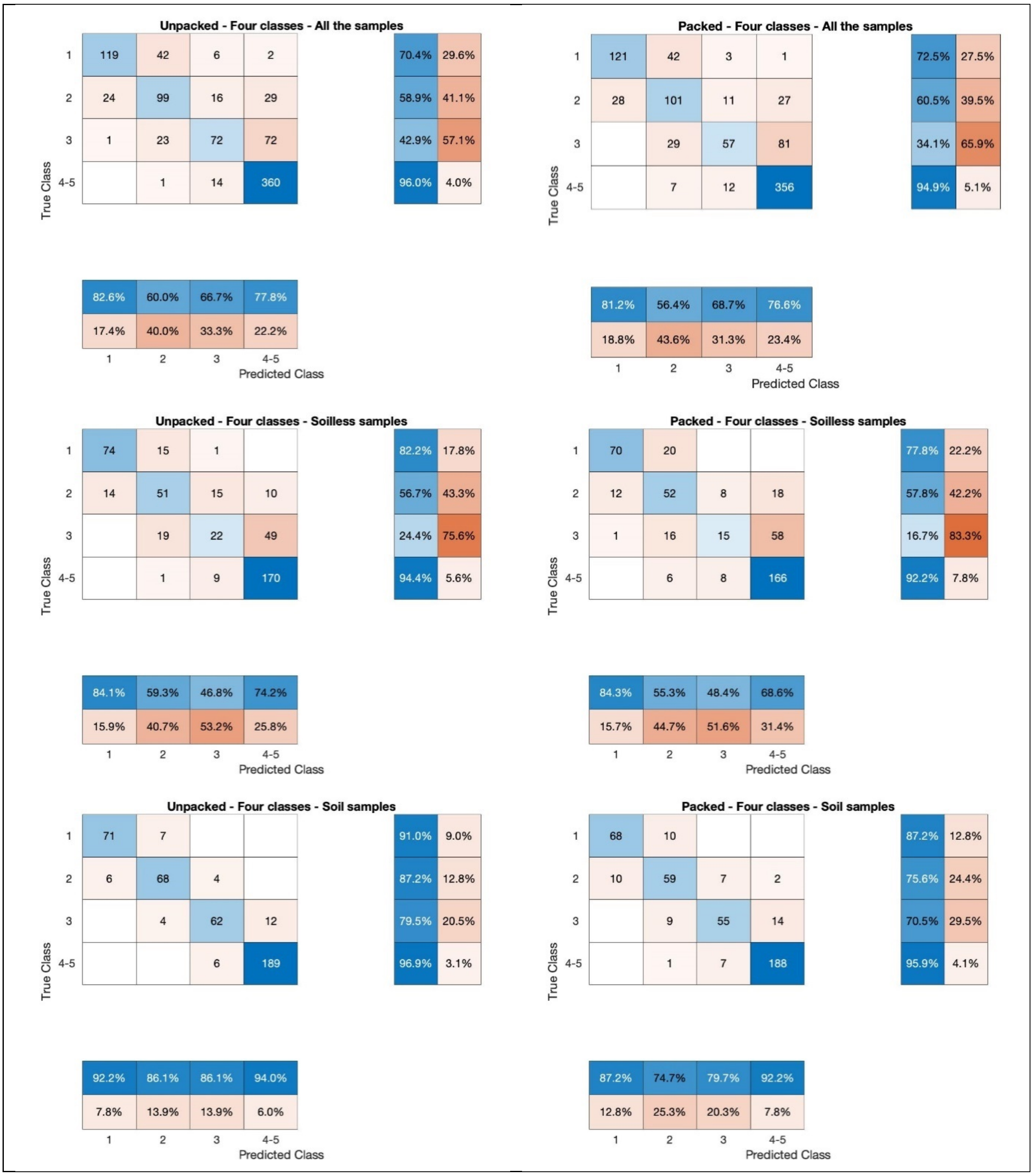
Table 5. Confusion matrix obtained separating 4 class (QL 1, QL2, QL 3, QL 4-5).

	Quality Level (QL)	*Unpacked*					*Packed*				
		QL				*Accuracy* (r)	QL				*Accuracy* (r)
		1	**2**	**3**	**4-5**		**1**	**2**	**3**	**4-5**	
All samples	**1**	119	42	6	2	0.738	121	42	3	1	0.724
	2	24	99	16	29		28	101	11	27	
	3	1	23	72	72		0	29	57	81	
	4-5	0	1	14	360		0	7	12	356	
Soil	**1**	71	7	0	0	0.909	68	10	0	0	0.86
	2	6	68	4	0		10	59	7	2	
	3	0	4	62	12		0	9	55	14	
	4-5	0	0	6	189		0	1	7	188	
Soilless	**1**	74	15	1	0	0.703	70	20	0	0	0.673
	2	14	51	15	10		12	52	8	18	
	3	0	19	22	49		1	16	15	58	
	4-5	0	1	9	170		0	6	8	166	

*Quality level (QL): 5=very good; 4= good; 3=fair; 2=poor; 1=very poor.

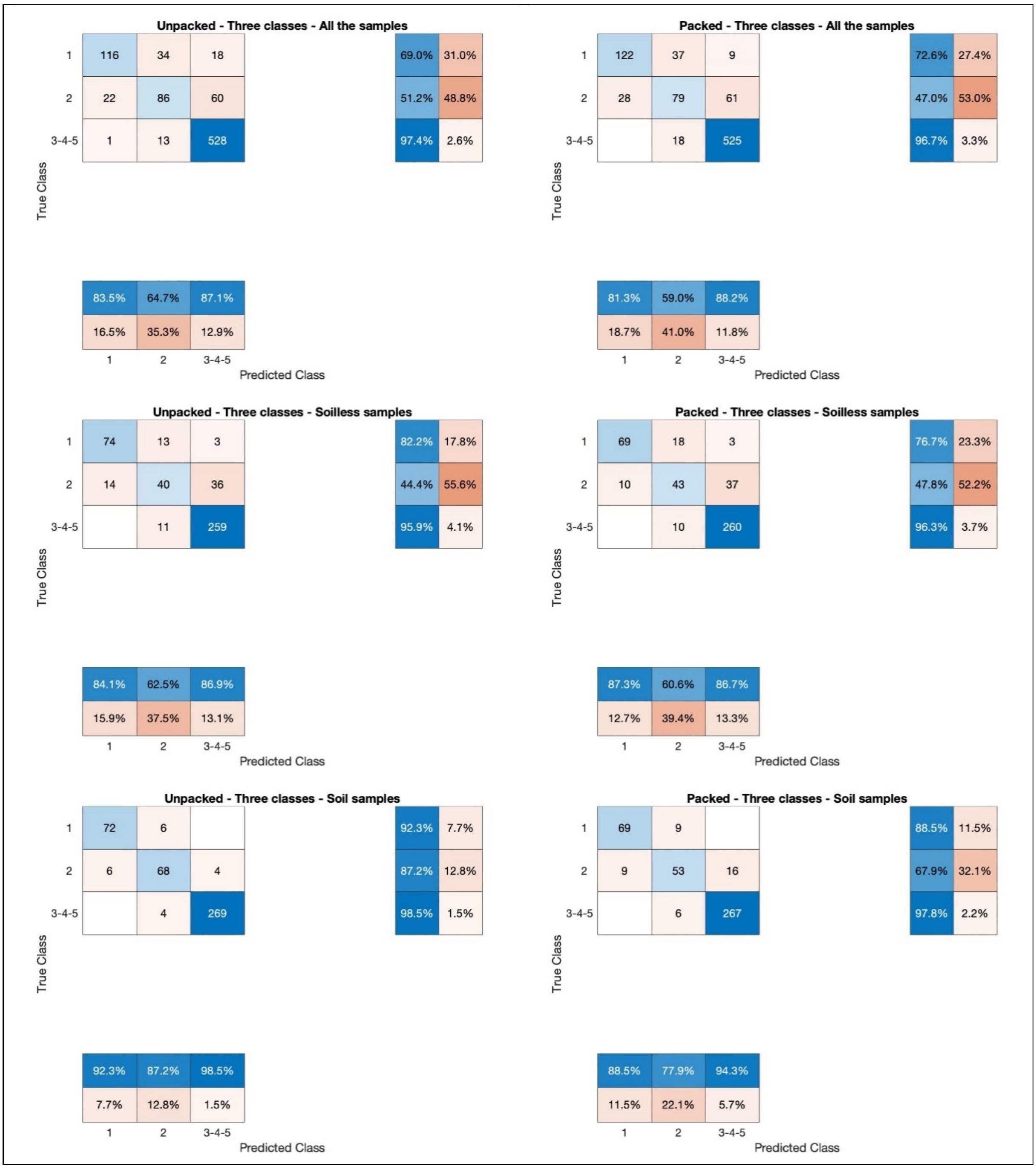
*QL3: limit of marketability

A



Unpacked - Four classes - All the samples
Packed - Four classes - All the samples
Unpacked - Four classes - Soilless samples
Packed - Four classes - Soilless samples
Unpacked - Four classes - Soil samples
Packed - Four classes - Soil samples
True Class
Predicted Class

B


Unpacked - Three classes - All the samples
1
2
3-4-5
116 34 18
22 86 60
1 13 528
69.0% 31.0%
51.2% 48.8%
97.4% 2.6%
True Class
83.5% 64.7% 87.1%
16.5% 35.3% 12.9%
1 2 3-4-5
Predicted Class
Packed - Three classes - All the samples
122 37 9
28 79 61
18 525
72.6% 27.4%
47.0% 53.0%
96.7% 3.3%
81.3% 59.0% 88.2%
18.7% 41.0% 11.8%
Unpacked - Three classes - Soilless samples
74 13 3
14 40 36
11 259
82.2% 17.8%
44.4% 55.6%
95.9% 4.1%
84.1% 62.5% 86.9%
15.9% 37.5% 13.1%
Packed - Three classes - Soilless samples
69 18 3
10 43 37
10 260
76.7% 23.3%
47.8% 52.2%
96.3% 3.7%
87.3% 60.6% 86.7%
12.7% 39.4% 13.3%
Unpacked - Three classes - Soil samples
72 6
6 68 4
4 269
92.3% 7.7%
87.2% 12.8%
98.5% 1.5%
92.3% 87.2% 98.5%
7.7% 12.8% 1.5%
Packed - Three classes - Soil samples
69 9
9 53 16
6 267
88.5% 11.5%
67.9% 32.1%
97.8% 2.2%
88.5% 77.9% 94.3%
11.5% 22.1% 5.7%

C

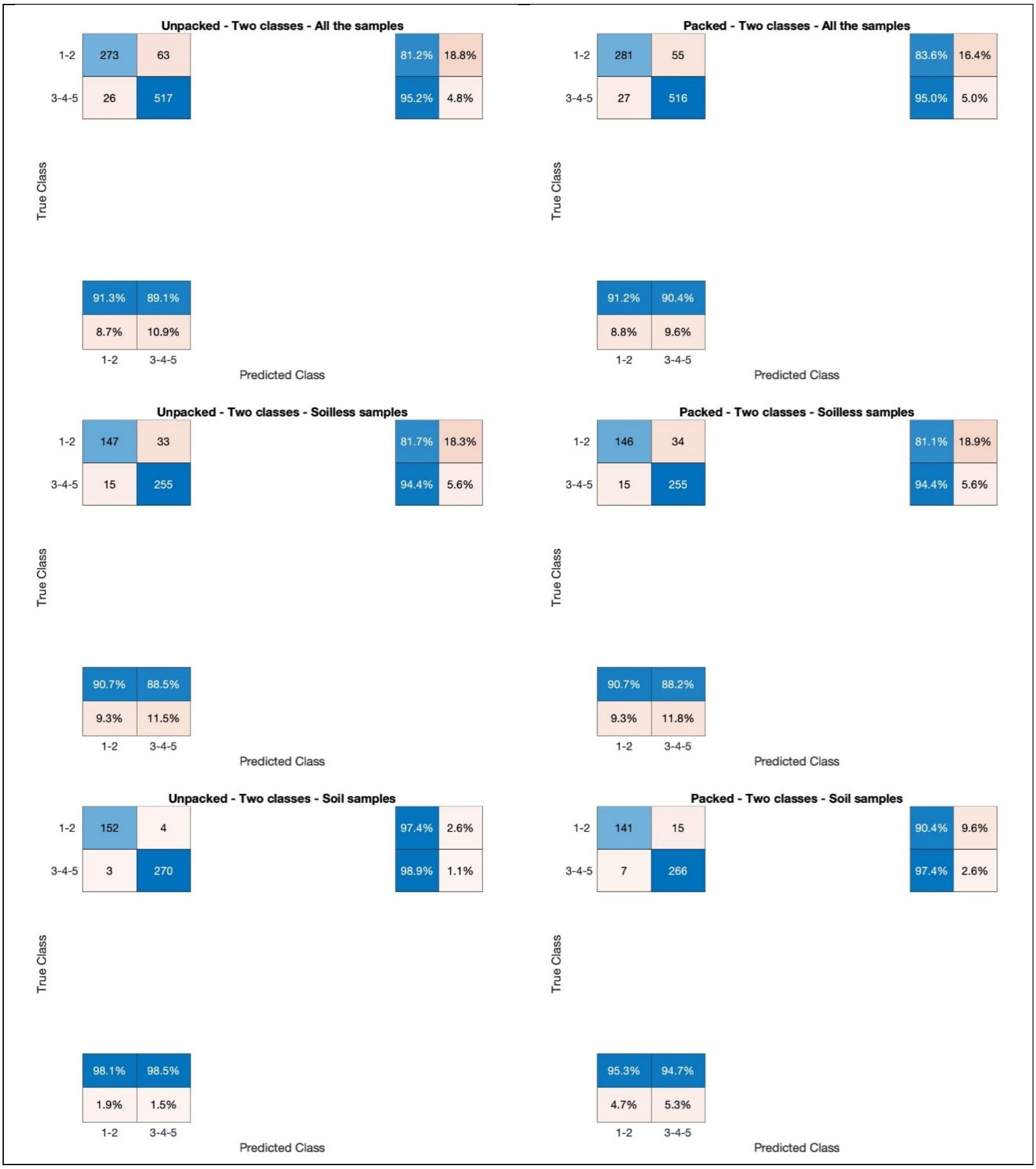


Figure S1. Confusion Matrix obtaining separating four (A), three (B) and two classes (C).

The results (Pearson's correlation coefficient and MSE) obtained by the CVS in estimating the chlorophyll and ammonia contents of all the samples and of packaged or unpackaged items cultivated on soil or soilless are reported in Table 6. For each value, the best result and the average of results obtained over 10 repetitions are reported. The negligible differences between the best values and the average ones show that the randomness used in the construction of the model does not affect significantly the performance of the method. The small differences between the values related to unpackaged and packaged products confirm that the method is able to operate also through the bag.

Table 6. Mean Squared Error (MSE) and Pearson's correlation coefficient (r) measured to predict by CVS total chlorophyll and ammonia content of unpackaged and packaged fresh-cut rocket ($p < 0.0001$).

Samples		***Total Chlorophyll***				***Ammonia Content***			
		Unpackaged		Packaged		Unpackaged		Packaged	
		Best	Average	Best	Average	Best	Average	Best	Average
Soil and Soilless	MSE	77.62	78.65	90.23	91.41	575.1	588.7	643.0	656.3
	r	0.87	0.87	0.84	0.84	0.92	0.92	0.91	0.91
Soil	MSE	69.55	70.45	72.20	74.05	746.9	782.1	872.5	908.0
	r	0.90	0.90	0.90	0.89	0.94	0.94	0.93	0.92
Soilless	MSE	66.52	68.42	83.81	84.68	284.3	289.0	314.6	326.0
	r	0.83	0.83	0.78	0.78	0.87	0.86	0.85	0.84

3.3. Estimation of visual quality level of packed fresh-cut rocket using as predictors total chlorophyll and ammonia measured by conventional methods or by CVS.

Three PLS models were built to predict the visual QL using as predictors the total chlorophyll and the ammonia contents obtained by destructive methods (Model I), CVS trough packaging (Model II) and CVS without packaging (Model III) (Table 7). The predictors estimated non-destructively and contactless by the CVS (Model II and Model III) provide better performances in terms of R^2 in predicting the visual QL than the ones measured by the destructive analysis in the laboratory, in both calibration and validation (Table 7). This is probably due to the wider area of product observed by the CVS, larger than the amount of product used by the destructive methods. It is also evident that the difference between Model II (related to the CVS applied to packaged product) and Model III (related to the CVS applied to unpackaged product) is negligible (Table 7). These results further confirm the ability of the CVS to evaluate the product also through the bag, even working only on the regions of the image that provide meaningful colour information about the product's surface.

Table 7. Root Mean Square Error (RMSE) and the coefficient of determination (R^2) in calibration ($_c$) or validation ($_v$) of the Partial Least Square (PLS) Models predicting visual quality of rocket leaves.

PLS Models	Predictors	$RMSE_c$	R^2_c	$RMSE_v$	R^2_v
I	Total chlorophyll and ammonia obtained by destructive methods	0.45	0.9	0.86	0.70
II	Total chlorophyll and ammonia obtained by CVS on packaged rocket leaves	0.46	0.89	0.75	0.77
III	Total chlorophyll and ammonia obtained by CVS on unpackaged rocket leaves	0.46	0.89	0.7	0.8

4. CONCLUSION

The quality parameters determined by destructive conventional methods during the fresh-cut rocket leaves storage significantly separated the visual quality levels (QL) and, among them, total chlorophyll and ammonia contents were very useful objective marker parameters in the assessment of all the considered QLs.

The CVS was able to operate with not relevant differences on unpackaged and packaged products, enabling these controls at all the steps of the whole supply chain. The proposed CVS is based on the use of calibrated colour images and on the proper combination of image processing techniques and machine learning models. It was able to solve the classification problem (assigning the visual quality level) and the regression problems (estimating the chlorophyll and ammonia contents) using the same supervised learning methodology (Random Forest) applied to proper training data. The results proved that the system can be a valid alternative to conventional destructive measures, offering the advantage of being non-destructive, contactless, fast and cheap. Moreover, experiments showed that the loss in performance due to the observation of product through the packaging is negligible. The PLS models built to predict the visual QL using as predictors total chlorophyll and ammonia content further confirmed the ability of the CVS to operate also through the packaging. The research was carried out using the polypropylene, widely used to store fresh-cut salad. Further experiments are needed to verify the performance of CVS on other type of plastic materials for packaging.

The developed CVS represents a simple, cheap, fast, non-destructive and contactless tool for a continuous monitoring of the quality and of the level of freshness of packaged salad along the whole supply chain, from harvest to final consumers, in an objective way and based on non-destructive measurement of biological markers (such as chlorophyll and ammonia) which have been shown to be strongly related to leaf senescence. It might represent a valid tool for producers and consumers to standardize quality levels and to timely detect senescence enabling waste reduction, sales optimization and customer satisfaction.

REFERENCES

Agarwal, A., Gupta, S.D., 2018. Assessment of spinach seedling health status and chlorophyll content by multivariate data analysis and multiple linear regression of leaf image features. *Computers and Electronics in Agriculture,* 152, 281-289.

Amodio, M.L., Colelli, G., Cantwell, M.I., 2018. Ammonia accumulation in plant tissues: a potentially useful indicator of postharvest physiological stress. *Acta Horticulturae*, 1194, 1511–1518.

Baresel, J. P., Rischbeck, P., Hu, Y., Kipp, S., Barmeier, G., Mistele, B., & Schmidhalter, U. (2017). Use of a digital camera as alternative method for non-destructive detection of the leaf chlorophyll content and the nitrogen nutrition status in wheat. *Computers and Electronics in Agriculture*, 140, 25-33.

Bonasia, A., Lazzizera, C., Elia, A., Conversa, G., 2017. Nutritional, biophysical and physiological characteristics of wild rocket genotypes as affected by soilless cultivation system, salinity level of nutrient solution and growing period. *Frontiers in Plant Science*, 8, 300.

Breiman, L., 1996. Bagging predictors. *Machine Learning*, 24, 123–140

Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5–32

Cantwell, M., Hong, G., Nie, X., 2009. Using tissue ammonia and fermentative volatile concentrations as indicators of beneficial and stressful modified atmospheres for leafy and floral vegetables. In: Amodio, M.L., Colelli, G. (Eds.), *X International Controlled and Modified Atmosphere Research Conference. Acta Horticulturae*, 876, pp. 165-172.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2018. Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material. *Journal of Food Engineering*, 223, 46-52.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2017. Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. *Computers and Electronics in Agriculture,* 140, 303-310.

Cefola, M., Pace, B., 2015. Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. *Journal of Food Processing*, 39(6), 2523-2532.

Cefola, M., Amodio, M.L., Rinaldi, R., Vanadia, S., Colelli, G., 2010. Exposure to 1-methylcyclopropene (1-MCP) delays the effects of ethylene on fresh-cut broccoli raab (Brassica rapa L.). *Postharvest Biology and Technology,* 58(1), 29-35.

Chandra, D., Matsui, T., Suzuki, H., Kosugi, Y., 2006. Postharvest changes in some physiological traits and activities of ammonia-assimilating enzymes in lettuce during storage. *Asian Journal of Plant Science*, 5 (2), 378–384.

Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., Li, B., 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment*, 114(9), 1987-1997.

Chibnall, A.C., 1939. The role of proteins in the respiration of detached leaves. In: Chibnall, A.C. (Ed), *Protein Metabolism in the Plant* (pp. 211–243). Yale University Press, New Haven, CT, USA: Yale University Press.

Colelli, G., Elia, A., 2009. I prodotti ortofrutticoli di IV gamma: aspetti fisiologici e tecnologici. *Italus Hortus*, 16(1), 55-78.

Fadda, A., Pace, B., Angioni, A., Barberis, A., Cefola, M., 2016. Suitability for ready-to-eat processing and preservation of six green and red baby leaves cultivars and evaluation of their antioxidant value during storage and after the expiration date. *Journal of Food Processing*, 40(3), 550-558.

Ferrante, A., Maggiore, T., 2007. Chlorophyll *a* fluorescence measurements to evaluate storage time and temperature of Valeriana leafy vegetables. *Postharvest Biology and Technology*, 45(1), 73-80.

Frezza, D., Logegaray, V.R., León, A.P., Harris, M., Chiesa, A., 2009. Rocket (*Eruca sativa* Mill.) quality affected by preharvest and postharvest factors. In: Kanlayanarat, S., McGlasson, W.B., Techavuthiporn, C. (Eds.), *Southeast Asia Symposium on Quality and Safety of Fresh and Fresh-Cut Produce* (pp. 357-364). *Acta Horticulturae*, 875.

Joy, K.W., 1988. Ammonia, glutamine, and asparagine: a carbon–nitrogen interface. *Canadian Journal of Botany*, 66(10), 2103-2109.

Kou, L., Yang, T., Luo, Y., Liu, X., Huang, L., Codling, E., 2014. Pre-harvest calcium application increases biomass and delays senescence of broccoli microgreens. *Postharvest Biology and Technology,* 87, 70-78.

Li, D., Li, L., Ge, Z., Limwachiranon, J., Ban, Z., Yang, D., Luo, Z., 2017. Effects of hydrogen sulfide on yellowing and energy metabolism in broccoli. *Postharvest Biology and Technology,* 129, 136-142.

Li, F., Mistele, B., Hu, Y., Chen, X., Schmidhalter, U., 2014. Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. *European Journal of Agronomy*, 52, 198-209.

Limantara, L., Dettling, M., Indrawati, R., Brotosudarmo, T.H.P., 2015. Analysis on the chlorophyll content of commercial green leafy vegetables. *Procedia Chemistry*, 14, 225-231.

Ling, Q., Huang, W., Jarvis, P., 2011. Use of a SPAD-502 meter to measure leaf chlorophyll concentration in Arabidopsis thaliana. *Photosynthesis Research*, 107(2), 209-214.

Liu, Z.A., Yang, J.P., Yang, Z.C., 2012. Using a chlorophyll meter to estimate tea leaf chlorophyll and nitrogen contents. *Journal of Soil Science and Plant Nutrition*, 12(2), 339-348.

Mastrandrea, L., Amodio, M.L., Cantwell, M.I., 2016. Modeling ammonia accumulation and colour changes of arugula (*Diplotaxis tenuifolia*) leaves in relation to temperature, storage time and cultivar. *Acta Horticulturae*, 1141, 275-282.

Matile, P., Hörtensteiner, S., Thomas, H., 1999. Chlorophyll degradation. *Annual Review of Plant Biology,* 50(1), 67-95.

Mohan, P.J., Gupta, S.D., 2019. Intelligent image analysis for retrieval of leaf chlorophyll content of rice from digital images of smartphone under natural light. *Photosynthetica*, 57(2), 388-398.

Novichonok, E.V., Novichonok, A.O., Kurbatova, J.A., Markovskaya, E.F., 2016. Use of the atLEAF+ chlorophyll meter for a nondestructive estimate of chlorophyll content. *Photosynthetica*, 54(1), 130-137.

Pace, B., Cavallo, D.P., Cefola, M., Burbaci, S., Attolico, G., 2019. Colour analysis to predict the total chlorophyll content of rocket leaves. In: Amodio, M.L., Colelli, G. (Eds.), *VI International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain*. Acta Horticulturae (pp. 107-112), Model-IT, 1311.

Pace, B., Cavallo, D.P., Cefola, M., Attolico, G., 2017. Automatic identification of relevant colours in non-destructive quality evaluation of fresh salad vegetables. *International Journal of Food Processing Technology*, 4(1), 1-5.

Pace, B., Cavallo, D.P., Cefola, M., Colella, R., Attolico, G., 2015. Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio. *Innovative Food Science and Emerging Technologies,* 32, 200-207.

Pace, B., Cefola, M., Da Pelo, P., Renna, F., Attolico, G., 2014. Non-destructive evaluation of quality and ammonia content in whole and fresh-cut lettuce by computer vision system. *Food Research Internatioanal*, 64, 647-655.

Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Serio, F., Colelli, G., Attolico, G., 2021. Self-configuring CVS to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. *Agronomy*, 11(7), 1353.3

Pathare, P.B., Opara, U.L., Al-Said, F.A.J., 2013. Colour measurement and analysis in fresh and processed foods: a review. *Food and Bioprocess Technology*, 6(1), 36-60.

Rigon, J.P.G., Capuani, S., Fernandes, D.M., Guimarães, T.M., 2016. A novel method for the estimation of soybean chlorophyll content using a smartphone and image analysis. *Photosynthetica*, 54(4), 559-566.

Santos do Amaral, E., Silva, D.V., Dos Anjos, L., Schilling, A.C., Dalmolin, Â.C., Mielke, M.S., 2019. Relationships between reflectance and absorbance chlorophyll indices with RGB (Red, Green, Blue) image components in seedlings of tropical tree species at nursery stage. *New Forests*, 50(3), 377-388.

Shi, J., Gao, L., Zuo, J., Wang, Q., Wang, Q., Fan, L., 2016. Exogenous sodium nitroprusside treatment of broccoli florets extends shelf life, enhances antioxidant enzyme activity, and inhibits chlorophyll-degradation. *Postharvest Biology and Technology* 116, 98-104.

Toivonen, P. M., & Brummell, D. A. (2008). Biochemical bases of appearance and texture changes in fresh-cut fruit and vegetables. *Postharvest Biology and Technology*, 48(1), 1-14.

Torales, A.C., Gutiérrez, D.R., Rodríguez, S.D.C., 2020. Influence of passive and active modified atmosphere packaging on yellowing and chlorophyll degrading enzymes activity in fresh-cut rocket leaves. *Food Packaging*, 26, 100569.

Tudela, J.A., Marín, A., Garrido, Y., Cantwell, M., Medina-Martínez, M.S., Gil, M.I., 2013. Off-odour development in modified atmosphere packaged baby spinach is an unresolved problem. *Postharvest Biology and Technology*, 75, 75-85.

Vesali, F., Omid, M., Mobli, H., Kaleita, A., 2017. Feasibility of using smart phones to estimate chlorophyll content in corn plants. *Photosynthetica*, 55(4), 603-610.

Wang, Y., Wang, D., Shi, P., Omasa, K., 2014. Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still colour camera under natural light. *Plant methods*, 10(1), 1-11.

Wang, Y., Wang, D., Zhang, G., Wang, J., 2013. Estimating nitrogen status of rice using the image segmentation of GR thresholding method. *Field Crops Research*, 149, 33-39.

Wang, Y.T., Yang, C.Y., Chen, Y.T., Lin, Y., Shaw, J.F., 2004. Characterization of senescence-associated proteases in postharvest broccoli florets. *Plant Physiology and Biochemistry*, 42(7-8), 663-670.

Watkins, C.B., 2006. The use of 1-methylcyclopropene (1-MCP) on fruits and vegetables. *Biotechnology*, Adv. 24(4), 389-409.

Wellburn, A.R., 1994. The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology*, 144(3), 307-313.

Yamauchi, N., Watada, A.E., 1991. Regulated chlorophyll degradation in spinach leaves during storage. *Journal of the American Society for Horticultural Science*, 116(1), 58-62.

Yang, S.F., Adams, D.O., Lizada, C., Yu, Y., Bradford, K.J., Cameron, A.C., Hoffman, N.E., 1982. Mechanism and regulation of ethylene biosynthesis. In: *Plant growth substances* (pp. 219-229). Springer, Berlin, Heidelberg.

Yuan, Z., Ata-Ul-Karim, S.T., Cao, Q., Lu, Z., Cao, W., Zhu, Y., Liu, X., 2016. Indicators for diagnosing nitrogen status of rice based on chlorophyll meter readings. *Field Crops Research*, 185, 12-20.

Yue, X. L., Hu, Y., Zhang, H. Z., & Schmidhalter, U. (2015). Green Window Approach for improving nitrogen management by farmers in small-scale wheat fields. *Journal of Agricultural Science and Technology*, 153(3), 446-454.

Chapter 5

MACHINE LEARNING FOR THE IDENTIFICATION OF COLOUR CUES TO ESTIMATE QUALITY PARAMETERS OF ROCKET LEAVES

Michela Palumbo[1,2], Maria Cefola[1], Bernardo Pace[1], Giancarlo Colelli[2], Giovanni Attolico[3],

[1] Institute of Sciences of Food Production, CNR, c/o CS-DAT, Via Michele Protano, 71121 Foggia

[2] Department of Agriculture, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Via Napoli 25, 71122 Foggia.

[3] Institute on Intelligent Industrial Systems and Technologies for Advanced Manufacturing, CNR, Via G. Amendola, 122/O, 70126 Bari

ABSTRACT

Computer Vision Systems (CVSs) have proved to be a powerful tool to evaluate the quality of agricultural products in a non-destructive, contactless and objective way. Machine learning techniques are increasingly relevant to simplify the development of CVS with better performance and greater flexibility in matching the requirements of different products and environmental characteristics. An interesting research field is to exploit the benefits of learning while keeping the resulting solutions simple, fast and interpretable by humans. One of these benefits would be to receive from learning techniques detailed hints about specific visual characteristics that correlate with relevant properties of products. By analysing a Random Forest model developed in previous experiments to classify visual quality and to estimate chlorophyll and ammonia contents in rocket leaves, the proposed method explicitly identifies specific informative colours, enabling interesting applications. First of all, the design of simpler and faster algorithms to extract relevant colour features related to significant visual characteristics of the product at hand. In addition, the objective identification of relevant colours can improve the objectivity and soundness of pictures and textual descriptions used to train human operators working in quality control. Finally, results obtained on real images of rocket leaves are shown, compared with previously obtained results, and discussed.

Keywords: machine learning techniques, colour features identification, regression techniques, chlorophyll and ammonia prediction.

1. INTRODUCTION

Recently, researchers have focused on the use of contactless, non-destructive, rapid, accurate and more sustainable techniques to objectively assess sensory and compositional quality of fruit and vegetables. Nevertheless, although these non-invasive methods offer significant advantages compared to analytical and destructive analysis, they cannot completely replace them. They are complementary, enabling time and cost saving, continuous and reliable monitoring and reduction of impact on environment along the supply chain (Chaudhry et al., 2020). Computer Vision Systems (CVSs) represent an innovative and contactless non-destructive technology suitable for in-line grading and quality assessment of fruit and vegetables (Fan et al., 2020). In a previous work (Palumbo et al., 2022), a machine learning model based on the Random Forest methodology was used to solve a classification problem (assessment of quality level of rocket leaves) and two regression problems (estimation of chlorophyll and ammonia content of rocket leaves). The system proved to be successful in working on packaged and unpackaged products. The Random Forest approach is based on the combination of several weak learners, each configured as a tree, to accomplish the classification and regression tasks of interest. Each weak learner is a single tree, trained to accomplish the same task of the complete forest using random subsets of the available training samples and of the considered features. For each tree, the training process selects the most relevant features, among the available ones, for achieving the required results (Breiman, 1984). The central hypothesis is that the combination of several weak learners can provide more powerful and robust results. The resulting Random Forest model is generally reasonably efficient and effective but exhibits a significant conceptual complexity that can prevent its application in some operational contexts. Moreover, it is hard to be read and interpreted by humans.

This paper aims to overcome these limitations, starting from the results achieved on packaged and unpackaged rocket leaves by the CVS reported in Palumbo et al. (2022). In detail, an analysis of the features selected by the Random Forest methodology is used to identify compact yet efficient colour cues that can be used by simpler classifiers and regressors. Due to the nature of the original universe of features (frequencies of colour occurrence in the product), the subsets correspond to colour regions that are most informative about the nature of the product. These regions can be exploited in several way: each of them can provide a single measure as a feature for more simple and efficient classifiers or regressors; each region can be associated to a specific colour that can be correlated to specific characteristics of the products and whose change during the storage can be correlated to well-known chemical or physical processes;

moreover, they can provide a sound and objective base for the description of quality marker parameters of the product, making more robust their textual and iconographic descriptions.

2. MATERIALS AND METHODS

2.1. Identification and preliminary selection of clusters

Images of packaged and unpackaged rocket leaves, acquired as reported in Palumbo et al. (2022), were used to extract specific informative colours using the methods reported below.

The set of trees composing the Random Forest model described in Palumbo et al. (2022) was analysed. The vocabulary of features provided to the Random Forest was composed by the elements of the histogram of the images in the *ab*-plane of the CIELab colour space: each feature represented the percentage of presence of a specific colour. The importance of each colour feature was estimated on the base of its use to generate splits in the trees and of its association with other features (Loh, 1997, 2002). The features were then sorted by decreasing order of importance. The n most important features (with n empirically set to 36) were selected providing a set of 36 colours (Figure 1) representing the most relevant colours used to build the trees of the Random Forest model. In fact, each feature is associated to a point in the *ab*-plane of the CIELab colour space. Features purposely do not contain information about the L component: past experiments have shown that the L channel is too sensitive to illumination levels and also to uneven distribution of light across the scene. Therefore, L has been removed from colour description while building the vocabulary of features for the Random Forest model.

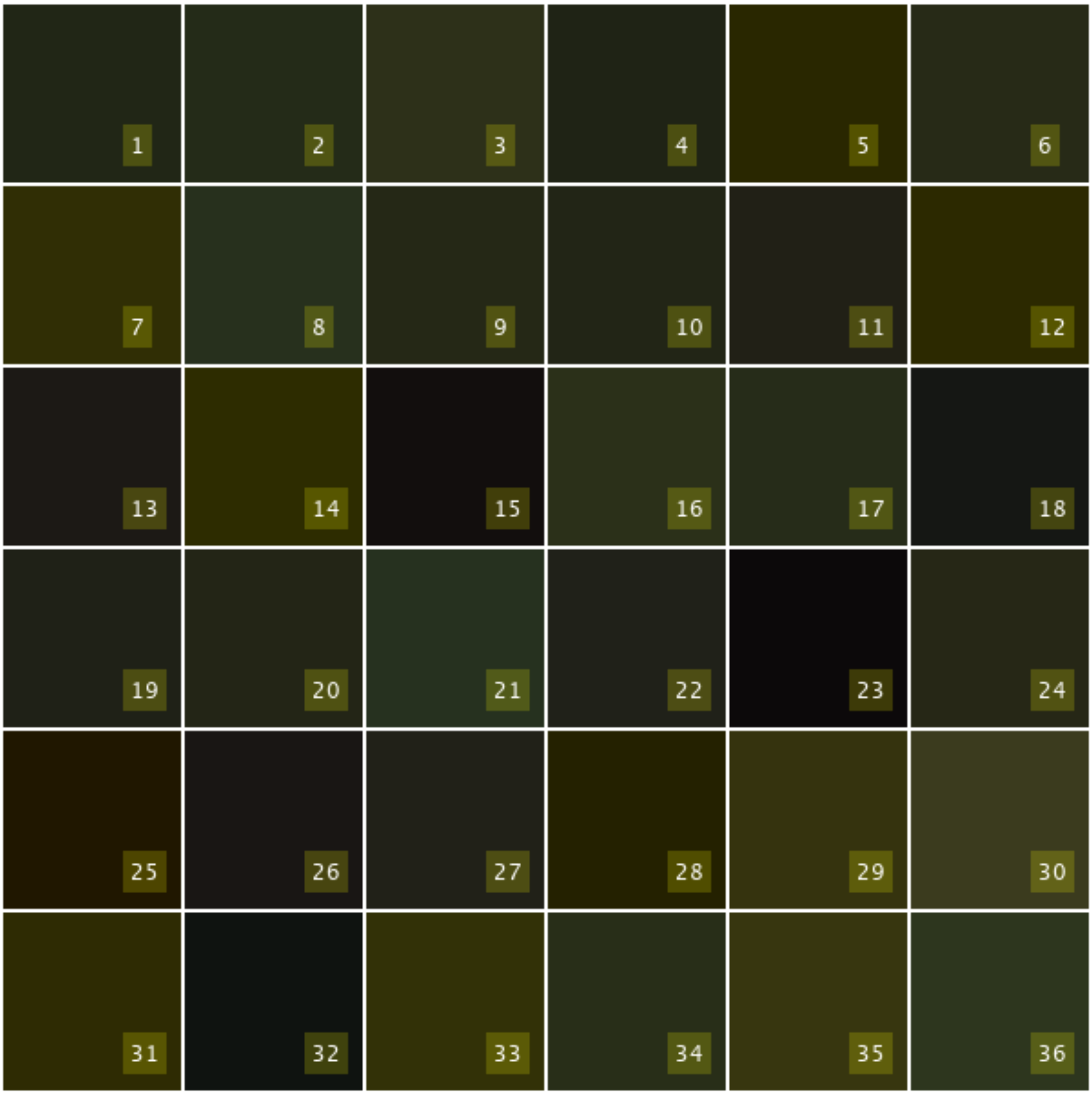


Figure 1. The 36 most important colours identified by the analysis of the Random Forest model.

A hierarchical clustering approach was used to identify 11 clusters out of the 36 colours and reported in Figure 2 using both rectangular (A) and polar (B) representations (Hastie, 2009). Since most clusters were composed by single isolated colours therefore, they have been considered less important to identify colour regions of interest for our work.

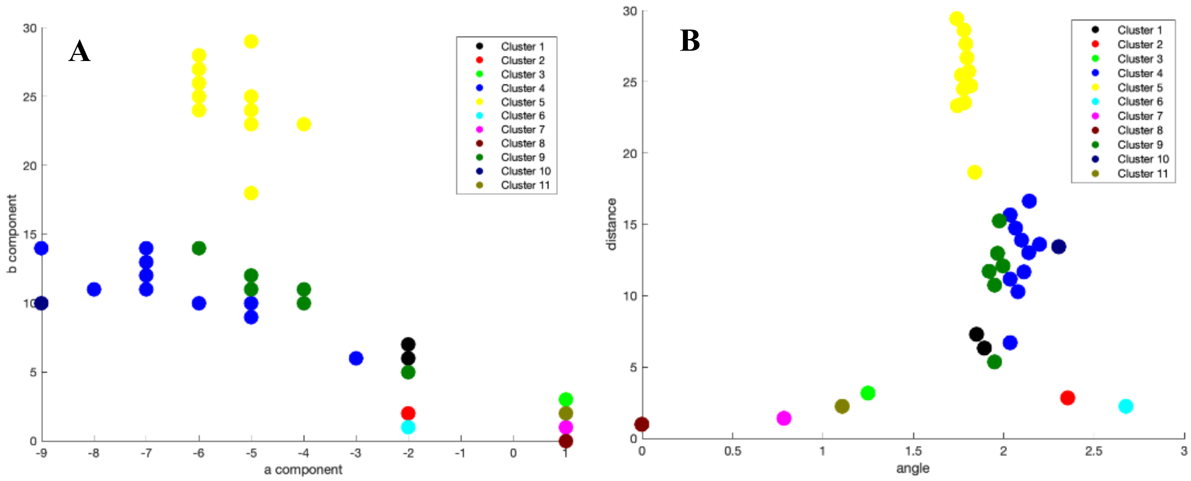
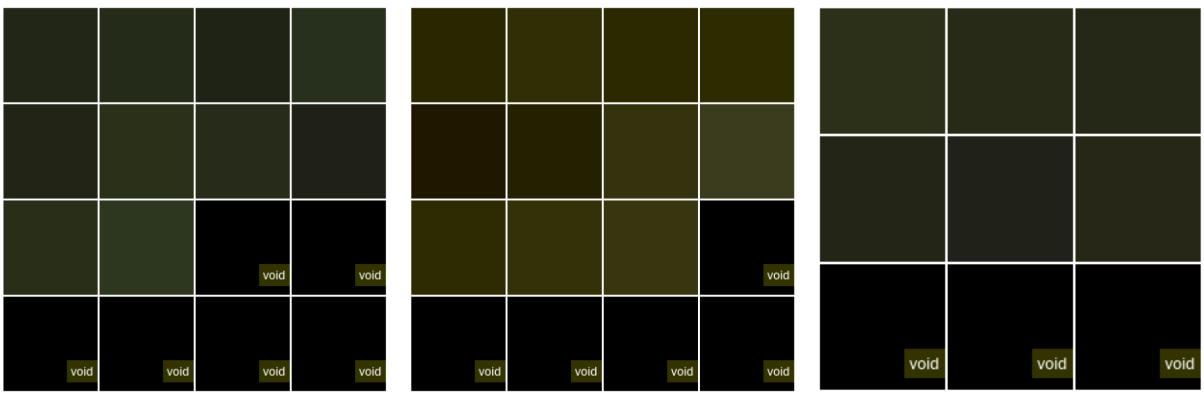


Figure 2. The clusters of the relevant colours as identified by the analysis of the Random Forest model. In A, the abscissa and ordinate axes represents respectively the *a* and *b* components of the colours in the rectangular representation of the *ab*-plane in the CIELab colour space. In B, the same points are represented in polar coordinates, where the ordinate represents the distance of the point from the origin (which is an achromatic point) while the abscissa represents the angle with respect to the line having *b* equal to 0. In the latter representation, the elements of each cluster lie inside a narrow region of the x axis.

Three clusters, whose colours are shown in Figure 3, collected a larger number of elements, namely the cluster 4 (C4) which contains 10 colours, the cluster 5 (C5) with 11 colours, and the cluster 9 (C9) with 6 colours. Moreover, the rectangles defined by the minimum and maximum values of the two components of the colours of each cluster in both rectangular and polar representations are shown in Figure 4.



Cluster 4 **Cluster 5** **Cluster 9**

Figure 3. Colours belonging to the most relevant clusters 4, 5 and 9. Statistical analysis showed that the clusters 5 and 9 are more relevant for quality assignment and parameters estimation.

Please note that the black cells marked as *void*, are empty and purposely introduced in the image to simplify its construction.

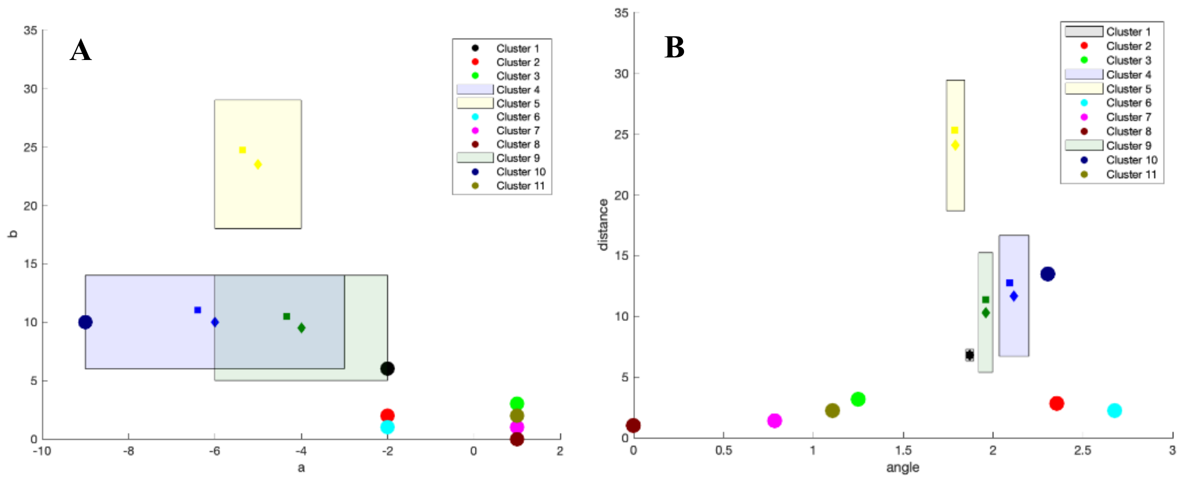


Figure 4. The figure shows the rectangular regions defined by the minimum and maximum values of the colours belonging to clusters 4, 5 and 9 in the rectangular (A) and polar representation (B). In both figures, the symbol of diamond represents the mean point of each region while the square represents the centroid of the same region. In the polar representation the regions related to different clusters have an empty intersection while they largely overlap in the rectangular representation.

This approach tried to associate a region of the *ab*-plane to each of these cluster: the aim was to extract a single feature related to this region and to use this value for the following processing with the purpose of reducing the number of features and of simplifying the processing required to evaluate them. The most natural value was the percentage of pixels of the image whose colours lied in the region.

2.2. Methodologies for the definition of the colour region corresponding to each cluster

A first statistical analysis (data not shown), done using a one-way ANOVA, pointed out the greater relevance of C5 and C9 with respect to C4 which was therefore discarded in the following experiments in which only two features, associated to the frequencies of colours belonging to the C5 and C9 respectively, were evaluated. It needed to define the mechanism to compute these frequencies from each image. In particular, it was critical to decide the strategy used to assign the colour of each pixel in each image to the proper cluster. Several strategies could be used and were analyzed to identify the best performing model. Firstly, it was observed that C5 was well separated by any other cluster both in the polar and rectangular representations of colours. On the contrary, C9 was very close to the C4 in both the representations. While to correctly assign colours to the C9 seemed feasible in the polar space, the same task appeared much more challenging in the rectangular space. Nevertheless, it was decided to analyze the behavior of each strategy in both representations to experimentally verify how much suitable was each of them for final tasks.

The second critical choice for computing the feature associated to each cluster, was the strategy to identify the extension and the shape of the region associated to each cluster: this choice proved to affect the efficiency both in terms of computational complexity and flexibility. Few approaches were compared. Let it define as (a_i, b_i) with $i = 1, \ldots, 36$, the rectangular coordinates of the selected colours in the *ab*-plane. Every point in the plane can be represented also using the polar representation:

$$distance_i = \sqrt{a_i^2 + b_i^2} \quad angle_i = arctg\ (\frac{b_i}{a_i})$$

For simplicity, all the formulas in the rest of the paper will be written only once using the symbols *x* and *y* but they remain valid and can be applied to both the representations, providing results that just need to be interpreted according to the used representation. For the rectangular representation *x* stands for *a* and *y* for *b* while for the polar representation *x* stands for *angle* and *y* for *distance*. In fact, every method has been applied into both the coordinates systems to compare their ability to characterize the colour regions of interest. Let it denote:

$$min_x^k = \min\ \{x_i | i \in cluster\ k\} \qquad max_x^k = \max\ \{x_i | i \in\ cluster\ k\}$$

$$min_y^k = \min\ \{y_i | i \in cluster\ k\} \qquad max_y^k = \max\ \{y_i | i \in\ cluster\ k\}$$

The method 1 (M1) associates to each cluster a rectangular region in the *ab*-plane whose limits are the minimum and maximum values of the two components of the colours, out of the 36, belonging to that cluster. The upper-left and bottom-right corners of the region have respectively coordinates:

$$(min_x^k, max_y^k) \quad \text{and} \quad (max_x^k, min_y^k)$$

It can easily be seen that this method can separate the clusters in the polar representation, but it does not work in the rectangular representation where the rectangular regions related to C4 and C9 largely overlap (Figure 4).

The method 2 (M2) represents each region using the central point, that is the point having as coordinates the mean values between the minimum and the maximum of x and y. The coordinates of the central point of the cluster k are:

$$\left(\frac{min_x^k + max_x^k}{2}, \frac{min_y^k + max_y^k}{2}\right)$$

The method 3 (M3) represents each region using the centroid of the colours belonging to the cluster. The coordinates of the centroid of the cluster k are:

$$\left(\frac{\sum_i x_i}{n_k}, \frac{\sum_i y_i}{n_k}\right)$$

where $i \in cluster\ k$ and n_k is the number of colours, out of the 36, associated to the cluster k. For the M3, two different variants have been used. In the first one (M3a), only the centroid points of C5 and C9 have been considered when looking for the proper cluster for a colour in the image: each pixel was assigned to the cluster (out of these two) whose representative point (centroid) was closest to the colour of the pixel. In the second one (M3b), all the centroid points of all the clusters have been considered when looking for the proper cluster for a colour in the image: each pixel was assigned to the cluster (out of all the 11 clusters) whose representative point (centroid) was closest to the colour of the pixel. Even in this last case, only the values corresponding to C5 and C9 were considered for further processing. The difference between the two variants is that in the first variant (a) all the pixels are assigned either to C5 or to C9, according to which one is the closest. In the second case (b), a significative number of pixels were assigned to other clusters (different from C5 or C9) and were not considered in further processing. In geometrical terms, variant (b) reduces the size and modifies the shape of the regions of the *ab*-plane assigned to each of the two clusters of interest (C5 and C9).

The method 4 (M4) assigns each pixel to the colour, out of the 36, which is the closest in the *ab*-plane. Then, all the pixels associated to colours belonging to the same cluster are cumulated to evaluate the number of pixels belonging to that cluster. This last method enables a finer definition of the colour region associated to each cluster.

For each method, two different versions, using the polar and the rectangular representations respectively, have been compared. All the methods reported above were applied on the images of unpackaged and packaged samples of rocket leaves acquired by the CVS in Palumbo et al. (2022). The features corresponding to the different clusters were normalized: they were divided by the total number of foreground pixels in the image.

2.3.Statistical analysis

The values corresponding to the most relevant clusters (C5 and C9) were subjected to a one-way ANOVA analysis to find significant relationships with the quality level (QL) scores of rocket leaves reported in Palumbo et al. (2022).

The mean values were separated using the Student-Newman-Keuls (SNK) test and Statgraphics Centurion (version 18.1.12, Warrenton, Virginia, USA) was used for statistical analyses.

Principal component analysis (PCA) was performed by the software Statistica (version 6.0, StatSoft, Inc., Tulsa, OK, USA), using as variables the values of the C5 and C9 obtained by all the methods described above, in both the polar or rectangular versions (Method 1, Method 2, Method 3a, Method 3b, Method 4) and the chemical data of total chlorophyll and ammonia content previously reported (Palumbo et al., 2022). While, as the case, data were mediated in two visual quality group: 5-4-3 (marketable) and 2-1 (unmarketable).

Significant correlations were highlighted between each method and the chemical data (total chlorophyll or ammonia content) reported in Palumbo et al., (2022). In particular, the correlation matrices based on the Pearson correlation coefficient were explored by an heatmap and the level $p = 0.05$ was assumed significant for the correlation coefficients. Data analysis was carried out using the software Statistica (version 6.0, StatSoft, Inc., Tulsa, OK, USA). Moreover, a partial least square regression (PLSR) analysis was carried out to predict the total chlorophyll or ammonia content using The Unscrambler X software (CAMO AS, Oslo, Norway).

3. RESULTS AND DISCUSSIONS

3.1. Selection of methods associated to rocket leaves marketability

Significant relationships among the values of C5 and C9 achieved by the 4 methods and the QL scores attributed to rocket leaves during the cold storage (Palumbo et al. 2021) are reported in Table 1.

Table 1. Relationships between the quality levels of packaged and unpackaged rocket leaves and the studied clusters 5 (C5) and 9 (C9) obtained by 4 different methods (M1, M2, M3a, M3b, M4) in both polar and rectangular representations. The values represent, for each method, the mean of presence (in terms of percentage, %) of image colours belonging to the studied clusters.

Method	*Representation*	*Clusters*	*Quality Level*										*P*-value
			5 *very good*		**4** *good*		**3*** *fair*		**2** *poor*		**1** *very poor*		
			Presence percentage (%)										
M1	**polar**	5	0.00053	c	0.00075	c	0.00089	c	0.00684	b	0.01638	a	****
		9	0.12830	c	0.16343	a	0.12125	c	0.16444	a	0.14662	bc	****
	rectangular	5	0.00074	c	0.00103	c	0.00131	c	0.00916	b	0.01927	a	****
		9	0.39882	b	0.44795	a	0.37086	c	0.37985	bc	0.34676	d	****
M2	**polar**	5	0.20719	e	0.22294	d	0.23866	c	0.33814	b	0.37348	a	****
		9	0.79280	a	0.77705	b	0.76133	c	0.66185	d	0.62652	e	****
	rectangular	5	0.15652	d	0.17772	c	0.18929	c	0.29372	b	0.33330	a	****
		9	0.84347	a	0.82227	b	0.81076	b	0.70627	c	0.66669	d	****
M3a	**polar**	5	0.16327	d	0.17472	d	0.18858	c	0.27137	b	0.30264	a	****
		9	0.83672	a	0.82527	a	0.81141	b	0.72863	c	0.69735	d	****
	rectangular	5	0.12073	d	0.13578	c	0.14621	c	0.23119	b	0.26607	a	****
		9	0.87926	a	0.86421	b	0.85378	b	0.76880	c	0.73392	d	****
M3b	**polar**	5	0.17793	d	0.18900	d	0.20609	c	0.29534	b	0.32366	a	****
		9	0.22341	a	0.19918	b	0.19192	b	0.14904	c	0.13697	c	****
	rectangular	5	0.12913	d	0.14746	c	0.16002	c	0.26392	b	0.30112	a	****
		9	0.16967	b	0.18999	a	0.14444	c	0.18574	a	0.19132	a	****
M4	**polar**	5	0.18734	e	0.20148	d	0.21638	c	0.30907	b	0.34268	a	****
		9	0.20359	bc	0.22448	a	0.19776	c	0.20940	b	0.20097	c	****
	rectangular	5	0.09040	d	0.11072	c	0.12132	c	0.24132	b	0.29544	a	****
		9	0.22129	c	0.29558	b	0.23595	c	0.36851	a	0.37268	a	****

*: limit of marketability. For each quality level, the mean values followed by different letters (a, b, c, d, e) are significantly different (P-value < 0.05) according to Student-Newman-Keuls (SNK) test. Significance: **** = significant at P-value ≤ 0.0001.

Results from the one-way ANOVA highlights that a separation of marketable samples (QL5, QL4 and QL3) from non-marketable ones (QL2 and QL1) is achieved by all the compared methods. This information is normally sufficient needed in most commercial applications where the QL3 represents the limit of marketability. It is interesting to note that M2, when applied in the polar representation, is able to separate all the QLs. The methods have different computational complexities and requires different times to assign a quality level to a sample: they offers a wide variety of possibilities to the designer of a Computer Vision System. The Random Forest model that has been analysed to identify the 36 most relevant colours required around 25 ms to assign a class to a sample. The methods derived by exploiting these colours reach this time only in the most complex version, allowing a significant reduction of computation time in the simpler versions. The method M1 takes only 1 ms to evaluate a sample. The method M2 requires 3 ms for assigning the class to a sample. The methods M3, in both its versions, takes 11 ms to achieve their results. M3a, applied in the polar representation, is able to separate all the QLs but not the QL5 and QL4 (corresponding to very good and good product, respectively). The method M4 involves a computation time of 22 ms.
Anyway, from QL5 to QL1, a general reduction in the C9, associated to green nuances, and an increase in the presence of yellow pigments (C5) was showed in all the 4 methods adopted. In rocket leaves, the reduction of green pigments and the simultaneous increase of yellow ones during the cold storage is due to biological degradation of chlorophyll (Cefola and Pace, 2015; Cefola et al. 2010; Watkins, 2006), as also described by Palumbo at al. (2022). Indeed, in Figure 5A, data of total chlorophyll and ammonia content reported in Palumbo et al. (2022) are presented. The total chlorophyll content of rocket leaves showed a significant reduction (42.4 %) during the storage.

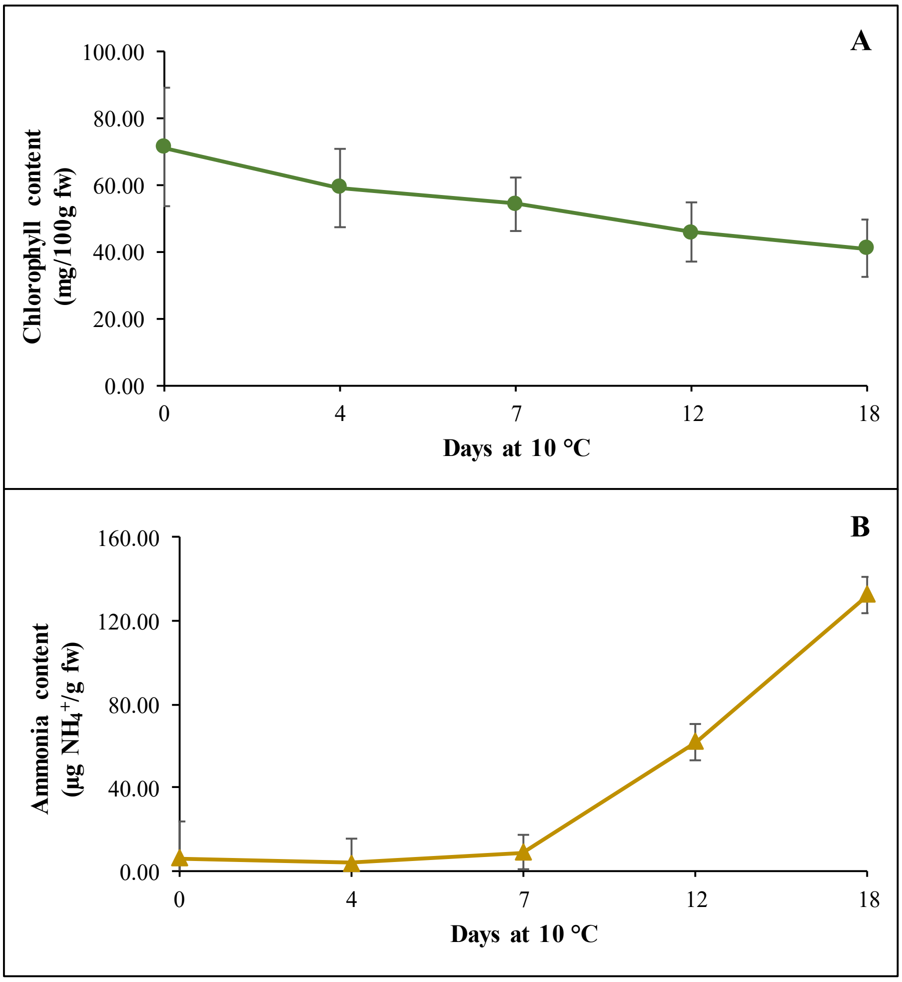


Figure 5. Changes in chlorophyll (A) and ammonia (B) contents of rocket leaves during 18 days of storage at 10 °C. Each data is the mean value of 60 samples ± standard deviation.

Additionally, postharvest chlorophyll breakdown may contribute to ammonia accumulation in vegetable tissues (Amodio et al. 2018) which is highly correlated to hue angle variations (related to leaves yellowing) in rocket leaves stored at 10 °C, already demonstrated by Palumbo et al. (2022) and Mastrandrea et al. (2016). As for ammonia content, at harvest samples showed very low values (6.11 ± 2.42 µg NH_4^+/g of fresh weight), but a significant increase was recorded at the end of storage (132.51 ± 8.67 µg NH_4^+/g of fresh weight) (Figure 5B). High levels of ammonia may cause tissue damage with visible senescence effects, influencing the overall quality of the product.

Because of their strict relation to the senescence of the product, both chlorophyll and ammonia content may be considered objective markers for quality loss of rocket leaves (Palumbo et al. 2022).

These results are clearly visible also by looking at the scoreS plot obtained by PCA analysis, that uses as variables the values of the C5 and C9 obtained by the 4 methods and the chlorophyll and ammonia data (Figure 6).

The first and the second components accounted for 66.2 % and 20.4 % of the total variance respectively, displaying a different distribution of marketable and non-marketable samples in the PCA quadrants (Figure 6A): the formers were mostly clustered at the left side, while the non-marketable ones at the right side along the first component.

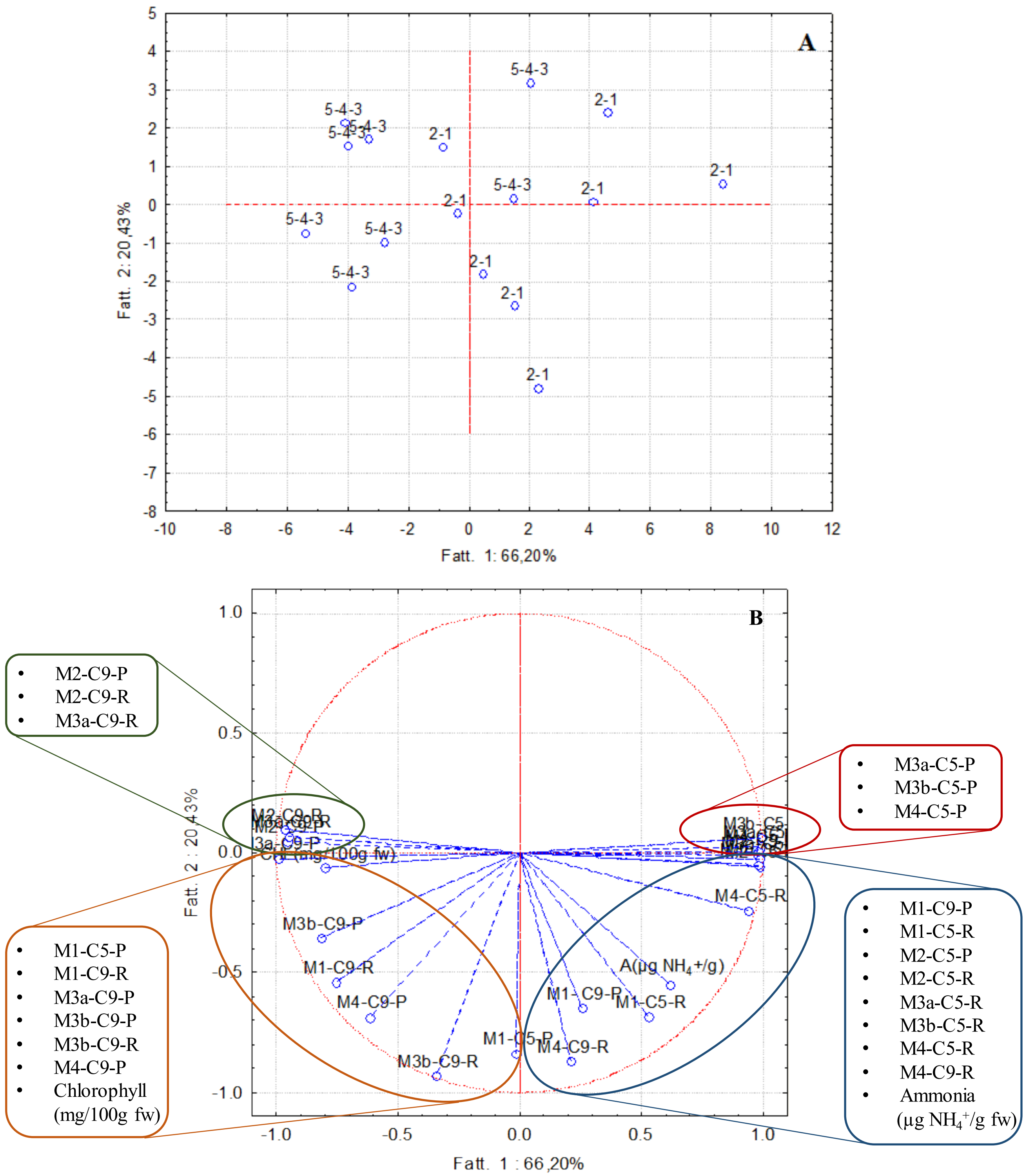


Figure 6. PCA loadingS plot (A) and scoreS plot (B) carried out on the values of clusters 5 (C5) and 9 (C9) in the polar (P) or rectangular (R) representation obtained by the 4 methods adopted (M1, M2, M3a, M3b, M4).

In the PCA scoreS plot, all the methods in the higher-left and lower-left quadrants showed a significant correlation with the chlorophyll content of rocket leaves which presented negative component 1 and 2 and was found in the lower-left quadrant of the scoreS plot; on the other

hand, the ammonia content, that is placed in the lower-right quadrant of the scoreS plot, was correlated significantly to the methods placed in the same quadrant (Figure 6B).
The relationships between clusters values of the 4 methods and the chemical attributes (chlorophyll and ammonia content) were explored by the heatmap reported in Table 2: the methods M2-C9-P, M2-C9-R and M3a-C9-R, in the higher-left quadrants, and M3a-C9-P and M3b-C9-P, in the lower-left quadrant, showed higher correlations with the chlorophyll content than the others in the same quadrants, while the ammonia content was highly correlated to the methods M1-C5-R, M4-C5-R and M4-C9-R. Even for what concern the correlation with chlorophyll content, M2 e M3 exhibit a high correlation at a very low computational cost (3 ms). The differences between C5 and C9 do not appear to be relevant. Instead, for what concern ammonia, the two clusters provide very different performances. C9 seems to poorly correlate with ammonia while C5 is much more effective in estimating this property of the product. In particular, the presence of C5 evaluated using the M1, applied to the rectangular representation, exhibits a very high correlation with ammonia content. The C5 is characterized by a strong separation from the other ones in the *ab*-plane of the CIELab colour space. The corresponding colour region defined by the M1 and the rectangular representation remains well separated and have a relevant extension in the colour plane that could explain this high correlation.

Table 2. The heatmap shows the correlations between values of clusters 5 (C5) and 9 (C9) in the polar (P) or rectangular (R) representation obtained with the 4 different methods (M1, M2, M3a, M3b, M4) and chlorophyll and ammonia contents in rocket leaves. A different colour code is used to represent the strong of correlations; **r** is the Pearson's correlation coefficient.

Method-Cluster-Representation	**Chlorophyll content (mg/100g fresh weight)**		**Ammonia content (µg NH_4^+/g fresh weight)**	
	r	*P-value*	**r**	*P-value*
M1-C5-P	0.1644	ns	0.3584	ns
M1- C9-P	-0.1460	ns	0.1905	ns
M1-C5-R	-0.4051	ns	0.9015	****
M1-C9-R	0.5509	*	-0.2107	ns
M2-C5-P	-0.7004	**	0.5708	*
M2-C9-P	0.7458	***	-0.5431	*
M2-C5-R	-0.7316	***	0.6222	**
M2-C9-R	0.7838	****	-0.5991	*
M3a-C5-P	-0.7578	***	0.5489	*
M3a-C9-P	0.7578	***	-0.5489	*
M3a-C5-R	-0.7388	***	0.6061	*
M3a-C9-R	0.7912	****	-0.5713	*
M3b-C5-P	-0.7633	***	0.5300	*
M3b-C9-P	0.8648	****	-0.3701	ns
M3b-C5-R	-0.7445	***	0.6313	**
M3b-C9-R	0.3585	ns	0.3041	ns
M4-C5-P	-0.7069	**	0.5633	*
M4-C9-P	0.4480	ns	-0.0754	ns
M4-C5-R	-0.7118	**	0.7622	***
M4-C9-R	-0.1898	ns	0.6831	**

r
-1
-0.5
0
0.5
1

Significance: ns= not significant; * significant for P ≤ 0.05; ** significant for P ≤ 0.01; *** significant for P ≤ 0.001; **** significant for P ≤ 0.0001.

3.2. Chlorophyll and ammonia content prediction

The methods that reported the highest correlations with chlorophyll and ammonia contents were used to build two PLS models to predict these two quality markers of rocket leaves (Table 3). Results showed good prediction of chlorophyll (Model 1) and ammonia (Model 2) content by using as predictors the values of the methods M3b-C9-P and M1-C5-R, respectively. In detail, the Model 1 with R^2 of 74 % in calibration and 70 % in validation was obtained for chlorophyll content. Higher performances were obtained with the Model 2 to predict the ammonia content (R^2 of 0.83 and 0.72 in calibration and validation, respectively). Similar performances were achieved by the prediction reported in Palumbo et al. (2022), in which random forest model was used. In this work, M3b-C9-P provides lower performances for chlorophyll prediction than Palumbo et al. (2022), but the methodology adopted provided simpler algorithms, easily interpretable by humans, and a lower computational speed (about 3 ms against the more than 20 ms of the random forest model). Additionally, while no relevant correlation was identified in Palumbo et al. (2022) for ammonia content, often used as another senescence indicator in leafy vegetables, the novel approach allowed to obtain a significant prediction of this parameter by M1-C5-R, the simplest model that has a computational time of 1 ms.

Table 3. Root Mean Square Error (RMSE) and the coefficient of determination (R^2) in calibration ($_c$) or validation ($_v$) of the partial least square regression (PLSR) models predicting chlorophyll and ammonia contents of rocket leaves.

PLSR models	Predictors	$RMSE_c$	R^2	$RMSE_v$	R^2
Model 1 (chlorophyll content prediction)	M3b-C9-P	6.23	0.74	7.04	0.70
Model 2 (ammonia content prediction)	M1-C5-R	20.27	0.83	27.58	0.72

4. CONCLUSIONS

The present research paper explores the possibility of using the information hidden into machine learning models developed to classify visual quality and to estimate internal properties of rocket leaves to develop methods that have lower computational costs but also that are more understandable by humans. In particular, a Random Forest model already used in previous experiments to classify visual quality and to estimate chlorophyll and ammonia contents in rocket leaves has been analysed. New methodologies are proposed to identify specific relevant and representative colours in the *ab*-plane of the CIELab colour space for the evaluation of significant traits of the product at hand and to identify colours whose changes induced by senescence are strictly related to chemical and physical properties. This set of relevant colours has been used to construct several methods, with increasing levels of computational complexity, for accomplishing the same tasks done by the Random Forest model.

Simple algorithms proved able to (i) identify relevant clusters of colours that are informative about the properties of the product at hand, (ii) further to select the clusters more significant to estimate the desired properties, (iii) to describe shape and size of regions of the *ab*-plane in the CIELab colour representation corresponding to the clusters of interest.

These results provided objective and sound bases for the design of different computational schemes with different execution times enabling the best trade-off between efficacy and efficiency, depending on the application constraints. In particular, two of the considered methods, M3b-C9-P and M1-C5-R, provided good prediction of chlorophyll and ammonia contents, that are able to assess the state of product in an objective and robust way. The two methods have computational time of 3 ms (M3b-C9-P) and of 1 ms (M1-C5-R) that favourably compare with the computational time of the Random Forest model (not less than 20 ms). The easily readable results of the experiments provided suggestion about their interpretation in terms of known processes occurring during the senescence of the product. Moreover, the identification of well-grounded objective colour cues could also be used to improve the significance of indications provided to human operators during their training on the quality evaluation task.

REFERENCES

Amodio, M.L., Colelli, G., Cantwell, M.I., 2018. Ammonia accumulation in plant tissues: a potentially useful indicator of postharvest physiological stress. Acta Hortic. 1194, 1511–1518. https://doi.org/10.17660/actahortic.2018.1194.211.

Cefola, M., Pace, B., 2015. Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. J. Food Process. 39(6), 2523–2532. https://doi.org/10.1111/jfpp.12502.

Cefola, M., Amodio, M.L., Rinaldi, R., Vanadia, S., Colelli, G., 2010. Exposure to 1-methylcyclopropene (1-MCP) delays the effects of ethylene on fresh-cut broccoli raab (*Brassica rapa* L.). Postharvest Biol. Technol. 58 (1), 29–35. https://doi.org/ 10.1016/j.postharvbio.2010.05.001.

Chaudhry, M.M.A., Babellahi, F., Amodio, M.L., Colelli, G., Sahar, A., 2020. Image analysis. In: Khan, M.K.I. (Ed.), Advances in Noninvasive Food Analysis (pp. 200-201). CRC Press, New York.

Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C., Huang, W., 2020. On line detection of defective apples using computer vision system combined with deep learning methods. J. Food Eng. 286, 110102. https://doi.org/10.1016/j.jfoodeng.2020.110102

Hastie, T., Tibshirani, R., Friedman, J., 2009. Hierarchical clustering. In: Hastie, T., Tibshirani, R., Friedman, J. (Eds), The Elements of Statistical Learning (2nd ed.) (pp. 520–8). Springer, New York. https://doi.org/10.1007/978-0-387-21606-5

Li, B., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees (CART). Biometrics, 40(3), 358-361.

Loh, W.Y., Shih, Y.S., 1997. Split Selection Methods for Classification Trees. Stat. Sin. 7, 815–840.

Loh, W.Y., 2002. Regression Trees with Unbiased Variable Selection and Interaction Detection. Stat. Sin. 12, 361–386.

Mastrandrea, L., Amodio, M.L., Cantwell, M.I., 2016. Modeling ammonia accumulation and colour changes of arugula (Diplotaxis tenuifolia) leaves in relation to temperature, storage time and cultivar. Acta Hortic. 1141, 275–282. https://doi.org/ 10.17660/ActaHortic.2016.1141.34.

Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Colelli, G., Attolico, G., 2022. Non-destructive and contactless estimation of chlorophyll and ammonia contents in packaged fresh-cut rocket leaves by a Computer Vision System. Postharvest Biol. Technol. 189, 111910. https://doi.org/10.1016/j.postharvbio.2022.111910

Palumbo, M., Pace, B., Cefola, M., Montesano, F. F., Serio, F., Colelli, G., Attolico, G., 2021. Self-configuring CVS to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. Agronomy, 11(7), 1353.

Watkins, C.B., 2006. The use of 1-methylcyclopropene (1-MCP) on fruits and vegetables. Biotechnol. Adv. 24 (4), 389–409. https://doi.org/10.1016/j. biotechadv.2006.01.005.

PART III: CONCLUSIONS

FINAL CONSIDERATIONS

The production of high quality foods is considered a key factor in horticultural sector even though the concept of quality has evolved significantly over the past few decades. Despite the traditional sensory attributes represent the main focus of most quality standards and regulations, more recently other important aspects are becoming increasingly relevant: high demands and attention of modern consumers towards nutritional value of fresh horticultural products and their growing sensibility toward the sustainability of production processes. Recent researches are working on rapid, precise and low-cost techniques for food analysis along the entire supply chain. Online sorting systems along the industrial lines may allow the inspection of large quantities of fruit and vegetables in a short time, may provide a good prediction of the external and internal quality traits of products, and may monitor the physiological postharvest state, thus reducing losses and wastes, with evident social advantages. Among the innovative techniques qualified for this task, image analysis by conventional computer vision systems (CVS), based on RGB imaging, represents a contactless and non-destructive tool to evaluate and monitor the quality of fruit and vegetables. They are non-expensive, fast, and effectively and extensively usable along the whole supply chain, even on packaged products which cannot be analysed using conventional tools.

The research activity carried out during this doctoral program proposed a CVS for a continuous monitoring of the freshness level and the quality of fresh rocket leaves from harvest to final consumers even when enclosed in plastic packaging. This technology worked in an objective and consistent way basing on non-destructive measurement of biological markers (such as chlorophyll and ammonia) which are strongly related to leaf senescence. The proposed CVS was able to automatically select, without human intervention, the most relevant colour traits using the Random Forest as machine learning model.

In the first study, the CVS achieved a high accuracy in the quality level assessment of rocket leaves during the cold storage. Moreover, the same system was used to identify traits related to the sustainability of the cultivation approach, discriminating among different water and fertilization management protocols. Even if performances were lower for this second objective and less relevant for practical applications, results were fully in accordance with the outcomes obtained by the conventional destructive methods, according to which the different cultivation approaches did not significantly affect the quality of the product. For this last task, further investigations are needed.

In the second study, the developed CVS was able to work on unpackaged and packaged rocket leaves without significant differences, properly combining the image processing techniques and the random forest model. In detail, it solved both the classification (i.e. assessment of the visual quality level) and the regression (i.e. estimation of the chlorophyll and ammonia contents) problems using the same methodology on the same training data. As results, experiments proved that the performance loss due to the presence of the packaging material was irrelevant in the chlorophyll and ammonia prediction by CVS. Additionally, the PLS models, built to predict the quality level using total chlorophyll and ammonia content as predictors, further confirmed the ability of the CVS to operate also through the packaging.

Finally, in the third study, new methodologies are proposed to identify specific relevant and representative colours in the *ab*-plane of the CIELab colour space for the evaluation of significant traits of the fresh rocket leaves and to identify colours whose changes induced by senescence are strictly related to chemical and physical properties. The results provided objective and good bases for the design of different computational schemes with shorter execution times, enabling the best trade-off between efficacy and efficiency and overcoming the performances of Random Forest models.

Conclusively, results reported in this doctoral Thesis may have a significant impact on advanced applications of the traditional vision systems, commonly used for the inspection of fruit and vegetables. As practical examples, portable diagnostic systems can be used directly on field, for a complete and non-invasive analysis of the physiological state of the crop or for the identification of the correct maturity stage at harvest; moreover, integrated systems can be installed inside supermarkets and household refrigerators to reduce food losses and preserve the safety of consumers; smartphone-based tools can help customers in verifying the quality of food at the time of purchase.

There are still many open challenges to perform the quality assessment in a faster and accurate way in industrial lines and during postharvest phase such as (i) the distribution of light on curve surfaces, (ii) the need for whole surface inspection, (iii) the need for new simpler statistical methodologies for algorithms construction, and, (iv) the requirements about the environment at acquisition time. Indeed, current technologies are adapted to industrial lines where the environment is controlled, but their transfer to less structured environments (logistic, retail facilities, households) might still require significative work in custom adaptation. Advanced solutions in which smarter algorithms may face the variations induced by environment conditions could simplify a widespread use of these technologies whenever quality control is required.

As for plastic packaged products, the identification of image regions affected by the interaction between light and the plastic material is a crucial issue for quality level assessment through the packaging. A valid and consistent segmentation approach which selects areas where colours can be measured correctly is required in order to obtain performances similar to those achieved on unpackaged samples.

Other future research perspectives in the field of non-destructive analysis by CVS should be focused on simplifying the existing models to support the industry for online implementation of sorting systems as it has been proposed in Chapter 5.

ACKNOWLEDGEMENTS

I am very grateful to everyone who gave me the strength and the serenity to complete my research work and who made this thesis possible.

Firstly, I offer my sincere thanks to my Tutor ***Prof. Giancarlo Colelli,*** *University of Foggia, for his constructive hints and support to complete my doctoral studies. All the Joint Meetings of Agriculture-oriented PhD Programs in which I was involved gave me the opportunity to challenge myself and to meet new friends with the same interest in the scientific research.*

My warmest thanks go to my Co-Tutor ***Dr. Maria Cefola****, Institute of Sciences of Food Production of National Research Council (ISPA-CNR), who was determinant in creating a comfortable and pleasant atmosphere during these years and under whose friendly supervision, clever guidance, constructive criticism and edifying discussions, I have been able to complete this challenging "journey" with very little difficulty. Sharing her knowledge, she allowed me to enhance my interest and passion for scientific research.*

Special thanks to ***Dr. Giovanni Attolico****, Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA), for helping me in the image analysis by CVS.*

I wish to extend my sincere thanks and appreciation to ***Dr. Bernardo Pace,*** *Institute of Sciences of Food Production of National Research Council (ISPA-CNR), for his kindly support and constant drive throughout my studies and laboratory works.*

Thanks to ***Dr. Francesco Serio****, Institute of Sciences of Food Production of National Research Council (ISPA-CNR), and* ***Dr. Francesco Fabiano Montesano****, University of Bari Aldo Moro, who provided me all the samples of rocket leaves every month for the experimental applications of this thesis work.*

I cannot forget my sincere colleagues and friends ***Drs. Imperatrice Capotorto****,* ***Antonia Corvino,*** *and* ***Ilde Ricci,*** *who helped me since the very first day of my PhD program.*

I am most grateful to all my colleagues of the ISPA-CNR, ***Drs. Sergio Pelosi****,* ***Vittorio Capozzi*** *and* ***Massimo Franchi****, for their friendly support, cooperation, encouragement, and love during all my study period, whenever I needed it.*

Finally, to my husband ***Gaetano****, to my son* ***Emanuele****, and my daughters* ***Maria Chiara*** *and* ***Monica****, to* ***my parents****, my* ***brother*** *and my* ***sister****, and to all the* ***Progetto Nazareth's fraternity*** *I want to say thank you very much for your love, care and encouragement in any situation in these years of my life. This achievement would not have been possible without you.*