Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material

Dario Pietro Cavallo[1a], Maria Cefola[1bc*], Bernardo Pace[1bc], Antonio Francesco Logrieco[bc], Giovanni Attolico[a]

[a] Institute on Intelligent Systems for Automation, CNR-National Research Council of Italy Via G. Amendola, 122/O – 70126 Bari, Italy.

[b]Institute of Sciences of Food Production, CNR-National Research Council of Italy Via G. Amendola, 122/O – 70126 Bari, Italy.

[c] Institute of Sciences of Food Production, CNR-National Research Council of Italy, URT c/o CS-DAT, Traversa Viale Fortore, 71121 Foggia, Italy.

[*] Corresponding Author: phone/fax: +39.080.5929304/9374; email address: maria.cefola@ispa.cnr.it

[1] **First Authorship is equally shared**

Abstract

Non-destructive evaluation of vegetables by Computer Vision Systems (CVSs) makes possible to check their quality level in an objective and consistent way along the whole supply chain up to the final users. CVSs have been proven to be successful when applied to unpackaged products. The proposed approach aimed to enable this analysis on packaged fresh-cut lettuce with minimum constraints on the acquisition phase and without any care to flatten the surface of the bag facing the camera. A deep-learning architecture, based on Convolutional Neural Networks (CNNs), was used to identify regions of the image where the vegetable was visible with minimum colour distortions due to packaging. To meaningfully assess the performance of the system, each lettuce's sample was

acquired both through packaging material and without packaging material. The image analysis was applied to both the resulting images to automatically grade their quality level. The results showed that the performance loss due to the presence of packaging is negligible (83% instead of 86%) and that the proposed system can be used to monitor the quality level of fresh-cut lettuce regardless of packaging at all the critical check points along the supply chain.

Keywords: non-destructive quality evaluation, automatic visual grading through packaging, deep learning, Convolutional Neural Network

1. Introduction

Fresh-cut products are particularly appreciated by consumers due to their nutritive properties, convenience and ease of use. During the post-packaging phase and along the distribution chain, the global quality of these products is generally determined by subjective sensory evaluations. To achieve optimal management of the packaged fresh-cut products and to inform consumers about the real quality level of each single bag available, industries and logistic operators need objective, non-destructive, simple, rapid and contactless methods to assess the quality level of fresh-cut packaged products without opening the bags.

Image analysis by Computer Vision Systems (CVSs) do not require sample preparation and represent a practical alternative to time-consuming analytical (chemical and physical) methods on unpackaged products. Recently, CVSs were used to assess quality and marketability of artichokes (Amodio, Cabezas-Serrano, Peri, & Colelli, 2011), fresh-cut nectarines (Pace, Cefola, Renna, & Attolico, 2011), fresh-cut lettuce (Pace et al., 2014), fresh-cut radicchio (Pace et al., 2015) and rocket leaves (Cavallo et al. 2017) in a completely contactless way. It is very important to enable the analysis of samples through the packaging that, generally, has transparent parts which size and shape depend on the specific brand. Images of packaged products are normally composed by three kinds of regions: opaque areas where the product is hidden by graphical elements where their colours and shapes strongly vary on the basis of brands; transparent areas where the visual appearance is heavily affected by reflections or other artefacts induced by the interaction between the packaging and the light (affected transparent areas); transparent areas where the product is observable with acceptable fidelity (unaffected transparent areas). Opaque and affected transparent areas need to be discarded and the image analysis must be restricted to regions of unaffected transparent areas. The separation of these regions is a critical step requiring robust and powerful segmentation approaches. To the best of our knowledge, there are no applications of CVS to estimate the quality level of a vegetable product through packaging. Multi-spectral reflective image analysis was applied to monitor the evolution and spoilage of leafy spinach covered by plastic materials (Lara et al., 2013). In detail, the leaves were placed inside plastic Petri dishes which were covered with three different plastic materials to compare the optical behaviour of the most commonly used packaging. It is important to note that, in such a situation, the plastic formed a regular planar surface: its interaction with the light (traversing the plastic both when going from the lamp to the product and from the product to the sensor) mainly reduced the measured spectra, affecting each wavelength on the base of the optical properties of the material. No reflection of other kinds of distortion were present, due to the controlled relative geometry among Petri dish, lamp and detector. Contrastingly, the bags used for fresh-cut products cannot be managed to offer such a favourable condition during the acquisition phase. In addition, Giovenzana et al. (2016) applied Vis/NIR and NIR spectroscopy to measure the quality of fresh-cut leaves of *Valerianella* through packaging, throughout the cold storage. In this case, the packaging was used to envelop the probe (a sensing optical-fiber with a ring of lighting optical fibers). The relative geometry among light, plastic surface and sensor was set and induced only a fixed attenuation on the hyperspectral measures (whose amount at each wavelength was dependent on the optical properties of the plastic material). In that paper, all the hyperspectral data were acquired by placing the probe in contact with the leaves.

This paper describes a non-destructive and contactless application of image analysis by CVS for the evaluation of the quality level of packaged fresh-cut lettuce. The final goal was to show that a careful selection of the area to be processed enables the achievement of similar performances on packaged and unpackaged products. The proposed approach is based on the acquisition of calibrated colour images without any contact with the product and without heavy constraints on the positioning of the bag: the method can therefore be used in a large number of points along the supply chain from the manufacturing process up to the point of sale and the consumer.

2. Materials and methods

2.1. Experimental setup and quality level evaluation

2.1.1. Plant material and processing

Fresh-cut iceberg lettuce (*Lactuca sativa* L.) was provided by a farm (Ortomad srl) located in Pontecagnano (southern Italy) and transported in cold conditions to the Postharvest Laboratory of the Institute of Sciences of Food Production. Fresh-cut leaves (already washed, sanitized and dried by the company) were placed in open polypropylene bags (25 × 30 cm, 30 μm, Carton Pack, Rutigliano, Italy) containing about 150 g of product each and stored at two different temperatures (8 °C and 15 °C). Eighty bags were prepared for each storage temperature (16 replicates X 5 quality levels). All items, at any time during storage, were graded according to a five-levels quality scale based on sensory evaluation, as reported below. For each bag, two images were acquired: the first was of the entire closed bag with the fresh-cut lettuce inside; the second was of the product alone, after removing the bag (Figure 1). Images were acquired by a Computer Vision System (CVS) using a 3CCD (Charged Coupled Device) digital camera (JAI CV-M9GE) with a dedicated CCD for each color channel to avoid the artefact introduced by interpolation over the Bayern pattern typical of single CCD cameras. The images were saved using the uncompressed TIFF format to prevent any colour deformation due to compression algorithms. The optical axis of the Linos MeVis

12 mm lens system was perpendicular to the black background. Eight halogen lamps (divided along two rows placed at the two sides of the imaged area) were oriented at a 45° angle with respect to the optical axis. A small colour-chart (X-Rite Colour Control Patches) was placed in the scene to enable colour correction. The samples were placed on the background without any care to flatten the surface of the bag facing the camera: which had created wrinkles that generated reflections and glares in the images. The robust segmentation step removed all these artefacts to make meaningful the following image analysis. Moreover, a chemical analysis evaluated the ammonium content of all the samples, as reported below.

2.1.2. Quality level classification and ammonium content

During the storage, fresh-cut iceberg lettuces were evaluated and classified using 5 quality levels (QL) according to the scale reported by Pace et al. (2014): 5 = very good (very fresh, no signs of wilting, decay or bruises), 4 = good (slight signs of shrivelling, bruises), 3 = limit of acceptability or marketability (moderate signs of shrivelling, browning, dryness, wilting, bruises), 2 = poor (severe bruises, evident signs of shrivelling, pitting, decay) and 1 = very poor (unacceptable quality due to decay, bruises, leaky juice). The QL 3 was considered the minimum level acceptable for the market (Nunes et al., 2009); therefore, samples with quality lower than 3 were considered waste product.

For ammonium measurement, the method reported by Weatherburn (1967) was used. In detail, 5 g of chopped sample was homogenised (Ultraturrax T-25, IKA Staufen Germany) with 20 mL distilled water for 2 min, centrifuged at 12,000 rpm × 5 min and a 0.5 ml extract was used for the analysis. Colour development, after the reaction with a phenol nitroprusside reagent and alkaline hypochlorite solution, was determined after incubation at 37 °C for 20 min, reading the absorbance at 635 nm (UV-1800, Shimadzu, Kyoto, Japan).

The effects of QL on ammonium content were tested by performing a one-way ANOVA with data means arranged in a completely randomized design. The mean values for QL were separated using the Student–Newman–Keuls (SNK) test.

2.2. Workflow of the non-destructive approach to assess the quality level of fresh-cut iceberg lettuce through packaging material

Separate images of each sample (n=160, eighty for each temperature) were acquired with the product inside and outside the packaging (Figure 1) to evaluate the effects of packaging on the performance of the method. Specifically, 5 acquisition days were chosen for each temperature, corresponding to the different quality levels. Therefore, two different datasets of images were obtained: one associated to the 160 samples acquired with packaging and the other with the same samples acquired without packaging. To make meaningful the comparison, the images of both the datasets went through mostly the same workflow shown in Figure 2. The only difference was the image segmentation step: the relevant part of packaged fresh-cut lettuce (visible and not affected by artefacts produced by the bag) was selected using a Convolutional Neural Network (CNN) trained to separate 3 different classes (package, product, artefact). All the pixels labelled as package or artefact were removed. The pixels associated to the product were further processed following the workflow steps reported in Figure 2 and detailed below.

2.2.1. Pre-processing

The pre-processing of each image consists of three main steps, as previously described in Pace et al. (2015 and 2017): (*i*) the colour-chart was automatically found and removed from the scene, (*ii*) the colours of the image were corrected using the information extracted from the colour-chart and (*iii*) the foreground was separated from the background.

2.2.2. Application of Convolutional Neural Network on packaged fresh-cut iceberg lettuce

The deep-learning approach based on a CNN was applied only to segment the images of packaged product, identifying and selecting only pixels belonging to the fresh-cut lettuce without artefacts (such as glares, reflections) that make unreliable the visual appearance of the product. Our hypothesis was that those pixels could enable the evaluation of quality level even if acquired through the packaging: to prove this hypothesis the proposed approach needed to achieve similar performances on both packaged and unpackaged samples by applying exactly the same processing. Therefore, the processing modules (apart from segmentation) were kept exactly the same (Fig. 2). The architecture used for segmentation should deal with commercially available bags without human intervention and without changing its structure. It should be able to select features suitable to identify graphical elements that are significantly different among different brands. It should also recognize regions where the visual appearance is distorted by artefacts: this makes the system insensitive to the effects of the bag being imaged without any care taken to flatten its surface. Moreover, the human intervention is limited to provide the ground-truth needed for the learning phase.

The CNN was trained and tuned on 40 images (selected by choosing eight samples for each quality level) and applied to segment the remaining 120 samples and to select in each image the part of product that was visible without any artefact. Those regions were used for the quality level classification phase. The complete separation, between samples used for training and tuning the segmentation, and samples used for classification, is required to provide robust and meaningful results.

2.2.2.1. *Patch-based CNN learning*

Three different classes (class 0: package; class 1: product; class 2: artefact) were learned by CNN, since artefacts and packages can exhibit quite different visual appearance even if they both need to be discarded. Three ground-truths were manually produced from each image of the training dataset by separating these three classes. Regions of transition between adjacent classes were ignored to improve the robustness of the process. A sliding window algorithm was developed to extract patches of size 3 by 3 pixels (3 pixels of height by 3 pixels of width). Each patch was classified using the ground-truth region it belonged to. A specific and efficient sliding window approach was developed to reduce the time complexity of managing this large number of patches during the training sessions: it combined numerical vectorization and a low-cost storage strategy (in terms of both number of read/write operations and memory occupation). The approach is detailed in Appendix A.

The following architecture of the CNN was used:

- Input layer: take as input n 3d-tensors (weight x height x channels) of 3x3x3 size (n is the number of the patches);
- Convolutional layer: 64 2x2 filters with rectify non-linear activation function (Nair et al., 2010) and Glorot uniform distribution to weight initialization (Sainath et al., 2015);
- Max-pooling layer with 2x2 filter;
- Dense layer (drop-out 0.5) with 256 units and rectify non-linear activation function (Nair et al., 2010; Srivastava et al., 2014);
- Dense layer (drop-out 0.5) with 3 units (number of classes to predict) and softmax non-linear activation function (Renals et al., 2014).

Parameters were optimized using a stochastic mini-batch gradient descent with Nesterov's moment (Bengio, 2012; LeCun et al., 2012).

2.2.2.2. Image segmentation by Convolutional Neural Network (CNN)

After the training, when used to segment an image, the CNN consider each 3 by 3 window of the image and returns three probabilities associated to the three different classes: these probabilities are assigned to all the 9 pixels of the window. Because each pixel can be part of a window up to 9 times (in different positions), all the received probabilities are cumulated and, at the end, normalized by the number of predictions made on that position. The result is a map in which, at each pixel, three cumulated and normalized probabilities are available, one for each class. For some pixels, one of the probabilities is much greater than the other (assessing a strong confidence that those pixels belong to a specific class) while in other pixels two or more probabilities are similar (ambiguous situations in which classification can be arbitrary). Several strategies can be applied to identify the regions of product that must be considered for further processing on the basis of this map. A conservative approach was adopted to reduce the risk of including pixels belonging to packaging or to artefacts with unreliable colours in the following processing . A very high threshold (equal to 1) has been applied and only the pixels with probability of belonging to products greater than or equal to this threshold were selected for the following processing.

2.2.3. Quality level classification comparison by k-Nearest Neighbors

The quality level classification part of the approach was developed and tuned on the 120 samples that were not involved in the training of CNN. During the classification phase, the corresponding images, acquired with and without the package on the same samples, were considered and the results were compared. The feature vector for each sample was built using this approach: (*i*) the colours of relevant pixels (all the foreground pixels for unpackaged product and the pixels selected by segmentation for the packaged ones) were converted in CIE Lab colour space; (*ii*) the L channel was discarded due to its greater sensitivity to the uncontrollable geometry between surface of lettuce, lights and camera; (*iii*) a 2d histogram of the other components (a and b) was built and used as a feature vector. To increase the efficiency of the algorithm, these two-dimensional histograms were reshaped as unidimensional vectors. A 3-Nearest Neighbours approach (Cover et al., 1967) was then used to predict the quality level experimentally determined by the chemical analysis of ammonium. Its performance was estimated using a Leave-1-Out Validation method (Kohavi, 1995). Classification accuracy was used to evaluate the predictive model. These accuracies (related to packaged and unpackaged products respectively) were compared.

3. Results and Discussion

3.1. Assessment of quality level and ammonium content of fresh-cut iceberg lettuce

Fresh-cut lettuce samples exhibited a decrease in the overall quality during storage, that was faster for samples stored at higher temperature (15 °C). In fact, samples stored at 8°C reached the QL 3 (limit of acceptability) after around 3 days, one day later than fresh-cut lettuce stored at 15 °C. The latter resulted poor (QL = 2) after roughly 3 days and unacceptable after about 4 days. On the other hand, samples stored at 8 °C reached the QL = 1 after about 7 days. The main factor which affected the loss of visual acceptability was the browning of fresh-cut tissues (Ares et al., 2008; Pace et al., 2014).

Ammonium content increased in fresh-cut samples from QL5 to QL1. Moreover, as previously reported by Pace et al. (2014) on fresh-cut lettuce, it proved to be able to discriminate the acceptable product (ranging from QL = 5 to QL = 3) from the waste (QL = 2 or 1). Even the two classes of waste were well discriminated by ammonium content (Figure 3). Ammonium is considered a reliable indicator of product freshness, since it was reported to accumulates while leaves become senescent (Cefola, Amodio, Rinaldi, Vanadia, & Colelli, 2010; Chandra,Matsui, Suzuki, & Kosugi, 2006).

3.2. Performance of CNN approach to predict the quality level of fresh-cut iceberg lettuce through package

The ground-truths of the 40 samples chosen for training and tuning the CNN were processed by the sliding window algorithm (described in Appendix A): this algorithm produced a dataset of 3 by 3 patches (each with a proper classification) that were used for tuning the segmentation phase. This dataset included: (*i*) 5.861.632 patches of class 0 (package), (*ii*) 3.554.759 patches of class 1 (product) and (*iii*) 359.781 patches of class 2 (artefacts). About 40Mb was required to store the indexes files representing 9.776.172 patches and 446 seconds were spent only once to identify all the patches and to store them. To balance the relevance of the three classes during the training phase, subsets with the same cardinality of class 2 were randomly extracted for class 0 (package) and class 1 (product).

This reduced and balanced dataset was further divided into training set (80% of the patches) and validation set (the remaining 20% of the patches). The training set was used to optimize the parameters of the CNN by stochastic mini-batch gradient descent while the validation set was used to evaluate the CNN. Specifically, after 100 training epochs, a classification accuracy of 0.979 was reached (it measured the number of correct predictions over the number of attempted predictions on validation data). Efficient Deep Learning libraries (Lasagne and Theano) were used that enable the acceleration by GPU: convergence was achieved in only 70 seconds.

The images corresponding to the remaining 120 packaged samples were segmented using the tuned CNN to classify each pixel of the foreground: the normalized probabilities were used to select the pixels belonging to the product. To be conservative, only the pixels with probability equal to 1 (product) were moved to the following step.

Figure 4 shows a compact representation of the output of the segmentation process: each band of the colour image encodes the probability that the corresponding pixel belongs to a specific class (the red band is associated to package, the green one to product and the blue one to artefacts). Crisp hues indicate unambiguous classification to a single class while less definite colour or shades describes uncertain situations that could produce arbitrary classification. Therefore, pixels with a strongly definite hue are associated to a single class (with maximum probability) while less definite colour and/or shades represent pixels where the classification is uncertain.

The pixels corresponding to product acquired in favourable conditions were used (as described in Section 2.2.2.2) to build a kNN (k = 3) classifier that identifies, for each sample, the 3 most similar samples in terms of (a,b) colour vector representations: by applying a voting algorithm to the quality levels of these 3 closer samples, the algorithm assigns a quality level to the sample at hand. Similarly, the images corresponding to the same 120 samples acquired without packaging were classified using the same approach: in this case, the two-dimensional histogram of the *a* and *b* components was built using all the foreground pixels (there was no need to discard graphical regions or area with unreliable colours). This complete symmetry between the processing of the two types of images enables a meaningful evaluation of the effects of packaging on the classification performance. A Leave-1-out validation method was used to fully exploit the limited number of samples available. The classification accuracy was 83% and 86% on images of packaged or unpackaged samples respectively. This minor performance loss, showed that a robust and effective segmentation phase can extend the quality level estimation by CVS also to products observed through their bag.

CNNs, while similar to Artificial Neural Networks in their basic elements, have a different architecture that generates significant advantages. ANNs take an input vector and process it through hidden layers (usually just one). In ANN each layer is a set of neurons fully-connected to neurons in the previous and following layers and each connection is associated to a learnable weight. A CNNs is composed by a long sequence of sparsely connected layers (mainly convolutional or pooling) having different functions. The lower levels evaluate multi-resolution features that are used to accomplish the classification task: it is important to note that these features are autonomously learned without human intervention as in the ANN. Normally, in a deep architecture, only the last layers are fully connected and implement the classification phase. These architectures are designed to naturally process multidimensional data (such as colour images) and to encode useful properties through efficient implementations (LeCun et al., 2015). Moreover, the neural architecture proposed, set all the weights without any specific tuning by humans. This aspect represents an advantage in comparison with other neural networks' algorithms (Zhang et al.,2014; Morais de Oliveira et al., 2016; Rong et al., 2017) that require a heavier configuration and tuning (features identification and selection, image processing algorithms, architecture of the neural network, …) instead of automatically learning from proper samples during a training phase.

3.3. Application on a commercial fresh-cut package

A preliminary check has been made about the generalisability of the proposed method: images of 11 commercial packaged samples of ready to eat corn salad leaves (*Valerianella locusta* L.) were acquired (Figure 5). This dataset posed several challenges: (*i*) the package represented a real industrial packaging, with several different colours (even similar to the ones of the product), graphical elements and banners; (*ii*) the package contained leaves of *Valerianella* (thus, a different product with respect to iceberg lettuce). Images were then used to train the CNN and the remaining one was used to evaluate the segmentation results. As shown in Figure 5, the proposed architecture has been able to identify the parts related to the packaging (in red), the artefacts (in blue) and the product (in green). It is important to note that the approach can recognize all the graphical elements (in spite of their colours and shape) and also the internal part of the bag without products (in the bottom) (Fig. 5). This application, on a commercial packaged salad, shows that the CNN is flexible and powerful enough to deal with such a challenging situation without any change in the architecture. Thus, the segmentation method can be applied to different type of bags by only changing the training data without modifying the architecture.

4. Conclusion

A non-destructive approach for the quality level evaluation of packaged fresh-cut products was proposed and experimentally validated.

The paper addressed two main questions. The first one was identification of the region of the bag where the product is visible without artefacts introduced by the interaction between packaging and light. A deep-learning approach, based on a CNN, was used to accomplish this task. The proposed architecture did not need human intervention to select features describing colour and texture or to fine tune the parameters of the network. They are automatically derived and have proven to be successful also on commercially available bags that exhibit large variation in visual appearance.

The CNN has been able to select the region suitable to be processed for classifying the quality level: its robustness also deals with artefacts generated by the three-dimensional complexity of the surface of the bag. The second question was to evaluate the effects of packaging on the assessment of quality level of fresh-cut lettuce. The classification was made using a 3-Nearest Neighbours classifier applied on a feature vector composed by the two-dimensional colour histogram on the *a* and *b* components in the CIELab space. The performance of the classification has been assessed using Leave-1-Out validation method that fully exploit the limited number of available samples. To make meaningful the comparison, the same workflow (except for segmentation by CNN) has been applied to the different image datasets obtained by imaging the same samples with and without packaging. The performance comparison showed a minor loss due to the effects of bags (83% classification accuracy instead of 86%). Therefore, a reliable quality evaluation can be done even through packaging provided that a robust segmentation of the meaningful parts of the images has been previously done. The architecture has been tested on commercial bags and provided good results. Therefore, the proposed approach can non-destructively and contactless assess the quality level of products through the packaging and be used along the whole supply chain from the manufacturer up to the final user to check continuously the quality of fresh-cut vegetables.

Acknowledgements

The research leading to these results has received funding from Puglia regional call "Aiuti a sostegno dei Cluster tecnologici regionali per l'innovazione". Continnova Project: Container innovativo isotermico intermodale equipaggiato con atmosfera controllata per il trasporto di prodotti ortofrutticoli freschi" (cod.VFQA3D0). The authors thank Arturo Argentieri and Michele Attolico for the technical support to the configuration of the experimental set-up.

Appendix A

Choosing the proper architecture and training the CNN requires to run several sessions each of which requires all the patches belonging to the training set to be loaded. The training set is composed of 3 by 3 patches extracted, for each class (product, packaging, artefact), from the corresponding mask. This selection of elements of the training can be done once and the result can be stored for any further need. To save storage space, it would be possible to store only the original image and the mask and to select the patches at every run: this approach saves space but can slow down the loading phase because the patches need to be selected again at every run. Contrastingly, it could be possible to extract all the patches once and to store them separately from the original image. This solution prevents the repetition of the selection phase but requires more storage and the loading phase becomes heavier due to the high level of redundancy between patches. It can therefore be useful to organize the data to optimize both the storage space and the time required to load the data. A specific approach has been designed and developed to achieve this result.

Specifically, using vectorization a very-fast indexing of patch coordinates was quickly obtained as follows:

- Let *row_inc* and *col_inc* the chosen pixel increasing steps for the sliding window along the the rows and the column respectively (1);
- Let *idx_rows* to contain all the possible indexes ranging from 1 to the number of row in the image by steps of *row_inc* and *idx_cols* to contain all the possible indexes ranging from 1 to the number of columns by steps of *col_inc*;
- Let $P = (x_i, y_i)$ be the set of all the possible top-left corner possible patches computed by a meshgrid of *idx_rows* and *idx_cols*;
- Let fg, bw_0, bw_1 and bw_2 be the binary masks corresponding to the foreground and the ground-truths of class 0, 1 and 2 respectively;
- Let I be a generic image: using a numerical environment (such as NumPy), the single operation $P(fg \,\&\, bw_x)$ can link to the coordinates of the upper-left pixel of the 3 by 3 patches completely

included in the foreground and in the bw_x mask (x class). In this way, patches belonging to each class can easily and rapidly be indexed.

The approach avoids the separate storage of patches as follows:

- The structures containing all the indexes of all the images in the dataset is stored in a very compact file (*.npy*);
- In a csv file, the image source locations (path) of each image in the dataset is saved with information about the number of patches belonging to each class;
- To get all the patches of a specific class from an image *I*, we need to read the entire image from location and select from the *.npy* file the indexes contained in the positions ranged from $x + 1$ to y, where x is the number of patches of the previous image in the file (0 is the first image) and y is the number of patches of the image *I*.

In this way, only ~40Mb has been needed to store the indexes files representing 9.776.172 patches. Moreover, reading all the patches belonging to each image requires only to load the original image (which source is extracted from the csv file).

References

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(2), 213-220.

Ares, G., Martínez, I., Lareo, C., & Lema, P. (2008). Failure criteria based on consumers' rejection to determine the sensory shelf life of minimally processed lettuce. Postharvest biology and technology, 49(2), 255-259.

Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In Neural networks: Tricks of the trade (pp. 437-478). Springer Berlin Heidelberg.

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., 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, 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, 378–384.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.

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.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai (Vol. 14, No. 2, pp. 1137-1145).

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.

LeCun, Y. A., Bottou, L., Orr, G. B., & Müller, K. R. (2012). Efficient backprop. In Neural networks: Tricks of the trade (pp. 9-48). Springer Berlin Heidelberg.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Morais de Oliveira E., Leme D. S., Groenner Barbosa B. H., Pereira Rodarte M., Pereira R. G. F. A. (2016). A computer vision system for coffee beans classification based on computational intelligemce tecniques, Journal of Food Engineering, 171, 22-27.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10) (pp. 807-814).

Nunes, M. C. N., Emondb, J. P., Rautha, M., Deac, S., & Chau, K. V. (2009). Environmental conditions encountered during typical consumer retail display affect fruit and vegetable quality and waste. Postharvest Biology and Technology, 51, 232–241.

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, 2017, 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 & 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(2), 178-183.

Renals, S. (2014). Introduction to neural networks.

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, Computer and Electronics in Agriculture, 137, 59-68.

Sainath, T. N., Vinyals, O., Senior, A., & Sak, H. (2015). Convolutional, long short-term memory, fully connected deep neural networks. In Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on (pp. 4580-4584). IEEE.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929-1958.

Zhang Y., Wang S., Ji G., Phillips P. (2014). Fruit classification using computer vision and feedforward neural network, Journal of Food Engineering, 143, 167-177.

Figure Captions

Figure 1. Typical images acquired with (on the left) and without (on the right) packaging. A color-chart enables color correction. Packaged products exhibit artefacts (such as reflections or glares) that must be removed before image analysis.

Figure 2. The figure emphasizes the complete symmetry between the processing of packaged and unpackaged products apart from the segmentation step required on the bags. In the central part the two learning processes are shown: one to tune the CNN to accomplish image segmentation and the other used to classify the samples on the base of the feature vectors (color histograms). The former is applied only on images if packaged samples while the latter is done on the images related to all the samples (packaged and unpackaged). At the extreme right and at the extreme left there are the processing modules applied to packaged and unpackaged unseen images respectively to classify at run-time.

Figure 3. Changes in ammonium content in fresh-cut iceberg lettuces stored at 8 and 15 °C from quality level 5 (very good) to quality level 1 (very poor).

Each quality level (n = 32) followed by different letters (a–c) are significantly different for $P < 0.05$. Method: 95.0 percent Student-Newman- Keuls.

Figure 4. Original image of fresh-cut lettuce (a) and the image segmented by Convolutional Neural Network (CNN) (b). The intensity of each band (R, G, B) is directly proportional to the probability of the corresponding pixel to belong to one of the three classes (red for bag, green for product, blue for artefacts).

Figure 5. Commercial bag of ready to eat corn salad leaves with plenty of graphical elements (a). The application of the proposed deep learning architecture (CNN): packaging is coded in red, artefacts in blue and useful product in green (b). Finally in c) is reported the region used to evaluate the quality level.

Figure

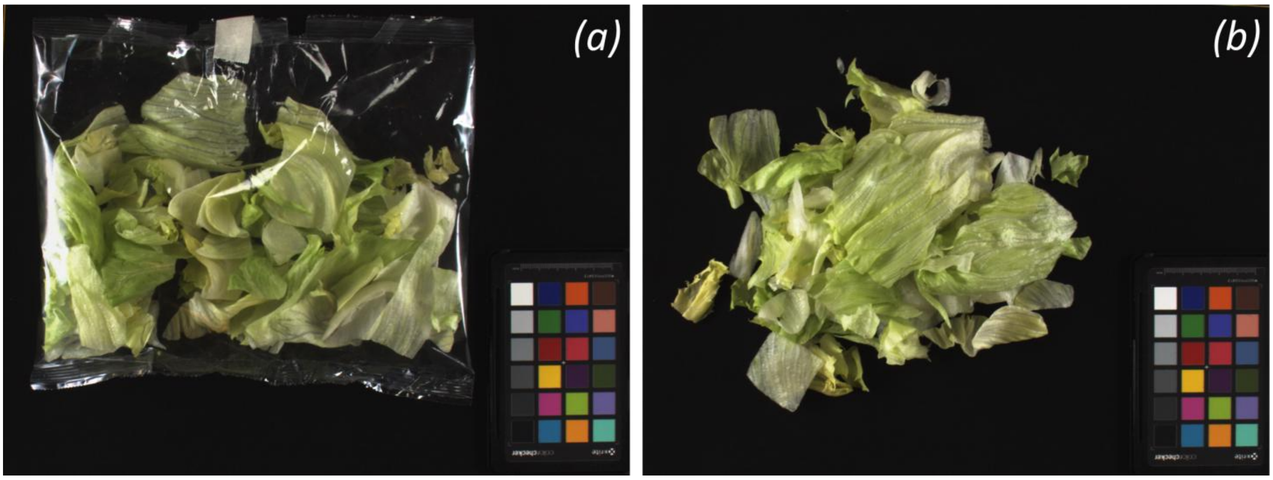

Figure 1. Typical images acquired with (on the left) and without (on the right) packaging. A color-chart enables color correction. Packaged products exhibit artefacts (such as reflections or glares) that must be removed before image analysis.

Figure

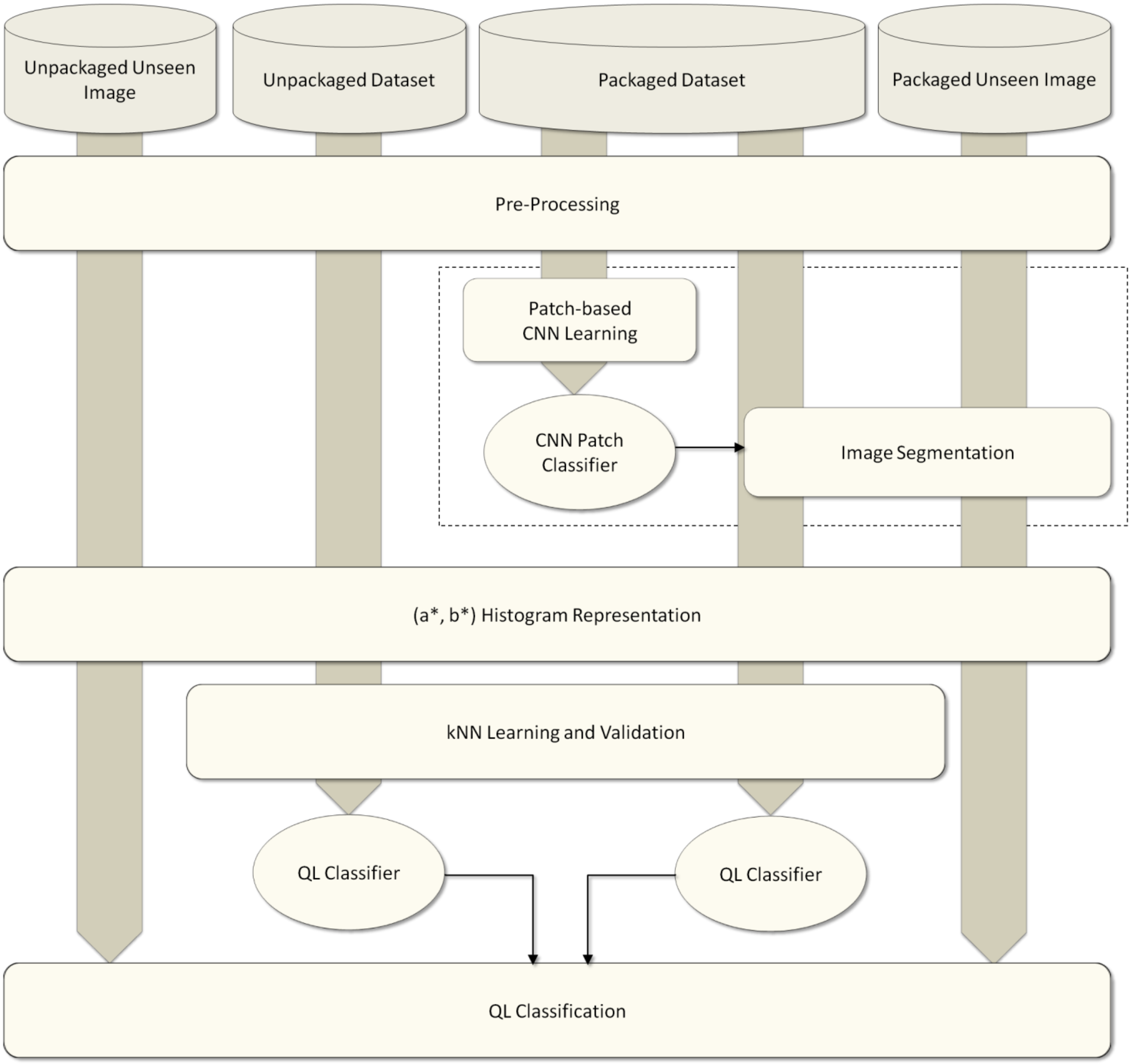

Figure 2. The figure emphasizes the complete symmetry between the processing of packaged and unpackaged products apart from the segmentation step required on the bags. In the central part the two learning processes are shown: one to tune the CNN to accomplish image segmentation and the other used to classify the samples on the base of the feature vectors (color histograms). The former is applied only on images if packaged samples while the latter is done on the images related to all the samples (packaged and unpackaged). At the extreme right and at the extreme left there are the processing modules applied to packaged and unpackaged unseen images respectively to classify at run-time.

Figure

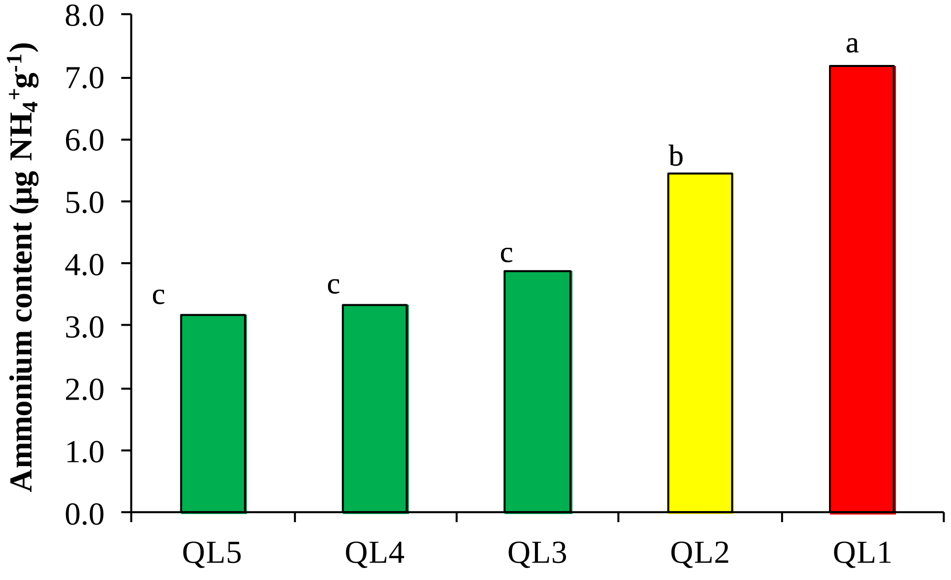

Figure 3. Changes in ammonium content in fresh-cut iceberg lettuces stored at 8 and 15 °C from quality level 5 (very good) to quality level 1 (very poor).

Each quality level (n = 32) followed by different letters (a–c) are significantly different for $P < 0.05$. Method: 95.0 percent Student-Newman- Keuls.

Figure

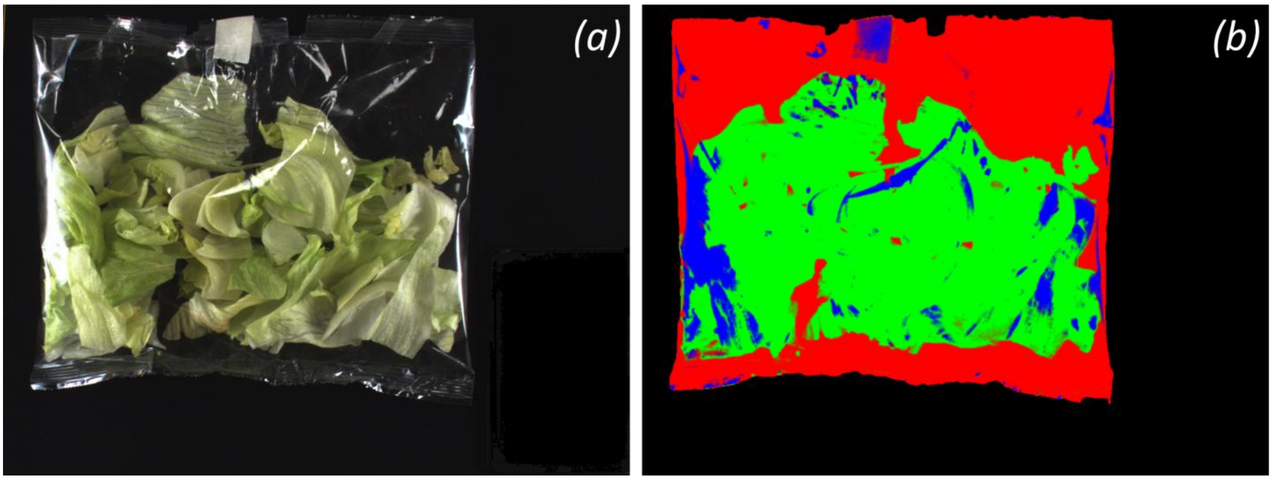

Figure 4. Original image of fresh-cut lettuce (a) and the image segmented by Convolutional Neural Network (CNN) (b). The intensity of each band (R, G, B) is directly proportional to the probability of the corresponding pixel to belong to one of the three classes (red for bag, green for product, blue for artefacts).

Figure

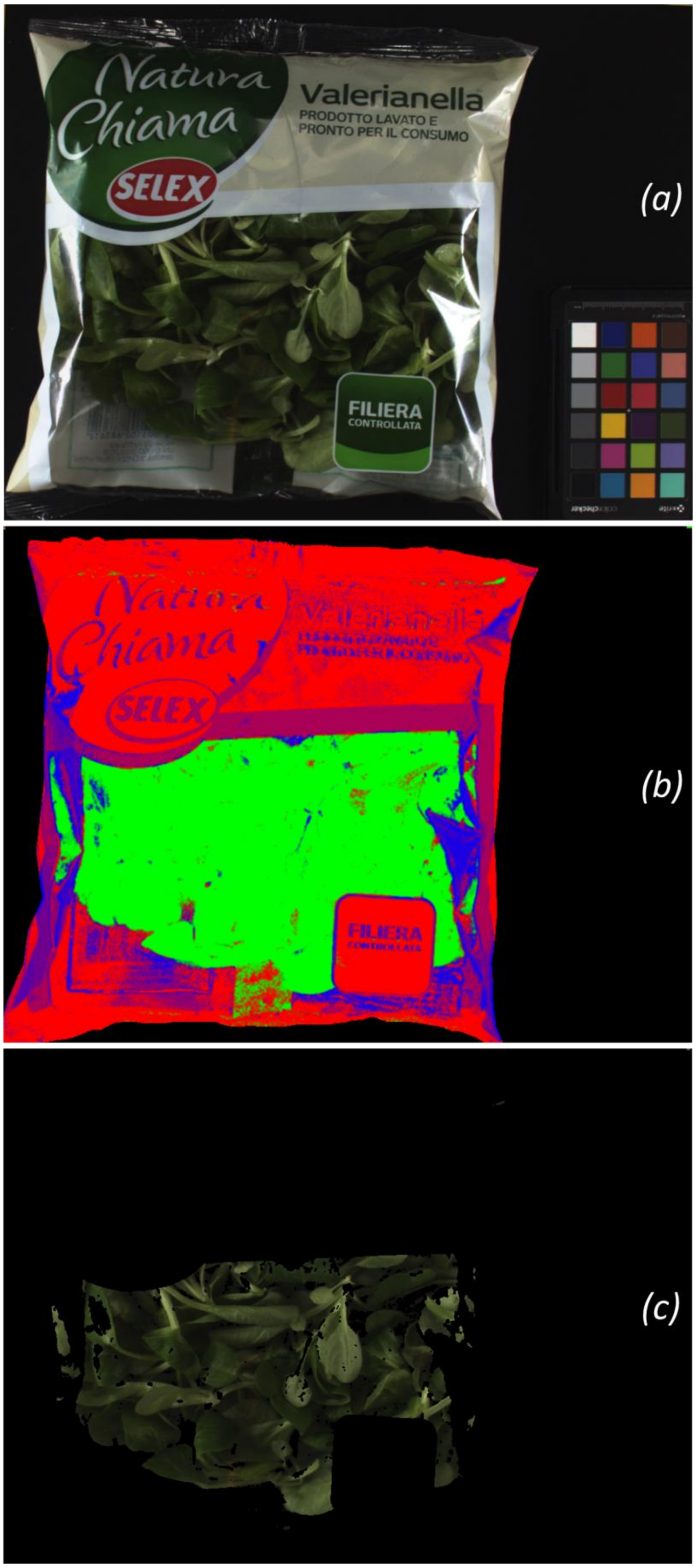

Figure 5. Commercial bag of ready to eat corn salad leaves with plenty of graphical elements (a). The application of the proposed deep learning architecture (CNN): packaging is coded in red, artefacts in blue and useful product in green (b). Finally in c) is reported the region used to evaluate the quality level.