Prog Artif Intell (2018) 7:119–127
https://doi.org/10.1007/s13748-017-0137-1

REGULAR PAPER

Study and classification of plum varieties using image analysis and deep learning techniques

Francisco J. Rodríguez[1] · **Antonio García**[2] · **Pedro J. Pardo**[1] · **Francisco Chávez**[1] · **Rafael M. Luque-Baena**[3]

Received: 10 July 2017 / Accepted: 5 October 2017 / Published online: 19 October 2017

Abstract Currently much of the pre-harvest fruit valuation is still done by farmers or technicians that visually inspect the pieces of fruit. However, this process has great limitations since their decisions have high subjectivity and a thorough analysis of the whole production, or even a significant part of it, is unapproachable. Therefore, computer vision and machine learning techniques are increasingly being introduced into this process. In this work, we deal with the problem of automatically identifying plum varieties at early maturity stages, which is even difficult for the human expert. To face that identification, we propose a two-step procedure. Firstly, captured images are processed to identify the region where the plum appears. Secondly, we determine the plum variety using a deep convolutional neural network. Experimental results show that the proposed system achieves a remarkable behavior, with accuracy values that range from 91 to 97%.

✉ Francisco J. Rodríguez
fjrodriguez@unex.es

Antonio García
antoniogv@unex.es

Pedro J. Pardo
pjpardo@unex.es

Francisco Chávez
fchavez@unex.es

Rafael M. Luque-Baena
rmluque@lcc.uma.es

[1] Department of Computer Science and Telematics, University of Extremadura, Mérida, Spain

[2] Mérida Campus, University of Extremadura, Mérida, Spain

[3] Department of Computer Languages and Computer Science, University of Málaga, Málaga, Spain

Keywords Deep convolutional neural networks · Image segmentation · Plum varieties

1 Introduction

Fruit consumption is essential for a healthy diet [20] thanks to the great contribution of benefits thereof. Therefore, the consumption of fruits and vegetables, both in the domestic market as well as abroad, increasingly demands strict quality parameters, as well as the objective of preserving for longer time the fresh product in the market. Technicians and farmers must be equipped with the necessary knowledge to ensure that their products have the highest possible quality and to predict early harvest characteristics and post-harvest behavior, as well as the date of probable harvest. These predictions will allow the farmer to take corrective agronomic measures (if necessary) and to know the characteristics of his harvest and thus to adapt to the demand of the markets.

Currently, much of the pre-harvest valuation of the fruits is done with the visual appreciation of the farmer or technician, which has great limitations, and their decisions have high subjectivity because they depend on the experience and criterion of the person who does it. The fruit selection that is made in the industry is still mainly based on visual selection by operators, although artificial vision systems are increasingly being introduced. The human factor is decisive in decision making, but a thorough analysis of a production or part of it is unapproachable. It seems reasonable with the technology currently available to implement image analysis and computer processing systems to facilitate decision making and integrate all the information available and relevant.

Computer vision and machine learning techniques have been successfully applied for food analysis in the last years [1–3,5,14,16,17,21]. However, traditional machine

Table 1 Description of the datasets

Dataset	No. images	Harvest time	Maturity weeks
First maturity week (*MW1*)	121	9–13 of May	6 (*Angelino*, *Owent*) and 7 (*BlackSplendor*)
Second maturity week (*MW2*)	147	23–27 of May	8 (*Angelino*, *Owent*) and 9 (*BlackSplendor*)
Third maturity week (*MW3*)	127	13–17 of June	11 (*Angelino*, *Owent*) and 12 (*BlackSplendor*)
Fourth maturity week (*MW4*)	130	4–9 of July	14 (*Angelino*, *Owent*) and 15 (*BlackSplendor*)
All datasets (*All_MW*)	525	–	–

learning techniques employed were limited up to a point since constructing a machine learning system required careful engineering and considerable domain expertise to extract a suitable set of features from the raw data to feed the learning system. This drawback has been overcome in the last few years by deep learning techniques [9] that automatically discover increasingly higher-level features instead of doing it manually and have dramatically improved state of the art in different fields such as speech recognition [6,12], object recognition [8,15,18] and many other domains [4,11].

In this work, we explore the application of a deep learning method for supervised learning known as deep convolutional neural networks (CNN) [18,19] to deal with the automatic identification of plum varieties at early maturity stages. Although the automatic variety identification is mainly performed after harvesting, we intend to assess our proposal's performance at pre-harvest stages since variety identification is much harder than after harvesting. Visual differences at pre-harvest stages are much more subtle, which poses a major challenge even for a human expert. It is important to point out that this work assumes the starting point for a broader project that aims to design a much wider range of pre-harvest validations, including aspects such as product quality, diseases detection and optimal maturity assessment.

The rest of the paper is structured as follows. Section 2.1 describes input data and hardware equipment employed in this work. Section 2 details the two-step method proposed to characterize plum varieties: image segmentation (Sect. 2.2) and plum variety classification (Sect. 2.3). Then, Sect. 3 shows the experimental study performed and the results obtained. Finally, Sect. 4 presents conclusions and further work.

2 Proposed system for studying and characterizing plums

In this section, we further describe the system developed in this work to study and characterize different varieties of plums. This goal is achieved by performing two main operations. Firstly, images obtained by the capture process are subjected to a segmentation process to extract the area of the image in which the plum appears (Sect. 2.2). Finally, we determine the variety of plums by means of a classification process based on deep convolutional neural networks (Sect. 2.3).

2.1 Experimental framework

In this section, we further describe the input data and the hardware devices used in this work. Regarding the input data, we have analyzed images of three plum varieties highly extended in Extremadura. These varieties are *Black Splendor*, *OwenT* and *Angelino* and are characterized by its estimated recollection date. We have included in this work plum varieties with a wide range of recollection dates, from early varieties, such as *Black Splendor*, whose estimated recollection date is mid June, to very late varieties recollected at the end of August, such as *Angelino*, through *OwenT* that is recollected at the beginning of July. Table 1 details the main features of each plum variety used in this work.

For the capture of images, a Retiga-Exi scientific digital camera (QImaging, Canada) has been used. This camera is equipped with a ICX285 progressive scan CCD color sensor (Sony, Japan) with a full resolution of 1392×1040 square pixels and a bit depth of 12 bits per pixel. The lens mount is type C and the communications interface is FireWire, providing up to 10 frames per second at full resolution and bit depth. The main characteristics of this camera are the great sensitivity and linearity in the response over a wide range of luminance values. It also has very little thermal noise due to its cooling system based on a Peltier thermoelectric cell.

All the image capture process was done in a GTI Mini-Matcher Model MM-4e lighting booth equipped with four light sources: Daylight 6500K simulator, Daylight 5000K simulator, illuminant A simulator and an ultraviolet source to measure fluorescence if it is needed. In particular, the D65 simulator, widely used in colorimetry, has been used for all captures. The light booth has a spectrally neutral gray background N7 that facilitates the tasks of image segmentation and is robust against changes of light source (Fig. 1).

Finally, we may highlight that the convolutional neural network has been implemented using the Caffe library[1] and

[1] http://caffe.berkeleyvision.org/.

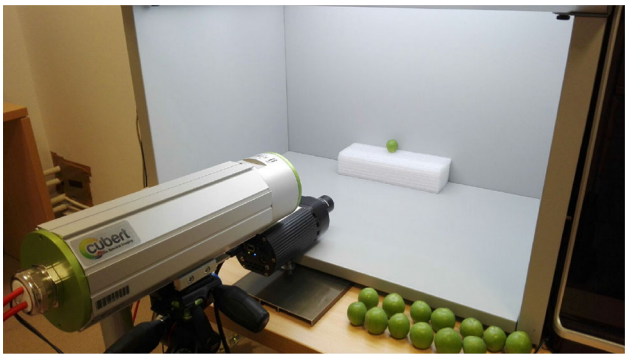


Fig. 1 Image of the capture data process. The scientific digital camera used is the one shown in black color in the center of the image (color figure online)

has been trained on a GPU Nvidia GTX 1080 with 8GB of RAM memory and 2560 computing cores.

2.2 Image segmentation

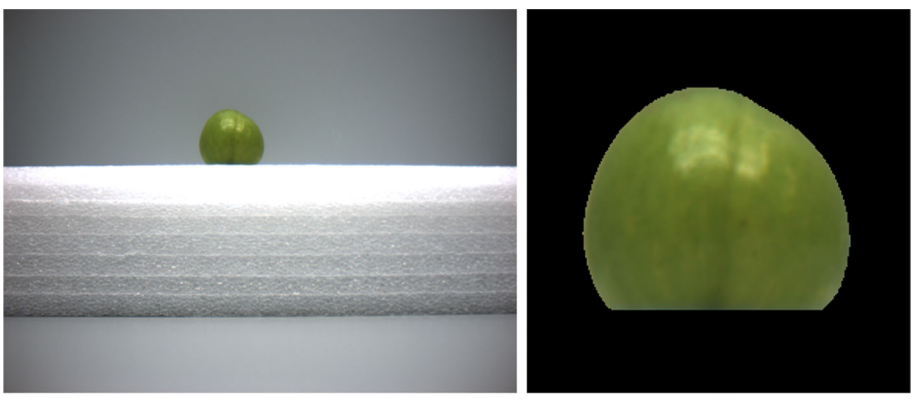
In this section, we summarize the segmentation process for input images (Fig. 2a). Firstly, before each image capture session, we obtain a background image of the scene without any piece of fruit. The aim is to locate the region of the image that contains the plum and keep just that part of the image, as it can be observed in the image on the left from Fig. 2. The region surrounding the plum can be filled with different values, depending on the applied strategy (see in Sect. 3). This process is called segmentation, and its main steps are detailed as follows:

- We subtract the captured image with the piece of fruit from the background image in RGB color space. This process implies that those pixels with the greatest tonality difference correspond to the regions where the plum is located.
- We perform image thresholding with a fixed threshold (40 in this work). This way, we obtain a preliminary region that will be improved later.
- The region obtained as described above still has a strip of background surrounding the plum. We calculate the median value of the pixels of this strip in order to obtain a mask closer to the real one and avoid variations in illumination between captures.
- The subtracted image is thresholded with the median value calculated before. Since the color background is gray and that of the plum is greenish, the color variations between channels will be larger than a relatively low value (we have set the threshold value to 10 in this work) for those pixels corresponding to the plum.
- Later, we clean the mask by filling holes and applying erosion and dilation.
- Finally, we obtain the final segmentation by multiplying the mask and the original image and put it in a region of 256×256, since the CNN will be fed with images of that size.

2.3 Classification using deep convolutional neural networks

In the last few years, a new area known as deep learning have emerged in the field of machine learning. This new area encompasses different techniques that are fundamentally characterized by an hierarchical learning process in which high-level structures are automatically build starting from low-level ones across multiple layers, starting from the raw data (i.e., the pixel values of an image).

Deep learning appears as an alternative to traditional machine learning methods, which require a carefully selection of hand-designed features from which the classifier can detect patterns by means of one, two at most, nonlinear transformations of those features. These classical methods have proven to be quite effective to solve simple or well-delimited problems, but encounter difficulties in dealing with real-world complex problems such as object and speech recognition. By contrast, deep learning techniques



(a) **(b)**

Fig. 2 **a** An example of the captured data. **b** Result of the segmentation process using zeros to fill the outside of the plum

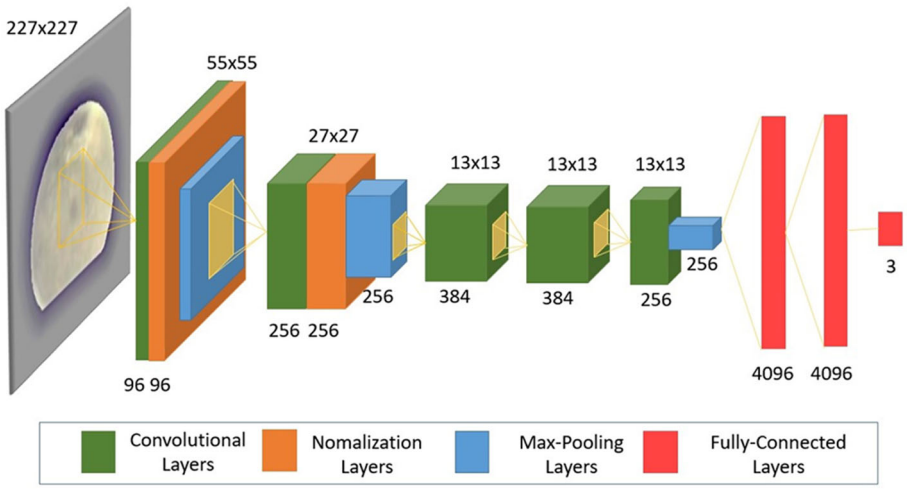


Fig. 3 Alexnet architecture

have hugely improved the state of the art in such complex tasks.

In this work, we focus on a deep learning technique for supervised learning known as deep convolutional neural networks (CNN) [7,10], which has shown an outstanding performance for visual objects recognition. CNN benefits from the spatial structure of input data, that is, the images. In doing so, CNNs use an architecture based on three key principles [13]:

- *Local receptive fields* Each neuron of intermediate layers is connected to a small region of the input layer. This kind of layers are called convolutional layers.
- *Shared weights and bias* All the neurons in a feature map share these parameters. This way, these neurons become specialized in identifying a particular feature in different regions of the image. At the same time, the number of parameters to be set during the network learning phase is reduced notably.
- *Pooling layer* This type of layers receive as input each of the feature maps and produce a new simplified feature map. For example, in *max-pooling layers*, each neuron produces an output with the maximum activation value from a small region of feature maps.

In the last few years, multiple deep CNNs models have been designed. For the purpose of the classification system developed in this work, we have chosen Alexnet network [8]. The latter was originally build to classify high-resolution images from the IMAGENET dataset.[2] Figure 3 depicts a simplified scheme of its architecture.

Alexnet network is composed of five convolutional layers and three fully connected. In contrast to convolutional layers, neurons in fully connected layers are linked to all neurons in previous layer. Local-response normalization layers follow the two first convolutional layers. Its aim is to promote competition among nearby groups of neurons by diminishing responses that are uniformly large in the neighborhood and increasing more pronounced responses [8]. Max-pooling layers are placed after the normalization layers and the fifth convolutional layer. Table 2 summarizes the filter size, the number of filters and stride for each layer.

Table 2 Alexnet's parameters for convolutional and max-pooling layers: filter size, number of filters and stride

Layer	Convolution			Max-pooling	
	Filter size	No. of filters	Stride	Filter size	Stride
First	11×11	96	4	3×3	2
Second	5×5	256	1	3×3	2
Third	3×3	384	1	–	–
Fourth	3×3	384	1	–	–
Fifth	3×3	256	1	3×3	2

3 Experimental results

In this section, the computational experiments to validate the performance of the proposed system for the study and characterization of plum varieties are performed. In order to check the viability of our approach, five different RGB color image datasets are created, featured by the number of plum varieties (*Angelino*, *BlackSplendor* and *Owent*) and level of maturity of each piece. Each dataset correspond to a specific and disjoint maturity time. Since no more than 10–12 plum pieces of each variety are available for dataset, each plum is photographed from 4 different angles to raise the number of patterns. These datasets were described in Table 1. It is important to note that the harvest time is different for each

[2] http://www.image-net.org/.

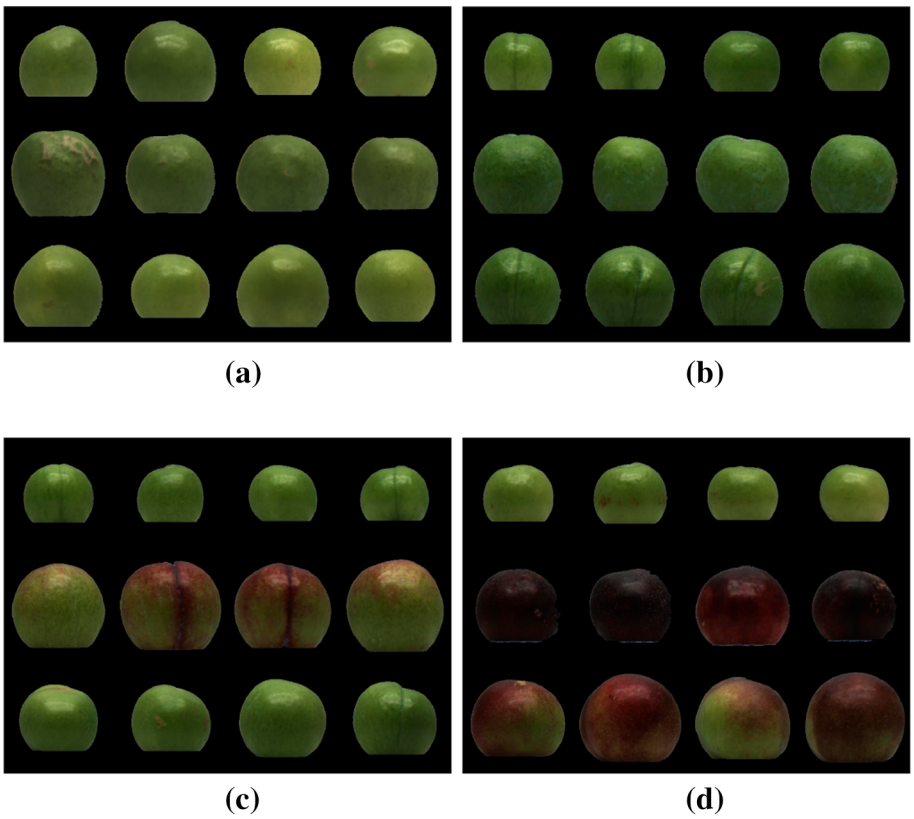


Fig. 4 Plum varieties distributed by maturity weeks. The rows of each figure are associated, from up to bottom, with the varieties: *angelino*, *blacksplendor* and *owent*, respectively. **a** First maturity week (MW1), **b** second maturity week (MW2), **c** third maturity week (MW3), **d** fourth maturity week (MW4)

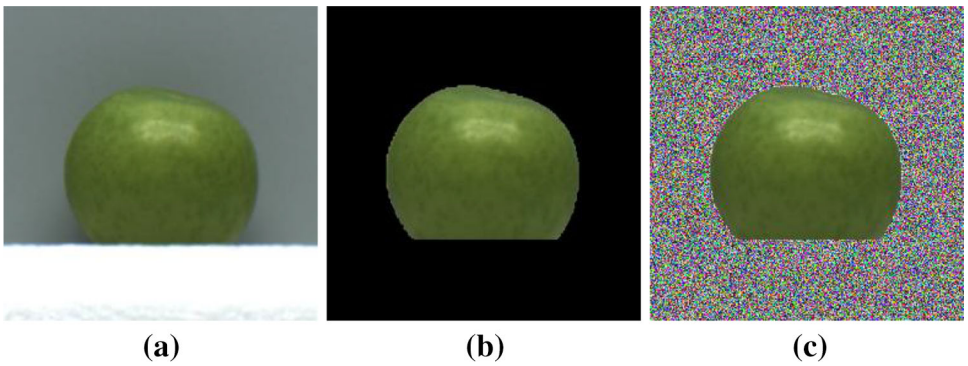


Fig. 5 Different types of input data, highlighting the image fill process. Initially, all approaches resize the image to 256 × 256 pixels. **a** The original ROI. The outer region of the plum is filled with **b** black pixels or **c** noise pixels (color figure online)

dataset which implies that the images for each plum variety could be different from one maturity week to another. It is particularly evident in Fig. 4, where a representation of the plum varieties of each dataset is displayed. Thus, the visual similarity among the three classes is very high for the two first maturity periods, specially in case of the *Angelino* and *Owent* classes in the MW1 dataset, where differences are practically indistinguishable.

A fivefold cross-validation strategy (5-cv) is applied in order to carry out the quantitative experiments. This distribution implies that the 80% of the images are for training and the remaining ones are for testing on each fold. Thus, the neural network approach is trained and tested with the image data of each fold. The accuracy of the 5-cv is the mean of the accuracy of the neural model of all the folds. In order to avoid the bias in the fold creation (distribution of the images on each fold), ten experiments of 5-cv are performed. Therefore, 10 × 5 neural models are computed and used to obtain a global accuracy value in the plum variety recognition.

An study of the hyper-parameters of the convolutional neural model is also performed. Only three hyper-parameters are considered, namely the global learning rate (λ) which modules the learning speed of the network, the number of epochs ($\mathcal{E}$) and the fill image method used to complete the input image. The last parameter is related to the way in which the plum image (input data to the model) is filled after the segmentation, considering in this work three possibilities: the region of interest around the plum previously segmented (named as *standard*); fill the surroundings of the segmented object with zeros (named *black* as strategy); and fill the same

Table 3 Test values for all the parameters

Learning rate	$\lambda = 0.01, 0.005, 0.001$
Epochs	$\mathcal{E} = 500, 750, 1000$
Image fill method	$\mathcal{T} = black, region, noise$
Dataset	$\mathcal{D} = MW1, MW2, MW3, MW4, All\_MW$

The three first rows correspond to the hyper-parameters of the CNN. The last row indicates the different datasets studied. In total, 135 configurations of the CNN are performed using an unbiased methodology

Table 4 Accuracy values (higher is better) for each dataset and plum variety or class (first and second columns)

Dataset	Class	Standard	Black	Noise
MW1	*Angelino*	0.8571 ± 0.016	*0.9143* $\pm$ *0.020*	0.8619 ± 0.019
	BlackSplendor	0.9857 ± 0.012	*0.9881* $\pm$ *0.013*	0.9619 ± 0.017
	Owent	0.8489 ± 0.020	*0.9178* $\pm$ *0.021*	0.8289 ± 0.030
	All	0.8960 ± 0.010	$\mathbf{0.9395 \pm 0.008}$	0.8829 ± 0.011
MW2	*Angelino*	0.9883 ± 0.011	*0.9917* $\pm$ *0.012*	0.9767 ± 0.018
	BlackSplendor	0.8822 ± 0.035	*0.9533* $\pm$ *0.016*	0.8622 ± 0.052
	Owent	*0.8976* $\pm$ *0.030*	0.8952 ± 0.044	0.8738 ± 0.039
	All	0.9299 ± 0.015	$\mathbf{0.9526 \pm 0.014}$	0.9123 ± 0.019
MW3	*Angelino*	0.9754 ± 0.012	*0.9930* $\pm$ *0.009*	0.9825 ± 0.000
	BlackSplendor	*1.0000* $\pm$ *0.000*	*1.0000* $\pm$ *0.000*	0.9944 ± 0.012
	Owent	*0.9424* $\pm$ *0.022*	0.9091 ± 0.000	0.9091 ± 0.000
	All	$\mathbf{0.9739 \pm 0.008}$	0.9731 ± 0.004	0.9669 ± 0.004
MW4	*Angelino*	*1.0000* $\pm$ *0.000*	*1.0000* $\pm$ *0.000*	*1.0000* $\pm$ *0.000*
	BlackSplendor	0.9604 ± 0.012	*0.9792* $\pm$ *0.014*	0.9521 ± 0.017
	Owent	0.8905 ± 0.023	*0.9143* $\pm$ *0.049*	0.9048 ± 0.050
	All	0.9674 ± 0.005	$\mathbf{0.9784 \pm 0.010}$	0.9667 ± 0.010
All_MW	*Angelino*	0.9626 ± 0.012	*0.9680* $\pm$ *0.011*	0.9475 ± 0.008
	BlackSplendor	0.9047 ± 0.016	*0.9170* $\pm$ *0.024*	0.8544 ± 0.022
	Owent	0.8241 ± 0.022	*0.8482* $\pm$ *0.027*	0.7248 ± 0.038
	All	0.9071 ± 0.010	$\mathbf{0.9198 \pm 0.014}$	0.8584 ± 0.014

Each row represents a different plum variety (*angelino*, *blacksplendor* and *owent*). The studied image fill proposals are shown from the third to the fifth columns. Last row for each dataset displays the global accuracy for all the classes. Best results for each dataset and plum class are highlighted in italics. Best results for each dataset and all plum classes are highlighted in bold

Table 5 Best accuracy values (higher is better) for each dataset

Dataset	Best accuracy	Parameters
MW1	0.9395 ± 0.008	$\mathcal{T} = black, \mathcal{E} = 1000, \lambda = 0.001$
MW2	0.9526 ± 0.014	$\mathcal{T} = black, \mathcal{E} = 1000, \lambda = 0.005$
MW3	0.9739 ± 0.008	$\mathcal{T} = region, \mathcal{E} = 750, \lambda = 0.01$
MW4	0.9784 ± 0.010	$\mathcal{T} = black, \mathcal{E} = 1000, \lambda = 0.005$
All_MW	0.9198 ± 0.014	$\mathcal{T} = black, \mathcal{E} = 1000, \lambda = 0.01$

Each row represents the best parameter, together with the metric value obtained

parts with noise pixels (named *noise*), i.e., RGB values from (0,0,0) to (255,255,255) obtained randomly from an uniform distribution. A representation of each strategy can be observed in Fig. 5. According to the learning of the convolutional neural model, it should be noted that the current learning rate is reduced during the learning process, decreasing by 10% each $\mathcal{E}/3$ times.

A summary of the values for each one of the studied parameters is presented in Table 3. Thus, the validation strategy to obtain fair measures of the performance (10 experiments $\times$ 5 folds) is applied for each of the 135 configurations considered.

Next, the results obtained by the classification approach should be analyzed. Tables 4 and 5 present these accuracy results by each dataset. Specifically, in Table 4 best results

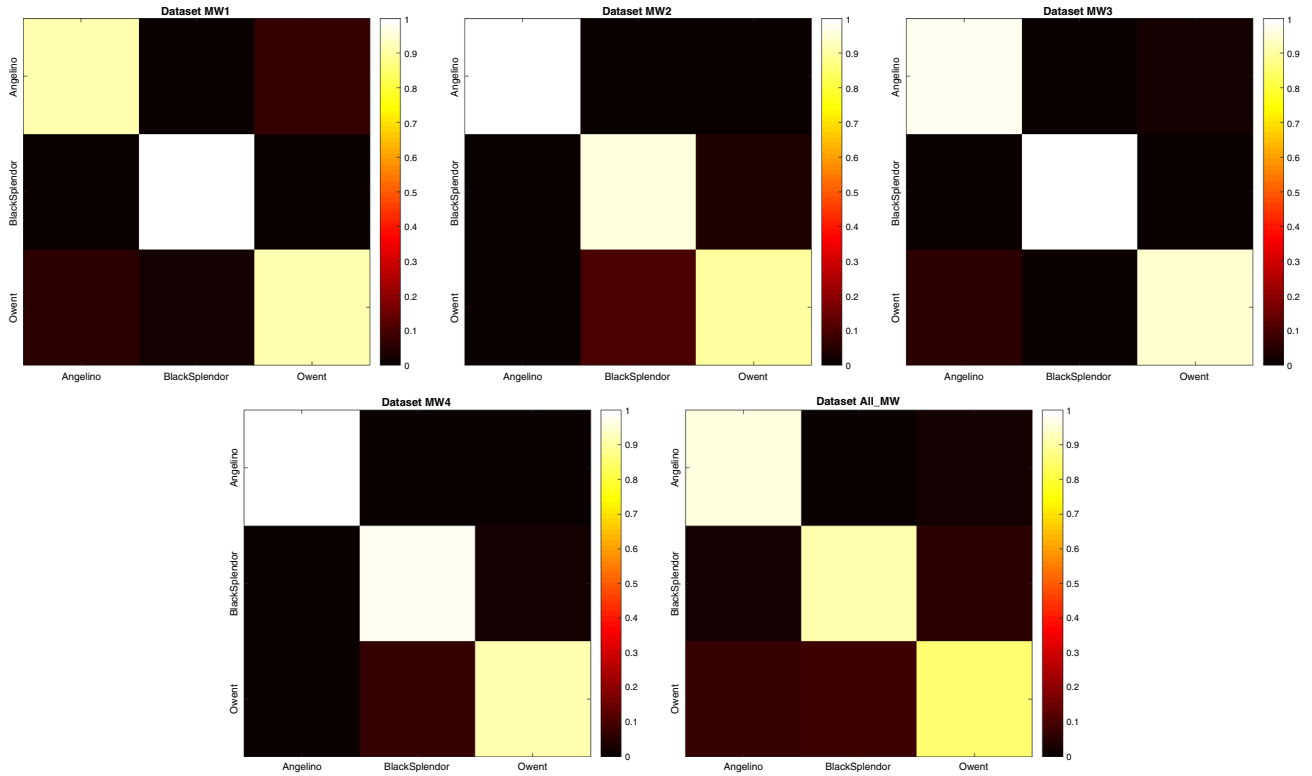


Fig. 6 Confusion matrices for all the datasets analyzed, by training the convolutional neural network with the best parameter configuration (Table 5). The row dimension indicates the real plum variety, while the column dimension reflects the predicted plum variety. The clearer the blocks on the diagonal, the greater the accuracy of the method

are obtained by plum variety (*Angelino*, *BlackSplendor* and *Owent*) and over the whole dataset (*All* label, in the table). This information is aggregated by the image fill method, in order to identify the best way to complete the image as input data. The shown accuracy is the best one taking into consideration the remaining hyper-parameters (learning rate and epoch). Table 5 summarizes the model with best accuracy for each dataset and indicates the values of the parameters. Furthermore, and also related with the experimental results, the confusion matrices for each dataset, which show the errors between the predicted plum variety by the model and the real one, are displayed in Fig. 6. The clearer the blocks on the diagonal, the greater the accuracy of the method. If a block outside the diagonal is not very dark, it implies that, given a plum image, there is an error between the output of the network (prediction) and the actual class to which it belongs. For example, by observing the results of MW1 and MW3, it is noted that the blocks relating to the varieties *Angelino* and *Owent* present more brown tones than blocks relating to the *BlackSplendor* class. Therefore, and in view of the results, we can make the following comments:

– The accuracy values for all datasets are very remarkable, all of them above 90%.
– The best image fill method is *black* since it improves the results in four of five of the datasets analyzed. In fact, the results outperform the 94% of success, and in some datasets as *MW1* the visual difference between varieties is practically indistinguishable (see Fig. 4a).
– Although the dataset *All_MW* is the aggregation of all the plum patterns of different maturity weeks (*MW1*, *MW2*, *MW3* and *MW4*) and has a greater number of patterns, there is a larger variability in the images of each variety, which implies a less accuracy rate.
– In general, the most complex plum variety to distinguish is *Owent*, due to its similarity in the first two weeks with *Angelino* and its similarity the last week with *BlackSplendor*.
– As discussed above, and according to Table 5, it is observed that the varieties are harder to differentiate in the first weeks of maturation than in the latter, as is visually observed in Fig. 4.

On the other hand, the results observed in Fig. 7 try to isolate the values of the hyper-parameters and check their effect on the performance of the CNN. For this, the success rates of the configurations with a specific value of each parameter are obtained and the mean and standard deviation are computed.

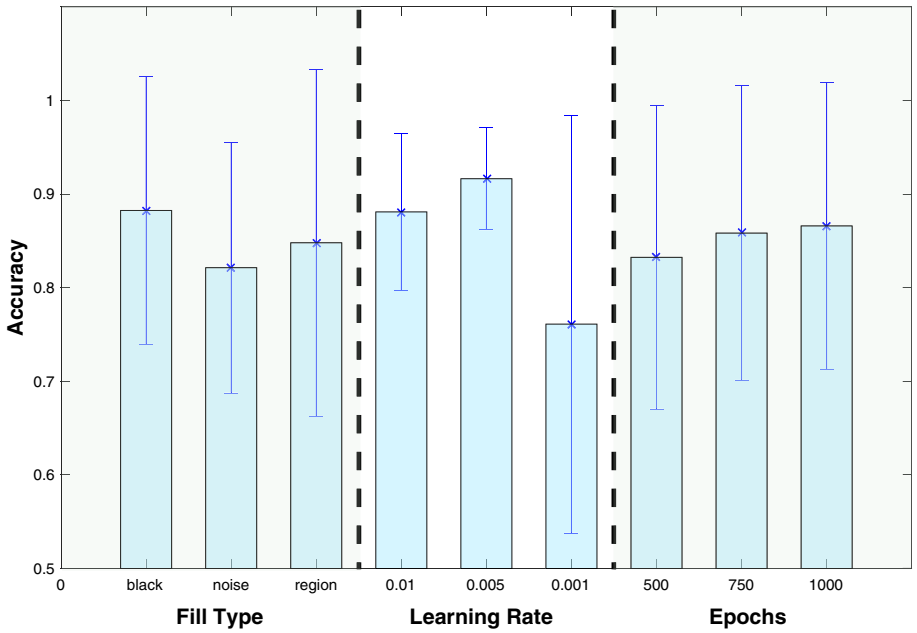


Fig. 7 Parameters robustness according to the accuracy. Each bar corresponds to a value of the three parameters studied, namely fill type, learning rate and epochs, while the blue line for each bar indicates the standard deviation

It is observed that the classification improves when the input image is filled with zeros, the learning rate is medium/high and the number of epochs is high, so that the network can effectively learn the discriminant features for the distinction of plum varieties.

4 Conclusions and future work

In this work, we have proposed a system to analyze and characterize three different plum varieties at early maturity stages: *Black Splendor*, *OwentT* and *Angelino*. The system is composed of a first stage to capture and segment plum images and, then, a second stage where a deep convolutional neural network classifies the images previously processed. The accuracy obtained ranges from 91 to 97%, which is very remarkable given the difficulty of this task at early maturity stages, even for a human expert.

The promising results achieved in this work make us consider other interesting avenues of research. On the one hand, we plan to improve the system to face other challenges in automatic variety identification of fruits using images such as occlusion caused by leaves, branches or other fruits and variance in natural illumination or scale. Secondly, we intend to include the automatic identification of other fruit issues such as maturity stage, quality, and diseases using additional information sources such as hypersprectal and thermal images.

Acknowledgements This work was supported by projects IB16035, IB16004, and GR15130 of the Regional Government of Extremadura, Department of Commerce and Economy, co-financed by the European Regional Development Fund, "A way to build Europe." Additionally, it is also partially supported by projects TIN2014-53465-R, TIN2016-75097-P and P12-TIC-657.

References

1. Abbott, J.A.: Quality measurement of fruits and vegetables. Postharvest Biol. Technol. **15**(3), 207–225 (1999)
2. Ariana, D., Guyer, D.E., Shrestha, B.: Integrating multispectral reflectance and fluorescence imaging for defect detection on apples. Comput. Electron. Agric. **50**(2), 148–161 (2006)
3. Cho, B.K., Kim, M.S., Baek, I.S., Kim, D.Y., Lee, W.H., Kim, J., Bae, H., Kim, Y.S.: Detection of cuticle defects on cherry tomatoes using hyperspectral fluorescence imagery. Postharvest Biol. Technol. **76**, 40–49 (2013)
4. Ciodaro, T., Deva, D., de Seixas, J.M., Damazio, D.: Online particle detection with neural networks based on topological calorimetry information. J. Phys. Conf. Ser. **368**(1), 012–030 (2012)
5. Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., Blasco, J.: Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food Bioprocess Technol. **4**(4), 487–504 (2011)
6. Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, R., Jaitly, A., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N., Kingsbury, B.: Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process. Mag. **29**(6), 82–97 (2012)
7. Kavukcuoglu, K., Ranzato, K., LeCun, M.Y.: What is the best multi-stage architecture for object recognition? In: 2009 IEEE 12th International Conference on Computer Vision, pp. 2146–2153 (2009)
8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105. Curran Associates, Inc (2012)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)

10. LeCun, Y., Huang, F.J., Bottou, L.: Learning methods for generic object recognition with invariance to pose and lighting. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004, vol. 2, pp. 97–104 (2004)
11. Ma, J., Sheridan, R.P., Liaw, A., Dahl, G.E., Svetnik, V.: Deep neural nets as a method for quantitative structure–activity relationships. J. Chem. Inf. Model. **55**(2), 263–274 (2015)
12. Mikolov, T., Deoras, A., Povey, D., Burget, L., Cernocky, J.: Strategies for training large scale neural network language models. In: Proceedings of Automatic Speech Recognition and Understanding, pp. 196–201 (2011)
13. Nielsen, M.A.: Neural Networks and Deep Learning. Determination Press, Oxford (2015)
14. Okamoto, H., Lee, W.S.: Green citrus detection using hyperspectral imaging. Comput. Electron. Agric. **66**(2), 201–208 (2009)
15. Olmos, R., Tabik, S., Herrera, F.: Automatic handgun detection alarm in videos using deep learning. Neurocomput. (2017). doi:10.1016/j.neucom.2017.05.012
16. Pathare, P.B., Opara, U.L., Al-Said, F.A.J.: Colour measurement and analysis in fresh and processed foods: a review. Food Bioprocess Technol. **6**(1), 36–60 (2013)
17. Riquelme, M., Barreiro, P., Ruiz-Altisent, M., Valero, C.: Olive classification according to external damage using image analysis. J. Food Eng. **87**(3), 371–379 (2008)
18. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9 (2015)
19. Tompson, J., Goroshin, R., Jain, A., LeCun, Y., Bregler, C.: Efficient object localization using convolutional networks. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 648–656 (2015)
20. Vicente, A.R., Manganaris, G.A., Sozzi, G.O., Crisosto, C.H.: Nutritional quality of fruits and vegetables. In: Florkowski, WJ., Prussia, S.E., Shewfelt, R.L., Brueckner, B. (eds.) Postharvest Handling, Food Science and Technology, Second Edition. pp. 57–106. Academic Press, San Diego (2009)
21. Wu, D., Sun, D.W.: Colour measurements by computer vision for food quality control: a review. Trends Food Sci. Technol. **29**(1), 5–20 (2013)