IOCAG 2022

MDPI

Proceeding Paper

Postharvest Authentication of Potato Cultivars Using Machine Learning to Provide High-Quality Products †

Ewa Ropelewska *

Fruit and Vegetable Storage and Processing Department, The National Institute of Horticultural Research, Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland; ORCID: https://orcid.org/0000-0001-8891-236X

* Correspondence: ewa.ropelewska@inhort.pl

† Presented at the 1st International Online Conference on Agriculture—Advances in Agricultural Science and Technology (IOCAG2022), 10–25 February 2022; Available online: https://iocag2022.sciforum.net/.

Abstract: The correct identification of potato cultivars is of great importance for both processing and cultivation due to differences in properties. The objective of this study was to discriminate potato cultivars 'Irga', 'Riviera' and 'Colomba' using models developed based on selected textures of tuber images converted to color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*. The highest accuracies of cultivar identification of potato tubers reached 99% for the IBk classifier and 98% for Multilayer Perceptron. The developed models can be used to avoid mixing potato cultivars. Postharvest cultivar authentication can contribute to providing consumers with high-quality products.

Keywords: potato tubers; cultivar identification; image features; discriminative models

Citation: Ropelewska, E. Postharvest Authentication of Potato Cultivars Using Machine Learning to Provide High-Quality Products. **2022**, *4*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: date

1. Introduction

Potatoes (*Solanum tuberosum* L.) are one of the most important staple foods. Potato tubers contain carbohydrates, phenolic compounds, fiber, minerals, vitamins. Potatoes provide phenolics and antioxidants. Due to its chemical composition, potato is characterized by potential health benefits including anti-carcinogenic, anti-diabetic, or anti-inflammatory effects. However, the cultivars bred for prioritizing high yield can have a lower content of micronutrients and worse taste. To avoid consumer complaints, potato cultivars with high quality and desirable flavors should be selected [1]. The cultivar influences potato quality and safety. The potato cultivars available in the market differ in properties. However, some potato cultivars can contain morphologically similar tubers, but with different qualities. The sorting of potato tubers requires classifying them into cultivars. The precise identification of cultivars often requires a destructive, expensive and time-consuming approach. The prevention of falsification of potatoes in retail marketing can be facilitated by the development of the robust, rapid, and universal technic of potato cultivar classification [2].

The potato cultivars may differ in chemical, physical, sensory and functional properties [3,4]. Individual cultivars may be characterized by differences in texture, color, size and shape. These features may be evaluated manually by experts using visual observation. However, the evaluation is subjective, labor-intensive, time-consuming and requires empirical knowledge [5]. The capability of artificial vision systems goes beyond human capacity. Artificial systems substitute human inspection and improve its capability [6]. Computer vision provides an objective, non-destructive, fast and accurate evaluation of quality characteristics. Therefore, computer vision can be successfully used for cultivar identification as well as shape classification, quality grading and defect detection. Machine vision is used to recognize vegetables effectively [7]. However, in some cases, the application of deep learning can improve the accuracy of image recognition [8]. Based on the available literature, machine learning algorithms were used, e.g., for the cultivar identification of whole potato tubers [9,10] as well as the cultivar discrimination of raw and processed flesh (slices) of potato [11].

Due to the potato cultivars may differ in chemical, physical, sensory and functional properties, the correct identification of potato cultivars is of great importance for both processing and cultivation. The application of machine learning enables the non-destructive, objective, repeatability and inexpensive quality evaluation. The objective of this study was to discriminate potato cultivars using models developed based on textures of tuber images.

2. Materials and Methods

The potatoes belonging to cultivars 'Irga', 'Riviera' and 'Colomba' were harvested from fields located in Poland. The washed, cleaned and air-dried tubers of each cultivar were imaged using a digital camera in one hundred repetitions. The potato tubers were imaged against a black background. The images were processed using the Mazda application (Łódź University of Technology, Institute of Electronics, Poland) [12]. The acquired images were converted to color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*. Due to the use of a black background, segmentation was facilitated, and the lighter tubers were separated from the background. Each tuber was a single region of interest (ROI). For each ROI (potato tuber image), about two thousand texture parameters based on the run-length matrix, co-occurrence matrix, autoregressive model, gradient map, histogram and Haar wavelet transform were computed. The discrimination analysis of potato tubers belonging to three cultivars was performed using the WEKA machine learning application (Machine Learning Group, University of Waikato) [13,14]. The attributes with the highest discriminative power were selected using the Best First with the CFS subset evaluator, Linear Forward Selection, Genetic Search and the Ranker in conjunction with OneR attribute evaluator. The selected textures were used to build discriminative models. Several models with a different number of textures were tested using the classifiers from the groups of Lazy, Functions, Rules, Trees, Bayes and Meta [15]. In the case of each model, the average accuracy (%), TP Rate—True Positive Rate, FP Rate—False Positive Rate, Precision, F-Measure and ROC Area—Receiver Operating Characteristic Area were determined.

3. Results and Discussion

The most successful model included 29 selected attributes (1 from color channel *R*, 2 from channel *G*, 1 from channel *B*, 7 from channel *a*, 2 from channel *b*, 1 from channel *X*, 3 from channel *Z*, 2 from channel *U*, 10 from channel *V*). No texture from images converted to color channels *L* and *Y* was included in the model.

The highest average accuracy of cultivar identification of potato tubers reached 99% for the IBk classifier from the group of Lazy. The values of TP Rate (0.987), Precision (0.987) and F-Measure (0.987) were the highest. The value of FP Rate was the lowest and was equal to 0.007. The Multilayer Perceptron (Functions) classified potato tubers 'Irga', 'Riviera' and 'Colomba' with an average accuracy equal to 98%. Also, correct classifications (97%) were obtained for models built using the PART (Rules), J48, LMT and Random Forest (Trees) and Logistic (Functions) classifiers. Slightly lower accuracies were obtained for models developed using the Bayes Net (Bayes) (96%) and Logit Boost (Meta) (95%) classifiers. In the case of the Logit Boost classifier, the TP Rate, Precision and F-Measure were characterized by the lowest value of 0.947. Whereas the FP Rate (0.027) was the highest.

Table 1. The results of potato cultivar discrimination of 'Irga', 'Riviera' and 'Colomba' using the model built based on selected textures of tuber images converted to individual color channels.

Classifier	Average Accuracy (%)	TP Rate	FP Rate	Precision	F-Measure	ROC Area
		(Weighted Average)				
IBk	99	0.987	0.007	0.987	0.987	0.990
Multilayer Perceptron	98	0.980	0.010	0.980	0.980	1.000
PART	97	0.973	0.013	0.974	0.973	0.980
J48	97	0.973	0.013	0.974	0.973	0.981
LMT	97	0.973	0.013	0.974	0.973	0.999
Random Forest	97	0.967	0.017	0.967	0.967	0.997
Logistic	97	0.967	0.017	0.968	0.967	0.986
Bayes Net	96	0.960	0.020	0.960	0.960	0.997
Logit Boost	95	0.947	0.027	0.947	0.947	0.991

TP Rate—True Positive Rate, FP Rate—False Positive Rate, ROC Area—Receiver Operating Characteristic Area.

The developed models discriminated three potato cultivars with high average accuracies. The results were very satisfactory. Due to this, the models can be used in practice for the postharvest cultivar authentication and avoid mixing potato cultivars. It can contribute to providing consumers with high-quality food products including only potato tubers belonging to cultivars with desired properties.

4. Conclusions

The application of machine learning allowed for the discrimination of three potato cultivars based on the tuber texture parameters with a very hight correctness of up to 99%. The effectiveness of image-based techniques for distinguishing different potato cultivars has been proven. Therefore, the results can be of great practical importance. The models can be used for postharvest authentication of potato cultivars to provide consumers with products containing ingredients with desirable properties. As a result, the products can be of high quality.

Author Contributions: Conceptualization, E.R.; methodology, E.R.; software, E.R.; validation, E.R.; formal analysis, E.R.; investigation, E.R.; resources, E.R.; data curation, E.R.; writing—original draft preparation, E.R.; writing—review and editing, E.R..; visualization, E.R.; supervision, E.R. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Yang, H.; Liao, Q.; Ma, L.; Luo, W.; Xiong, X.; Luo, Y.; Yang, X.; Du, Ch.; He, Y.; Li, X.; Gao, D.; Xue, X.; Shang, Y. Features and genetic basis of chlorogenic acid formation in diploid potatoes. *Food Chem. Mol. Sci.* **2021**, *3*, 100039.
2. Kasampalis, D.S.; Tsouvaltzis, P.; Ntouros, K.; Gertsis, A.; Moshou, D.; Siomos, A.S. Rapid Nondestructive Postharvest Potato Freshness and Cultivar Discrimination Assessment. *Appl. Sci.* **2021**, *11*, 2630.
3. Pardo, J.E.; Alvarruiz, A.; Perez, J.I.; Gomez, R.; Varon, R. Physical-chemical and sensory quality evaluation of potato varieties (*Solanum tuberosum* L.). *J. Food Qual.* **2000**, *23*, 149–160.
4. Mbougueng, P.D.; Dzudie, T.; Scher, J.; Tchiégang, C. Physicochemical and functional properties of some cultivars of Irish potato and cassava starches. *J. Food Technol.* **2008**, *6*, 139–146.

5. Yamamoto, K.; Ninomiya, S.; Kimura, Y.; Hashimoto, A.; Yoshioka, Y.; Kameoka, T. Strawberry cultivar identification and quality evaluation on the basis of multiple fruit appearance features. *Comput. Electron. Agric.* **2015**, *110*, 233–240.
6. 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.* **2011**, *4*, 4870–504.
7. Brosnan, T.; Sun, D.-W. Improving quality inspection of food products by computer vision—A review. *J. Food Eng.* **2004**, *61*, 3–16.
8. Oishi, Y.; Habaragamuwa, H.; Zhang, Y.; Sugiura, R.; Asano, K.; Akai, K.; Shibata, H.; Fujimoto, T. Automated abnormal potato plant detection system using deep learning models and portable video cameras. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102509.
9. Przybył, K.; Górna, K.; Wojcieszak, D.; Czekała, W.; Ludwiczak, A.; Przybylak, A.; Boniecki, P.; Koszela, K.; Zaborowicz, M.; Janczak, D.; et al. The recognition of potato varieties using neural image analysis method. In Proceedings of the Seventh International Conference on Digital Image Processing: ICDIP 2015, Proc of SPIE, Los Angeles, CA, USA, 9–10 April 2015; Falco, C.M., Jiang, X., Eds.; Volume 9631, p. 963116.
10. Azizi, A.; Abbaspour-Gilandeh, Y. Identifying Irregular Potatoes by Developing an Intelligent Algorithm Based on Image Processing. *Tarim Bilimleri Derg. J. Agric. Sci.* **2016**, *22*, 32–41.
11. Ropelewska, E. Effect of boiling on classification performance of potatoes determined by computer vision. *Eur. Food Res. Technol.* **2021**, 247, 807–817.
12. Szczypinski, P.M.; Strzelecki, M.; Materka, A.; Klepaczko, A. MaZda—A software package for image texture analysis. *Comput. Meth. Prog. Biomed*. **2009**, *94*, 66–76.
13. Bouckaert, R.R.; Frank, E.; Hall, M.; Kirkby, R.; Reutemann, P.; Seewald, A.; Scuse, D. *WEKA Manual for Version 3-9-1*; The University of Waikato: Hamilton, New Zealand, 2016.
14. Frank, E.; Hall, M.A.; Witten, I.H. *The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed.; Morgan Kaufmann: 2016.
15. Witten, I.H.; Frank, E. Data mining. In *Practical Machine Learning Tools and Techniques*, 2nd ed.; Elsevier: San Francisco, CA, USA, 2005.