Effect of boiling on classification performance of potatoes determined by computer vision

Ewa Ropelewska¹

Received: 27 October 2020 / Revised: 23 November 2020 / Accepted: 29 November 2020 / Published online: 30 January 2021
© The Author(s) 2021

Abstract

The aim of this study was to evaluate the effect of potato boiling on the correctness of cultivar discrimination. The research was performed in an objective, inexpensive and fast manner using the image analysis technique. The textures of the outer surface of slice images of raw and boiled potatoes were calculated. The discriminative models based on a set of textures selected from all color channels (R , G , B , L , a , b , X , Y , Z , U , V , S), textures selected for color spaces and textures selected for individual color channels were developed. In the case of discriminant analysis of raw potatoes of cultivars ‘Colomba’, ‘Irga’ and ‘Riviera’, the accuracies reached 94.33% for the model built based on a set of textures selected from all color channels, 94% for Lab and XYZ color spaces, 92% for color channel b and 92.33% for a set of combined textures selected from channels B , b , and Z . The processed potatoes were characterized by the accuracy of up to 98.67% for the model including the textures selected from all color channels, 98% for RGB color space, 95.33% for color channel b , 96.67% for the model combining the textures selected from channels B , b , and Z . In the case of raw and processed potatoes, the cultivar ‘Irga’ differed in 100% from other potato cultivars. The results revealed an increase in cultivar discrimination accuracy after the processing of potatoes. The textural features of the outer surface of slice images have proved useful for cultivar discrimination of raw and processed potatoes.

Keywords Raw potatoes · Boiled potatoes · Cultivars · Image processing · Texture parameters · Discriminative classifiers

Introduction

Potato (*Solanum tuberosum* L.) is considered one of the major food crops. Potato is a very nutritive vegetable. Potato may contain about 75% water. The cooked potatoes consist of 20% carbohydrate, 1.8% fiber, 1.7% protein, vitamin C, calcium, riboflavin, niacin, other minerals, and provide 86 kcal energy. Due to the presence of antioxidants, phenolic compounds and low cholesterol, eating potatoes may protect people against cancer, cardiovascular diseases and disorders and trauma of cells [1]. Potato may be consumed after processing, in different forms, such as, for example, salads, potato soups, mashed potatoes, French fries, potato chips [2]. Due to great importance in the human diet, research

regarding various aspects of potato evaluation may prove to be very significant [3].

Potatoes are grown in hundreds of countries around the world. Therefore, due to the diverse climate and soil properties in different sites, the development of new cultivars of potato was desired. Crossbreeding resulted in the occurrence of thousands of potato cultivars [4]. The potato cultivars may be different in terms of physical, chemical and functional properties, processing potentials, as well as post-processing parameters [1]. Heat processing in water (boiling) or oil (frying) before the consumption of potatoes causes the gelatinization of the starch resulting in the softening of the inner core [2]. The sensory properties, such as, for example, hardness, adhesiveness, cohesiveness, graininess, mealiness, moistness of boiled potato may differ depending on cultivar. Furthermore, the sensory texture quality of cooked potatoes may be predicted by image analysis of raw potatoes [5].

A machine vision with the grading system may be used for automated inspection of potatoes. Postharvest classification and sorting may be important for ensuring consumer expectations in terms of potato quality. Potato grading is commonly

✉ Ewa Ropelewska
ewa.ropelewska@inhort.pl

¹ Fruit and Vegetable Storage and Processing Department, The National Institute of Horticultural Research, Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland

performed based on the shape, size, and defects of tubers [6]. Imaging can be performed using various devices, such as, for example, CCD camera, hyperspectral camera, ultraviolet camera, NMR, X-ray CT. It may allow the evaluation of the external, as well as internal features of potatoes [7, 8]. Computer vision ensures high accuracy with low costs, high repeatability, and flexibility. Due to these facts, inspection systems that use machine vision can be successfully applied in modern manufacturers, for example, as a part of a food processing plant for real-time quality control [9]. Systems using machine vision can be used besides other techniques and methods of the identification of potato cultivar, for example, involving human experts, molecular markers, spectroscopy, some of which may be subjective with a higher risk for human error, time-consuming, more expensive or require sophisticated laboratory facilities [3, 4, 10].

The identification of cultivar may be a significant factor at each step in the agricultural production chain. It can be important, among others, for farmers, growers, variety registration agencies, plant breeders, processors, bulk handlers, marketers, end-users [3]. For example, in industrial usages, for the production of potato chips, identification of the cultivar of potato tubers is very important for further cutting and frying [3]. However, after cutting the potatoes intended for the production of chips, the identification of cultivar may also be necessary because different potato cultivars may require different frying conditions and mixing of slices belonging to different cultivars could affect the final product (potato chips). Therefore, the development of models for the cultivar determination of raw potato slices based on slice features may have practical application to verify cultivar authenticity, detect adulteration and avoid cultivar mixing in potato processing. According to Bărăscu et al. [11], the potato cultivars can be grouped into potato cooking types with different usages, such as A, A/B with properties suitable for salad and other dishes; B, B/C for multiple uses; C for chips, French fries; D for starch and alcohol industries. Thus, cooked potato subjected to further processing also should belong to the appropriate potato cooking type for a specific application. Therefore, identifying the cultivar in an intermediate step of meal preparation can also prove useful.

The aim of this study was to evaluate the effect of potato boiling on the correctness of cultivar discrimination. The discriminative models based on textures of slice images of raw and boiled potatoes were developed. Due to the application of image analysis, the research was performed in an objective, inexpensive and fast manner.

Materials and methods

Materials

The tubers of potatoes of three cultivars: ‘Colomba’, ‘Irga’ and ‘Riviera’ were applied in the experiments. The samples of potatoes with a size larger than 35 mm were collected from north-eastern Poland. The studies were performed with the use of two groups of potatoes: raw and processed. Raw potatoes were washed and cleaned before the experiment. The potatoes intended for processing were also washed and cleaned, and then whole potatoes were subjected to boiling. The potatoes were cooked in boiling water until tender for 25 min. After cooking, the potatoes were air-dried and cooled down at room temperature. The slices of raw and processed potatoes were obtained by cutting each potato with a knife. The slices were cut from the middle part of the tubers. One hundred raw potatoes and 100 boiled potatoes of each cultivar were tested. For each potato, one slice was selected and subjected to tests. The image analysis was carried out with the use of 100 slices (one slice from each potato) of raw and 100 slices (one slice from each potato) of processed potatoes of each cultivar, as follows:

- 100 slices of raw potato tuber ‘Colomba’,
- 100 slices of raw potato tuber ‘Irga’,
- 100 slices of raw potato tuber ‘Riviera’,
- 100 slices of processed potato tuber ‘Colomba’,
- 100 slices of processed potato tuber ‘Irga’,
- 100 slices of processed potato tuber ‘Riviera’.

Image analysis

The image acquisition system consisted of a flatbed scanner and a computer. The imaging device (scanner) was connected to the computer. The slice images of raw and processed potatoes were acquired using the Epson Perfection flatbed scanner. The slices were scanned against a black background at the 800 dpi resolution. The obtained images were saved in TIFF format. The slice images of potatoes were subjected to analysis with the use of the Mazda software (Łódź University of Technology, Institute of Electronics, Poland) [12]. The available version of the Mazda software allowed to convert the images to individual color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V* and *S*. The original images and images from *B*, *b* and *Z* color channels of slices of potatoes ‘Colomba’, ‘Irga’ and ‘Riviera’ are presented in Fig. 1 for raw potatoes and in Fig. 2 for processed potatoes. The images from *B*, *b* and *Z* color channels are presented in figures because the differences

Fig. 1 The original images and images from different color channels of slices of raw potatoes

between the slices of individual cultivars were most visible in the case of images from these channels and the discriminative models built using textures selected from these color channels provided the highest accuracies.

The regions of interest (ROIs) for images were determined. One ROI included one whole slice of potato (flesh without skin). The black background of the slice images facilitated the determination of ROIs. The ROI defined as a set of pixels of slice was clearly separated from the black background by segmenting the image into brightness regions using the manually determined brightness threshold. Thus, the ROI including a lighter slice than the black background was determined. For each ROI in each color channel about 200 textures of the outer surface of the image based on the co-occurrence matrix, run-length matrix, histogram, gradient map, Haar wavelet transform and autoregressive model were calculated [12].

Statistical analysis

The cultivar discrimination of potatoes was carried out using the WEKA 3.9 application (Machine Learning Group, University of Waikato) [13]. A discriminant analysis was performed separately for raw and boiled potatoes and the results were compared to determine the effect of processing on the discriminative power of the textural features of the outer surface of slice images for distinguishing the potato cultivars. The selection of attributes (textures with the highest discriminative power) was carried out with the use of the Best First with the CFS (correlation-based feature selection algorithm) subset evaluator. The analysis was carried out in several steps. In the first step, the discrimination was performed using models including a set of textures selected from all color channels (R , G , B , L , a , b , X , Y , Z , U , V , S). In

Fig. 2 The original images and images from different color channels of slices of processed potatoes

the next step, the analysis was carried out for RGB, Lab and XYZ color spaces. In this case, the texture parameters were selected separately for each color space. The following step included building discriminative models for individual channels from each color space (channels *R*, *G*, *B* for RGB color space; channels *L*, *a*, *b* for Lab color space; channels *X*, *Y*, *Z*, for XYZ color space). The results for *B*, *b*, and *Z* color channels were included in this paper because the classification accuracies based on textures selected from these color channels were the highest. Additionally, the discriminative model was built for a set combining the textures selected from three color channels (*B*, *b* and *Z*). The discriminant analysis was carried out using tenfold cross-validation mode and the classifiers from Bayes (Bayes Net), functions (multilayer perceptron) and Lazy (IBk) groups [14].

Results

In the case of raw and boiled potatoes, the accuracies of cultivar discrimination were computed based on confusion matrices. The results were presented in the form of confusion matrices for three potato cultivars containing a number of cases that were correctly and incorrectly classified with an indication of average accuracy (%). Additionally, discrimination accuracies (%) for pairs of potato cultivars: 'Irga' vs. 'Colomba', 'Irga' vs. 'Riviera', 'Colomba' vs. 'Riviera' were determined. The raw potatoes of three cultivars 'Colomba', 'Irga' and 'Riviera' were discriminated with high accuracy reaching 94.33% for the Bayes Net and IBk classifiers based on discriminative models including a set of textures selected from all color channels (*R*, *G*, *B*,

Table 1 The confusion matrix and discrimination accuracy of raw potato slices of different cultivars based on a set of textures selected from all color channels (*R, G, B, L, a, b, X, Y, Z, U, V, S*)

Classifier	Predicted class (number of cases)			Actual class	Average accuracy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
Bayes (Bayes Net)	92	0	8	‘Colomba’	94.33	100	100	91.50
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Functions (Multilayer perceptron)	90	0	10	‘Colomba’	93.67	100	100	90.50
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Lazy (IBk)	90	0	10	‘Colomba’	94.33	100	100	91.50
	0	100	0	‘Irga’				
	7	0	93	‘Riviera’				

L, a, b, X, Y, Z, U, V, S) of slice images (Table 1). In the case of the Bayes Net classifier, all cases of ‘Irga’ were correctly included in the class ‘Irga’. Among the cases of potatoes ‘Colomba’, 92 cases were correctly classified as ‘Colomba’, and eight cases were incorrectly included in the class ‘Riviera’. For potatoes ‘Riviera’, 91 cases were correctly classified as ‘Riviera’ and nine cases—incorrectly as ‘Colomba’. In the case of the IBk classifier, also all cases of ‘Irga’ were correctly classified. For ‘Colomba’, 90 cases were correctly classified as ‘Colomba’, and ten cases as ‘Riviera’. For ‘Riviera’, 93 cases were correctly included in the class ‘Riviera’ and seven cases—incorrectly in the class ‘Colomba’. Considering the accuracy of discrimination performed for pairs of potato cultivars, ‘Irga’ vs. ‘Colomba’, as well as ‘Irga’ vs. ‘Riviera’ were distinguished with 100% correctness. Potatoes ‘Colomba’ vs. ‘Riviera’ were discriminated with the accuracy of up to 91.50% (Bayes Net and IBk classifiers).

In the case of discriminative models built for color spaces, the highest accuracy of discrimination of raw potato ‘Colomba’, ‘Irga’ and ‘Riviera’ was equal to 94% for Lab color space (Bayes Net classifier) (100 cases of ‘Irga’ were correctly classified; 91 cases belonging to ‘Colomba’ were correctly included in the class ‘Colomba’ and nine cases in the class ‘Riviera’; 91 cases of ‘Riviera’ were correctly classified as ‘Riviera’ and nine cases—as ‘Colomba’) and XYZ color space (Multilayer Perceptron classifier) (100 cases belonging to ‘Irga’ were correctly classified; 93 cases of ‘Colomba’ were classified as ‘Colomba’ and seven cases—as ‘Riviera’; 89 cases of ‘Riviera’ were classified as ‘Riviera’ and 11 cases—as ‘Colomba’) (Table 2). In the case of textures selected from RGB color space, the average accuracy for discrimination of three potato cultivars reached 91.33% (multilayer perceptron). All cases of ‘Irga’ were correctly classified. Potatoes ‘Colomba’ were correctly classified in

86% (86 cases from 100) and ‘Riviera’—in 88% (88 cases from 100). For pairs of cultivars, ‘Irga’ vs. ‘Colomba’ and ‘Irga’ vs. ‘Riviera’ were characterized by 100% discrimination accuracy in the case of all color spaces and all classifiers. For ‘Colomba’ vs. ‘Riviera’, the correctness reached 91% (Lab color space, Bayes Net classifier and XYZ color space, Multilayer Perceptron).

For discriminative models built based on textures selected from color channels (Table 3), the accuracies were lower than for color spaces (Table 2). In the case of individual color channels, the highest average discrimination accuracy for three potato cultivars was equal to 92% and was observed for channel *b* and the Bayes Net classifier, from which all slice images of potatoes ‘Irga’ were correctly classified, 87 cases of ‘Colomba’ were classified as ‘Colomba’ and 13 cases as ‘Riviera’, and 89 cases of ‘Riviera’ were correctly classified and the remaining 11 cases were classified as ‘Colomba’ (Table 3). Both, channels *B* and *Z* provided the average correctness of up to 86% for multilayer perceptron. The potatoes ‘Irga’ were completely correctly classified. The lowest accuracies were observed for potatoes ‘Riviera’. In the case of discriminative models built based on textures selected from color channel *B*, only 78 from 100 cases of ‘Riviera’ were correctly classified and for color channel *Z*—77 from 100 cases of ‘Riviera’ were included in class ‘Riviera’. Combining the selected textures from channels *B, b* and *Z* resulted in an increase in discrimination accuracy to 92.33% for three cultivars and 88.5% for ‘Colomba’ vs. ‘Riviera’. In the case of discrimination of ‘Irga’ vs. ‘Colomba’ and ‘Irga’ vs. ‘Riviera’, the accuracy of 100% was obtained (Table 3).

Very high accuracies of discrimination of processed (boiled) potatoes ‘Colomba’, ‘Irga’ and ‘Riviera’ were observed using models built based on a set of textures selected from all color channels (*R, G, B, L, a, b, X, Y, Z, U, V, S*) of slice images (Table 1). In the case of the Bayes Net classifier, all cases of ‘Irga’ were correctly included in the class ‘Irga’. Among the cases of potatoes ‘Colomba’, 92 cases were correctly classified as ‘Colomba’, and eight cases were incorrectly included in the class ‘Riviera’. For potatoes ‘Riviera’, 91 cases were correctly classified as ‘Riviera’ and nine cases—incorrectly as ‘Colomba’. In the case of the IBk classifier, also all cases of ‘Irga’ were correctly classified. For ‘Colomba’, 90 cases were correctly classified as ‘Colomba’, and ten cases as ‘Riviera’. For ‘Riviera’, 93 cases were correctly included in the class ‘Riviera’ and seven cases—incorrectly in the class ‘Colomba’. Considering the accuracy of discrimination performed for pairs of potato cultivars, ‘Irga’ vs. ‘Colomba’, as well as ‘Irga’ vs. ‘Riviera’ were distinguished with 100% correctness. Potatoes ‘Colomba’ vs. ‘Riviera’ were discriminated with the accuracy of up to 91.50% (Bayes Net and IBk classifiers).

Table 2 The confusion matrix and discrimination accuracy of raw potato slices of different cultivars based on textures selected from RGB, Lab and XYZ color spaces

Classifier	Predicted class (number of cases)			Actual class	Average accu- racy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
RGB color space								
Bayes (Bayes Net)	80	0	20	‘Colomba’	86.67	100	100	80
	0	100	0	‘Irga’				
	20	0	80	‘Riviera’				
Functions (Multilayer perceptron)	86	0	14	‘Colomba’	91.33	100	100	87
	0	100	0	‘Irga’				
	12	0	88	‘Riviera’				
Lazy (IBk)	82	0	18	‘Colomba’	87.33	100	100	81
	0	100	0	‘Irga’				
	20	0	80	‘Riviera’				
Lab color space								
Bayes (Bayes Net)	91	0	9	‘Colomba’	94	100	100	91
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Functions (Multilayer perceptron)	88	0	12	‘Colomba’	92.67	100	100	89
	0	100	0	‘Irga’				
	10	0	90	‘Riviera’				
Lazy (IBk)	87	0	13	‘Colomba’	92	100	100	88
	0	100	0	‘Irga’				
	11	0	89	‘Riviera’				
XYZ color space								
Bayes (Bayes Net)	84	0	16	‘Colomba’	89.67	100	100	84.50
	0	100	0	‘Irga’				
	15	0	85	‘Riviera’				
Functions (Multilayer perceptron)	93	0	7	‘Colomba’	94	100	100	91
	0	100	0	‘Irga’				
	11	0	89	‘Riviera’				
Lazy (IBk)	85	0	15	‘Colomba’	88.33	100	100	82.50
	0	100	0	‘Irga’				
	20	0	80	‘Riviera’				

V, S). The accuracies were equal to 98.67% for the Multilayer Perceptron classifier, 98.33% for Bayes Net and 97.67% for IBk (Table 4). These results are higher than for discriminative models based on textures selected from all color channels built for three cultivars (‘Colomba’, ‘Irga’ and ‘Riviera’) of raw potatoes, for which the correctness reached 94.33% (Table 1). In the case of boiled potato, for the multilayer perceptron and Bayes Net classifiers, 100 cases from the actual class ‘Colomba’ were correctly included in predicted class ‘Colomba’ and 100 cases from class ‘Irga’ were correctly classified as ‘Irga’. A discriminant analysis of pairs of potato cultivars revealed the complete distinction of potatoes ‘Irga’ from ‘Colomba’ and ‘Irga’ from ‘Riviera’ for all classifiers. Potatoes ‘Colomba’ and ‘Riviera’ were discriminated with

an accuracy of up to 98% (multilayer perceptron), for which slices of ‘Colomba’ were correctly classified in 100%, and 96 cases belonging to cultivars ‘Riviera’ were included in class ‘Riviera’ and four cases—in class ‘Colomba’.

Also, high discrimination accuracies for boiled potato of three cultivars were obtained for color spaces (Table 5). The highest accuracy was equal to 98% for RGB color space and the Multilayer Perceptron classifier. Discriminative models built based on textures selected from Lab color space provided an accuracy of up to 97% (Bayes Net), and for models built using textures from XYZ color space, the correctness was equal up to 97.67% (multilayer perceptron). In the case of all color spaces and all classifiers, the potatoes ‘Irga’ were completely distinguished from other potato cultivars

Table 3 The confusion matrix and discrimination accuracy of raw potato slices of different cultivars based on textures selected from color channels *B*, *b*, *Z* and combined set of textures selected from channels *B*, *b* and *Z*

Classifier	Predicted class (number of cases)			Actual class	Average accu- racy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
Color channel <i>B</i>								
Bayes (Bayes Net)	75	0	25	‘Colomba’	82.67	100	100	74
	0	100	0	‘Irga’				
	27	0	73	‘Riviera’				
Functions (Multilayer perceptron)	80	0	20	‘Colomba’	86	100	100	79
	0	100	0	‘Irga’				
	22	0	78	‘Riviera’				
Lazy (IBk)	75	0	25	‘Colomba’	82	100	100	73
	0	100	0	‘Irga’				
	29	0	71	‘Riviera’				
Color channel <i>b</i>								
Bayes (Bayes Net)	87	0	13	‘Colomba’	92	100	100	88
	0	100	0	‘Irga’				
	11	0	89	‘Riviera’				
Functions (Multilayer perceptron)	88	0	12	‘Colomba’	90.33	100	100	85.50
	0	100	0	‘Irga’				
	17	0	83	‘Riviera’				
Lazy (IBk)	81	0	19	‘Colomba’	88.67	100	100	83
	0	100	0	‘Irga’				
	15	0	85	‘Riviera’				
Color channel <i>Z</i>								
Bayes (Bayes Net)	72	0	28	‘Colomba’	84.33	100	100	76.50
	0	100	0	‘Irga’				
	19	0	81	‘Riviera’				
Functions (Multilayer perceptron)	81	0	19	‘Colomba’	86	100	100	79
	0	100	0	‘Irga’				
	23	0	77	‘Riviera’				
Lazy (IBk)	73	0	27	‘Colomba’	83.67	100	100	75.50
	0	100	0	‘Irga’				
	22	0	78	‘Riviera’				
Color channels <i>B</i> , <i>b</i> , <i>Z</i>								
Bayes (Bayes Net)	87	0	13	‘Colomba’	91.33	100	100	87
	0	100	0	‘Irga’				
	13	0	87	‘Riviera’				
Functions (Multilayer perceptron)	89	0	11	‘Colomba’	92.33	100	100	88.5
	0	100	0	‘Irga’				
	12	0	88	‘Riviera’				
Lazy (IBk)	87	0	13	‘Colomba’	90.67	100	100	86
	0	100	0	‘Irga’				
	15	0	85	‘Riviera’				

(‘Colomba’, ‘Riviera’). The pair of cultivars ‘Colomba’ and ‘Riviera’ was discriminated with the correctness of up to 97% in the case of RGB color space and multilayer perceptron classifier.

In the case of discriminant analysis performed based on textures selected for individual color channels (Table 6), the highest accuracy of 95.33% was observed for color channel *b* and Multilayer Perceptron classifier. The slice images of

Table 4 The confusion matrix and discrimination accuracy of boiled potato slices of different cultivars based on a set of textures selected from all color channels (*R, G, B, L, a, b, X, Y, Z, U, V, S*)

Classifier	Predicted class (number of cases)			Actual class	Average accuracy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
Bayes (Bayes Net)	100	0	0	‘Colomba’	98.33	100	100	97.50
	0	100	0	‘Irga’				
	5	0	95	‘Riviera’				
Functions (Multilayer perceptron)	100	0	0	‘Colomba’	98.67	100	100	98
	0	100	0	‘Irga’				
	4	0	96	‘Riviera’				
Lazy (IBk)	96	0	4	‘Colomba’	97.67	100	100	96.50
	0	100	0	‘Irga’				
	3	0	97	‘Riviera’				

potatoes ‘Irga’ were characterized by the 100% accuracy of discrimination. Potatoes ‘Colomba’ were correctly classified in 95% (95 cases of ‘Colomba’ were classified as ‘Colomba’, and five cases as ‘Riviera’). Potatoes belonging to cultivar ‘Riviera’ were classified with an accuracy of 91% (91 cases of ‘Riviera’ were classified as ‘Riviera’, and nine cases as ‘Colomba’). For comparison of pairs of potato cultivars, the accuracy of 100% was noted between ‘Irga’ and ‘Colomba’, as well as ‘Irga’ and ‘Riviera’, and the accuracy of up to 93% (color channel *b*, Multilayer Perceptron) was determined for ‘Colomba’ vs. ‘Riviera’. The increase in the correctness of discrimination was achieved by combining the textures selected from channels *B, b*, and *Z* (Table 6). The accuracy reached 96.67% for the Bayes Net classifier, from which potatoes ‘Irga’ were correctly classified in 100% cases, and ‘Colomba’, as well as ‘Riviera’—in 95% cases.

Discussion

The obtained results revealed that the boiling resulted in an increase in cultivar discrimination of potato slices. In the case of the discrimination performed using models built for a set of textures selected from all color channels (*R, G, B, L, a, b, X, Y, Z, U, V, S*), for color spaces and color channels, the accuracies were the higher for the boiled than raw potatoes. Additionally, the usefulness of textures of the outer surface of slice images of potato for cultivar discrimination of potato tuber was proved. Based on literature data, research on the image analysis of potatoes focused mainly on whole tubers. Machine vision combined with artificial neural networks was used by Azizi et al. [3] for the discrimination of ten potato cultivars based on color, morphological and textural features of whole tubers. The obtained accuracy ranged from 40 to 93.3% depending on the cultivar for

stepwise discriminant analysis (DA). Using neural network analysis, the correctness increased up to 100%. Mercurio and Hernandez [15] applied image processing combined with a convolutional neural network with the use of size and skin properties to classify sweet potatoes belonging to five cultivars achieving an average accuracy of 96.33% (from 95 to 98.33% for individual cultivars). Cultivar identification of sweet potato using NIR hyperspectral imaging and FT-MIR microspectroscopy revealed that two cultivars may be distinguished with 100% correctness [16]. Identification of two cultivars of potato tubers using neural image analysis based on aspect factors, geometric and color features was successfully performed with the quality of testing of above 0.99 by Przybył et al. [17]. In addition to cultivar identification, image analysis based on morphological features was also applied for the detection of potatoes with an irregular shape [18], the detection of potato tubers with damages, diseases and defects [19–22], the prediction of shape parameters and mass of normal and deformed potatoes [7], the classification of the potato tubers based on moisture content [23], the prediction of changes in moisture content and color of sweet potato during drying [24], the prediction of texture and color in cooked and cooked freeze-dried rehydrated potatoes [25]. The own studies regarding the evaluation of the effect of boiling on cultivar discrimination of potato based on the textural features of the outer surface of slice images completed the literature data on image analysis of potato. The results demonstrated the usefulness of slices in addition to whole tubers for potato cultivar discrimination. Further research may include the development of models based on texture parameters of slices for cultivar discrimination of potatoes subjected to other methods of processing. Further studies can also be intended to evaluate the robustness of classifiers by examining samples of the same potato cultivars collected from different regions and different growing seasons.

Table 5 The confusion matrix and discrimination accuracy of boiled potato slices of different cultivars based on textures selected from RGB, Lab and XYZ color spaces

Classifier	Predicted class (number of cases)			Actual class	Average accu- racy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
RGB color space								
Bayes (Bayes Net)	88	0	12	‘Colomba’	92.67	100	100	89
	0	100	0	‘Irga’				
	10	0	90	‘Riviera’				
Functions (Multilayer perceptron)	97	0	3	‘Colomba’	98	100	100	97
	0	100	0	‘Irga’				
	3	0	97	‘Riviera’				
Lazy (IBk)	84	0	16	‘Colomba’	91.67	100	100	87.50
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Lab color space								
Bayes (Bayes Net)	96	0	4	‘Colomba’	97	100	100	95.50
	0	100	0	‘Irga’				
	5	0	95	‘Riviera’				
Functions (Multilayer perceptron)	92	0	8	‘Colomba’	94.67	100	100	92
	0	100	0	‘Irga’				
	8	0	92	‘Riviera’				
Lazy (IBk)	96	0	4	‘Colomba’	96.33	100	100	94.50
	0	100	0	‘Irga’				
	7	0	93	‘Riviera’				
XYZ color space								
Bayes (Bayes Net)	88	0	12	‘Colomba’	90.67	100	100	86
	0	100	0	‘Irga’				
	16	0	84	‘Riviera’				
Functions (Multilayer perceptron)	99	0	1	‘Colomba’	97.67	100	100	96.50
	0	100	0	‘Irga’				
	6	0	94	‘Riviera’				
Lazy (IBk)	84	0	16	‘Colomba’	91.33	100	100	87
	0	100	0	‘Irga’				
	10	0	90	‘Riviera’				

Conclusions

The obtained results revealed that the analysis of slice images acquired with the use of a flatbed scanner allows for cultivar discrimination of raw and processed potatoes in an objective, inexpensive and fast manner with high accuracy. The discriminative models built based on a set of textures selected from color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, *S* produced the highest results, 94.33% for raw potatoes ‘Colomba’, ‘Irga’ and ‘Riviera’, and 98.67% for processed potatoes. These models revealed that both in the case of raw and processed potatoes, 100% discrimination accuracy was achieved between potatoes ‘Irga’ vs. ‘Colomba’ and ‘Irga’ vs. ‘Riviera’. Raw potatoes ‘Colomba’ vs. ‘Riviera’

were distinguished with the correctness of up to 91.50% and processed potatoes ‘Colomba’ vs. ‘Riviera’—of up to 98%. The results may be applied in practice to verify the cultivar authenticity of sliced potato, detect adulteration, and avoid cultivar mixing in potato processing. The identification of cultivar may be a significant factor at each step in the agricultural production chain. It can be important, among others, for farmers, growers, cultivar registration agencies, plant breeders, processors, bulk handlers, marketers, end-users. Further research can be intended to develop the discriminative models for potatoes subjected to other methods of processing and can relate to examining potato samples collected from different growing seasons and several regions.

Table 6 The confusion matrix and discrimination accuracy of boiled potato slices of different cultivars based on textures selected from color channels *B*, *b*, *Z* and combined set of textures selected from channels *B*, *b* and *Z*

Classifier	Predicted class (number of cases)			Actual class	Average accu- racy (%)	Accuracy (%) for pairs of potato cultivars		
	‘Colomba’	‘Irga’	‘Riviera’			‘Irga’ vs. ‘Colomba’	‘Irga’ vs. ‘Riviera’	‘Colomba’ vs. ‘Riviera’
Color channel <i>B</i>								
Bayes (Bayes Net)	83	0	17	‘Colomba’	91	100	100	86.50
	0	100	0	‘Irga’				
	10	0	90	‘Riviera’				
Functions (Multilayer perceptron)	83	0	17	‘Colomba’	91.33	100	100	87
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Lazy (IBk)	77	0	23	‘Colomba’	86	100	100	79
	0	100	0	‘Irga’				
	19	0	81	‘Riviera’				
Color channel <i>b</i>								
Bayes (Bayes Net)	80	0	20	‘Colomba’	90.33	100	100	85.50
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Functions (Multilayer perceptron)	95	0	5	‘Colomba’	95.33	100	100	93
	0	100	0	‘Irga’				
	9	0	91	‘Riviera’				
Lazy (IBk)	85	0	15	‘Colomba’	91	100	100	86.50
	0	100	0	‘Irga’				
	12	0	88	‘Riviera’				
Color channel <i>Z</i>								
Bayes (Bayes Net)	86	0	14	‘Colomba’	90	100	100	85
	0	100	0	‘Irga’				
	16	0	84	‘Riviera’				
Functions (Multilayer perceptron)	92	0	8	‘Colomba’	94.67	100	100	92
	0	100	0	‘Irga’				
	8	0	92	‘Riviera’				
Lazy (IBk)	79	0	21	‘Colomba’	87.67	100	100	81.50
	0	100	0	‘Irga’				
	16	0	84	‘Riviera’				
Color channels <i>B</i> , <i>b</i> , <i>Z</i>								
Bayes (Bayes Net)	95	0	5	‘Colomba’	96.67	100	100	95
	0	100	0	‘Irga’				
	5	0	95	‘Riviera’				
Functions (Multilayer perceptron)	92	0	8	‘Colomba’	95.33	100	100	93
	0	100	0	‘Irga’				
	6	0	94	‘Riviera’				
Lazy (IBk)	95	0	5	‘Colomba’	96.33	100	100	94.50
	0	100	0	‘Irga’				
	6	0	94	‘Riviera’				

Compliance with ethical standards

Conflict of interest The author declares no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbasi KS, Qayyum A, Mehmood A, Mahmood T, Khan SU, Liaquat M, Sohail A, Ahmad A (2019) Analysis of selective potato varieties and their functional assessment. *Food Sci Technol* 39(2):308–314
- Liu J, Wen Ch, Wang M, Wang S, Dong N, Lei Z, Lin S, Zhu B (2020) Enhancing the hardness of potato slices after boiling by combined treatment with lactic acid and calcium chloride: mechanism and optimization. *Food Chem* 308(124832):1–9
- Azizi A, Abbaspour-Gilandeh Y, Nooshyar M, Afkari-Sayah A (2016) Identifying potato varieties using machine vision and artificial neural networks. *Int J Food Prop* 19:618–635
- Morey R, Ermolenkov A, Payne WZ, Scheuring DC, Koym JW, Vales MI, Kurouski D (2020) Non-invasive identification of potato varieties and prediction of the origin of tuber cultivation using spatially offset Raman spectroscopy. *Anal Bioanal Chem* 412:4585–4594
- Thybo AK, Szczypinski PM, Karlsson AH, Dønstrup S, Stødkilde-Jørgensen HS, Andersen HJ (2004) Prediction of sensory texture quality attributes of cooked potatoes by NMR-imaging (MRI) of raw potatoes in combination with different image analysis methods. *J Food Eng* 61:91–100
- Al Riza DF, Suzuki T, Ogawa Y, Kondo N (2017) Diffuse reflectance characteristic of potato surface for external defects discrimination. *Postharvest Biol Technol* 133:12–19
- Su Q, Kondo N, Li M, Sun H, Al Riza DF (2017) Potato feature prediction based on machine vision and 3D model rebuilding. *Comput Electron Agric* 137:41–51
- Sanchez PDC, Hashim N, Shamsudin R, Nor MZM (2020) Applications of imaging and spectroscopy techniques for non-destructive quality evaluation of potatoes and sweet potatoes: a review. *Trends Food Sci Technol* 96:208–221
- ElMasry G, Cubero S, Moltó E, Blasco J (2012) In-line sorting of irregular potatoes by using automated computer-based machine vision system. *J Food Eng* 112:60–68
- Nováková A, Šimáčková K, Bárta J, Čurn V (2009) Potato variety identification by molecular markers based on retrotransposon analyses. *Czech J Genet Plant Breed* 45(1):1–10
- Bărăscu N, Donescu V, Hermeziu M, Cioloca M, Nițu S, Ștefan FM (2019) Evaluation of yield quality and culinary characteristics of different potato varieties. *J Hortic For Biotechnol* 23(1):14–21
- Szczypinski PM, Strzelecki M, Materka A, Klepaczko A (2009) MaZda- a software package for image texture analysis. *Comput Methods Prog Biomed* 94(1):66–76
- Bouckaert RR, Frank E, Hall M, Kirkby R, Reutemann P, Seewald A, Scuse D (2016) WEKA manual for version 3-9-1. The University of Waikato, Hamilton
- Witten IH, Frank E (2005) Data mining. In: Practical machine learning tools and techniques, 2nd edn. Elsevier, San Francisco, CA, p 525
- Mercurio DI, Hernandez AA (2019) Classification of sweet potato variety using convolutional neural network. In: 2019 IEEE 9th international conference on system engineering and technology (ICSET), pp 120–125
- Su W-H, Bakalis S, Sun D-W (2019) Chemometrics in tandem with near infrared (NIR) hyperspectral imaging and Fourier transform mid infrared (FT-MIR) microspectroscopy for variety identification and cooking loss determination of sweet potato. *Biosyst Eng* 180:70–86
- Przybył K, Górna K, Wojcieszak D, Czekala W, Ludwiczak A, Przybylak A, Boniecki P, Koszela K, Zaborowicz M, Janczak D, Lewicki A (2015) The recognition of potato varieties using neural image analysis method. In: Falco CM, Jiang X (eds) Seventh international conference on digital image processing: ICDIP 2015, Proc of SPIE, vol 9631, p 963116
- Azizi A, Abbaspour-Gilandeh Y (2016) Identifying irregular potatoes by developing an intelligent algorithm based on image processing. *Tarim Bilimleri Dergisi J Agric Sci* 22:32–41
- Przybylak A, Kozłowski R, Osuch E, Osuch A, Rybacki P, Przygodziński P (2020) Quality evaluation of potato tubers using neural image analysis method. *Agriculture* 10(112):1–11
- Oppenheim D, Shani G, Erlich O, Tsror L (2019) Using deep learning for image-based potato tuber disease detection. *Phytopathology* 109:1083–1087
- Zhang W, Zhu Q, Huang M, Guo Y, Qin J (2019) Detection and classification of potato defects using multispectral imaging system based on single shot method. *Food Anal Methods* 12:2920–2929
- Razmjoo N, Mousavi BS, Soleymani F (2012) A real-time mathematical computer method for potato inspection using machine vision. *Comput Math Appl* 63:268–279
- Ebrahimi E, Mollazade K, Arefi A (2012) An expert system for classification of potato tubers using image processing and artificial neural networks. *Int J Food Eng* 8(4):1–18
- Onwude DI, Hashim N, Abdan K, Janius R, Chen G (2018) Combination of computer vision and backscattering imaging for predicting the moisture content and colour changes of sweet potato (*Ipomoea batatas* L.) during drying. *Comput Electron Agric* 150:178–187
- Pieniazek F, Messina V (2017) Texture and color analysis of freeze-dried potato (cv. *Spunta*) using instrumental and image analysis techniques. *Int J Food Prop* 20(6):1422–1431

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.