



# Flesh of pumpkin from ecological farming as part of fruit suitable for non-destructive cultivar classification using computer vision

Ewa Ropelewska<sup>1</sup> · Wioletta Popińska<sup>1</sup> · Kadir Sabanci<sup>2</sup> · Muhammet Fatih Aslan<sup>2</sup>

Received: 19 July 2021 / Accepted: 11 December 2021 / Published online: 23 December 2021  
© The Author(s) 2021

## Abstract

The aim of this study was to build the discriminative models for distinguishing the different cultivars of flesh of pumpkin ‘Bambino’, ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’ based on selected textures of the outer surface of images of cubes. The novelty of research involved the use of about 2000 different textures for one image. The highest total accuracy (98%) of discrimination of pumpkin ‘Bambino’, ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’ was determined for models built based on textures selected from the color space Lab and the IBk classifier and some of the individual cultivars were classified with the correctness of 100%. The total accuracy of up to 96% was observed for color space RGB and 97.5% for color space XYZ. In the case of color channels, the total accuracies reached 91% for channel *b*, 89.5% for channel *X*, 89% for channel *Z*.

**Keywords** Pumpkin flesh · Cultivars · Digital image analysis · Discriminative classifiers

## Introduction

Pumpkin (squash) belongs to the genus *Cucurbita* and the family Cucurbitaceae. Especially economically important are *Cucurbita pepo* L., *Cucurbita maxima* Duchesne and *Cucurbita moschata* Duchesne ex Poiret which have high production and are cultivated worldwide [1]. America, Mexico and Peru are considered to be the primary center of origin of pumpkin [2]. The pumpkin is selective harvesting including recognition of fruit in the first stem, and then stem cutting and harvesting the fruit [3]. Pumpkin was characterized by high yields, high nutritive value and long storage life. It may be grown under a wide range of agro-climatic conditions. The ripe fruits have sweet yellow or orange flesh [4]. The fruits may be different in terms of color, size, weight and shape. The pumpkin fruits consist of edible flesh, skin and seed cavity. The pumpkin crops are seasonal and fresh fruits, which are sensitive to microbial spoilage, should be dried or frozen. The physicochemical characteristics of

pumpkin and its products may differ, among others, depending on the variety and conditions in the region of cultivation [5]. Additionally, the physicochemical properties may change during the growth and ripening of pumpkin [6]. The pumpkin may be considered as a functional food [7]. Due to the presence of phytonutrients, important functional components, e.g., dietary fibers, carotenoids, minerals, zeaxanthin, vitamins, ascorbic acids, linoleic acid, consuming pumpkin is beneficial to human health. The pumpkin can be consumed in various forms, including processed products, e.g., jelly, jam, marmalade, sauce, puree, candy, chutney, cookies, pickles, pies, halwa, weaning mix, beverages and pumpkin flour may be added to prepare instant noodles, bakery products, pasta, soups [1, 4]. Also seeds of pumpkin have high nutritional value and they may be used, for example, to produce cold pressed oils [7, 8].

Pumpkin flesh (pulp) includes phytochemicals with health-promoting action. The content of many compounds depends on the cultivar and species of pumpkin. The significant differences can be determined in the content of carotenoids, vitamin C, B1, tocopherols, folates, flavonols, phenolic acids [9]. Therefore, the selection of the cultivars with desired properties for consumption and processing can be important. Determination of the chemical compounds of pumpkin can be expensive and time-consuming. According to Sim et al. [10] also testing the cultivars based on phenotypes may be time-consuming and laborious. The authors

✉ Ewa Ropelewska  
ewa.ropelewska@inhort.pl

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

<sup>2</sup> Electrical and Electronics Engineering, Karamanoglu Mehmetbey University, Karaman, Turkey

[10] reported that the application of the molecular markers for pumpkin cultivar identification may be fast and useful.

In this study, the authors assumed that it is possible to distinguish the flesh of different pumpkin cultivars using an image analysis technique. It would provide an objective way to avoid mixing different cultivars of pumpkin indistinguishable by the naked eye. Additionally, the application of image processing may partially replace more expensive or more laborious methods of pumpkin cultivar identification, e.g., phenotypic or molecular. The aim of this study was to build the classification models based on texture parameters computed from digital images of the pumpkin flesh.

## Materials and methods

### Materials

The research material comprised of the pumpkin of four cultivars ‘Bambino’, ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’. The cultivars ‘Bambino’, ‘Uchiki Kuri’ and ‘Orange’ belonged to pumpkin species *Cucurbita maxima* Duchesne. Whereas cultivar ‘Butternut’ belonged to species *Cucurbita moschata* Duchesne ex Poiret. The samples were collected from ecological farming from Świętokrzyskie Voivodeship (Holy Cross Province), Poland. The soil in classes II and III and fertilizers of potassium sulfate and Bioilsa were used. In the case of pumpkin ‘Bambino’, the triticale was applied as forecrop, straw was left, the mulching was used. For ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’, the cabbage was forecrop, and cattle manure was used. The pumpkin was harvest after reaching consumer maturity. After harvesting the fruits were stored under room conditions (temperature about 18–20 °C) for two weeks. Then, the pumpkin flesh was cut into cubes (1 cm × 1 cm) using a knife. The pumpkin cubes were put into plastic boxes and were frozen at a temperature of – 29 °C. The samples were subjected to frozen storage. Afterwards the pumpkin cubes were thawed under room conditions. Excess water was removed from cubes using a paper towel and the pumpkin cubes were subjected to experiments using image analysis.

### Image analysis

The flesh of pumpkin fruit was subjected to image processing. Image acquisition of cubes of the flesh placed on the black background was performed using a flatbed scanner. One image included ten cubes of pumpkin belonging to one cultivar. For each cultivar, five images were obtained. A total of 50 cube images for each cultivar were acquired. Therefore, the image analysis was performed for two hundred cube images (50 cube images for each of the four cultivars). The images of pumpkin flesh were subjected to processing using

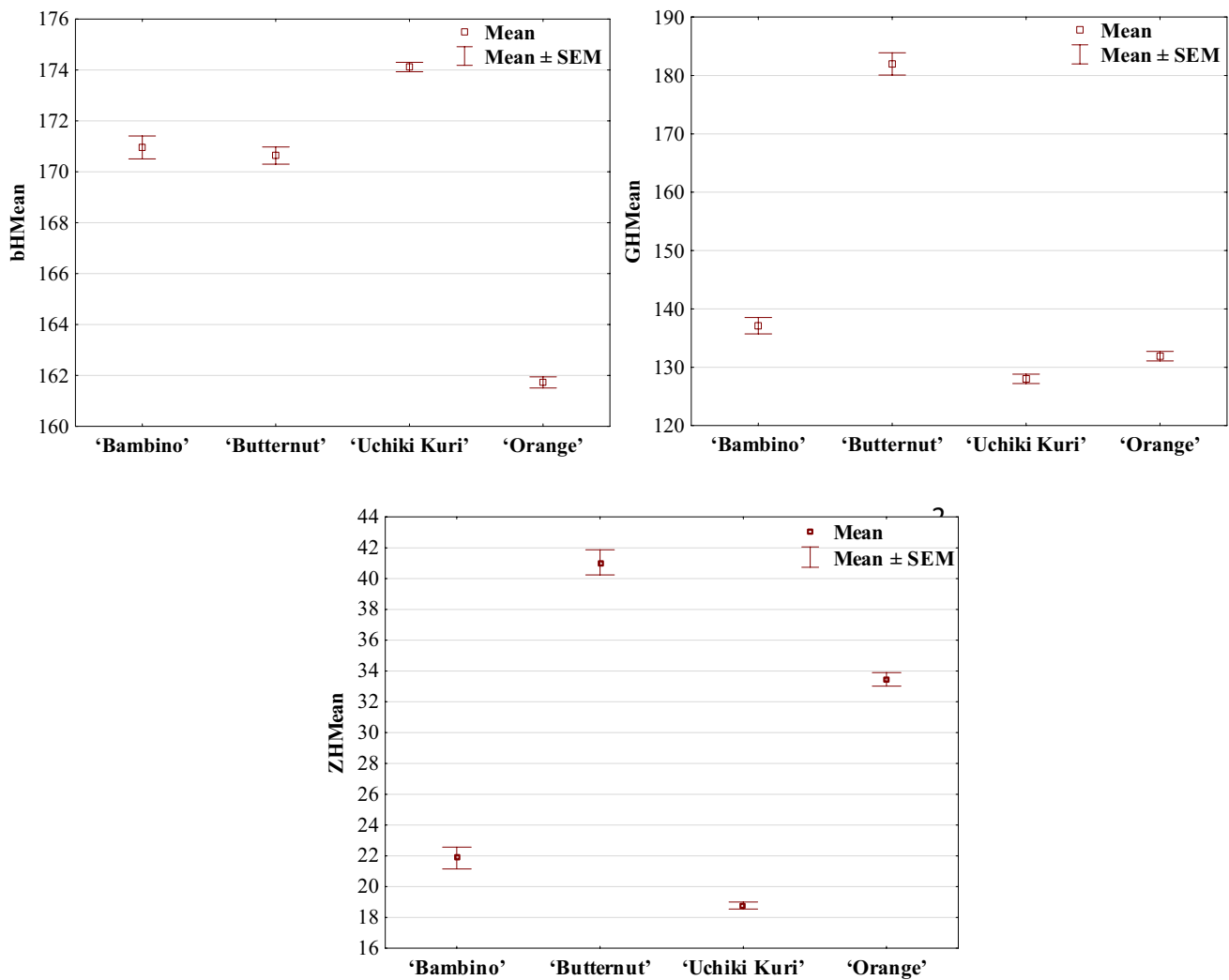
the Mazda software (Łódź University of Technology, Institute of Electronics, Poland) [11]. The obtained images were converted to individual color channels  $L, a, b, R, G, B, X, Y, Z$ . For each cube image in each color channel, the region of interest (ROI) as the whole scanned surface, was determined. For each ROI about 200 textures of the outer surface of the image based on the histogram, run-length matrix, co-occurrence matrix, gradient map and autoregressive model were calculated [11].

### Statistical analysis

The discriminant analysis was performed using the WEKA 3.9 application (Machine Learning Group, University of Waikato) [12] to develop the models for distinguishing the cultivar of pumpkin based on selected textures of the outer surface of images of cubes of flesh. The Best First with the CFS (Correlation-based Feature Selection algorithm) subset evaluator was used for the selection of textures with the highest discriminative power. The texture selection was performed separately for the sets including textures of images of pumpkin flesh of four cultivars ‘Bambino’, ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’ from color space Lab and color channels  $L, a, b$ , color space RGB and color channels  $R, G, B$ , color space XYZ and color channels  $X, Y, Z$ . The discriminative models were built separately based on the sets selected textures from color space Lab and channels  $L, a, b$ , color space RGB and channels  $R, G, B$ , color space XYZ and channels  $X, Y, Z$ . The analyzes were carried out using the of Multilayer Perceptron (group of Functions), IBk (Lazy), PART (Rules) and Random Forest (Decision Trees) classifiers and ten-fold cross-validation mode [13]. These procedures ensured the highest correctness of discrimination.

## Results and discussion

The differences in texture parameters of different cultivars of pumpkin flesh were observed. The mean values of selected textures from different color channels are presented in Fig. 1. The complete differentiation of pumpkin cultivars in terms of some textures is noticeable. It resulted in high discrimination accuracies. The accuracies (%) of cultivar discrimination of pumpkin of four cultivars ‘Bambino’, ‘Butternut’, ‘Uchiki Kuri’ and ‘Orange’ based on textures selected from color space Lab and channels  $L, a, b$  are presented in Table 1. The results were very high. The discriminative models built based on textures selected from the color space Lab provided a total accuracy of up to 98% for the IBk classifier. For the individual cultivars, the accuracies reached 100% for ‘Orange’, 98% for ‘Butternut’, ‘Uchiki Kuri’ and 96% for ‘Bambino’. The correctness of 100% for pumpkin ‘Orange’ was also observed in the case of Multilayer



**Fig. 1** The mean values and SEM (standard error of the mean) of selected textures (bHMean from color channel *b*, GHMean from color channel *G*, ZHMean from color channel *Z*) of pumpkin flesh

Perceptron (total accuracy of 97%) and Random Forest (total accuracy of 97.5%). It indicated that this cultivar was the most different from the others. In the case of the models built for textures selected from images from individual color channels, the highest total accuracy was determined for color channel *b* and was equal to 91% for the Random Forest classifier. The pumpkin 'Orange' was characterized by 100% accuracy. It means that all images of flesh of pumpkin 'Orange' were correctly classified as 'Orange'. The same situation (100% accuracy for 'Orange') was also found for the Multilayer Perceptron (total accuracy of 90.5%) and IBk (total accuracy of 89.5%) classifiers. The flesh pumpkin belonging to cultivar 'Orange' differed most from the others in terms of selected textures from the color channel *b*, which was reflected in the high correctness of discrimination of this cultivar for models built for color space Lab. The total accuracies of discrimination of pumpkin 'Bambino', 'Butternut',

'Uchiki Kuri' and 'Orange' for color channel *L* ranged from 84.5% for the IBk classifier to 88% for the Random Forest classifier, and for color channel *a* were in the range of 79% for the PART classifier to 85% for the Random Forest. None of the cultivars obtained 100% correctness of discrimination. However, the highest accuracies of up to 98% (PART and Random Forest for color channel *L* and IBk for color channel *a*) were determined for 'Butternut'.

The results of discriminative models built based on textures selected from color space RGB and color channels *R*, *G*, *B* are shown in Table 2. The discrimination accuracies for color space RGB were very high. The flesh of pumpkin 'Bambino', 'Butternut', 'Uchiki Kuri' and 'Orange' was distinguished with the total accuracy reaching 96% for the IBk and Random Forest classifiers. In the case of individual cultivars, the correctness of 100% was for 'Orange' for the Multilayer Perceptron, IBk and Random Forest classifiers, as

**Table 1** Discrimination accuracy of pumpkin flesh based on textures selected from color space Lab and channels *L*, *a*, *b*

Classifier	Accuracy (%)				Total
	'Bambino'	'Butternut'	'Uchiki Kuri'	'Orange'	
Color space Lab					
Multilayer perceptron (functions)	92	98	98	100	97
IBk (lazy)	96	98	98	100	98
PART (rules)	88	100	94	96	94.5
Random forest (decision trees)	94	96	100	100	97.5
Color channel <i>L</i>					
Multilayer perceptron (functions)	88	96	82	78	86
IBk (lazy)	88	96	80	74	84.5
PART (rules)	82	98	82	84	86.5
Random forest (decision trees)	86	98	80	88	88
Color channel <i>a</i>					
Multilayer perceptron (functions)	84	96	70	72	80.5
IBk (lazy)	76	98	68	78	80
PART (rules)	82	96	60	78	79
Random forest (decision trees)	84	96	78	82	85
Color channel <i>b</i>					
Multilayer perceptron (functions)	78	92	92	100	90.5
IBk (lazy)	82	88	88	100	89.5
PART (rules)	84	76	90	94	86
Random forest (decision trees)	80	88	96	100	91

**Table 2** Discrimination accuracy of pumpkin flesh based on textures selected from color space RGB and channels *R*, *G*, *B*

Classifier	Accuracy (%)				Total
	'Bambino'	'Butternut'	'Uchiki Kuri'	'Orange'	
Color space RGB					
Multilayer perceptron (functions)	92	90	96	100	94.5
IBk (lazy)	92	94	98	100	96
PART (rules)	84	94	94	96	92
Random forest (decision trees)	90	94	100	100	96
Color channel <i>R</i>					
Multilayer perceptron (functions)	84	58	74	54	67.5
IBk (lazy)	84	58	58	66	66.5
PART (rules)	86	50	52	66	63.5
Random forest (decision trees)	94	64	64	76	74.5
Color channel <i>G</i>					
Multilayer perceptron (functions)	80	90	82	86	84.5
IBk (lazy)	76	92	72	80	80
PART (rules)	78	92	80	72	80.5
Random forest (decision trees)	88	92	88	82	87.5
Color channel <i>B</i>					
Multilayer perceptron (functions)	80	68	96	80	81
IBk (lazy)	88	76	100	72	84
PART (rules)	78	72	94	78	80.5
Random forest (decision trees)	84	84	94	80	85.5

well as for 'Uchiki Kuri' for the Random Forest. The accuracies for models built based on textures selected from color channels *R*, *G*, *B* were lower than color space RGB. The

lowest results ranged from 63.5% (PART) to 74.5% (Random Forest) were obtained for color channels *R*. The color channel *R* turned out to be unsuitable for cultivar discrimination

of pumpkin flesh. For other color channels, the accuracies were higher and were equal to 80% (IBk)—87.5% (Random Forest) for color channel *G* and 80.5% (PART)—85.5% (Random Forest) for color channel *B*.

The accuracies of discrimination of pumpkin flesh based on textures selected from color space XYZ and channels *X*, *Y*, *Z* are presented in Table 3. The discriminative models built for textures selected from color space XYZ produced the highest total accuracy equal up to 97.5% for the Multilayer Perceptron classifier. For the individual cultivar, the result of 100% was observed only for pumpkin ‘Uchiki Kuri’ for the Multilayer Perceptron. The discrimination accuracies for color channel *X* were from 85% (IBk) to 89.5% (Random Forest). In the case of color channel *Y*, the correctness ranged from 80.5% (PART) to 86% (Random Forest). For color channel *Z*, the accuracies were in the range of 80.5 (PART) to 89% (Random Forest).

The results revealed the usefulness of digital image analysis based on texture parameters for the cultivar classification of the pumpkin flesh. In the available literature, there are data about the application of image processing in various studies of pumpkin. For example, Wittstruck et al. [14] used UAV (Unmanned Aerial Vehicle)-Based RGB imagery for the detection of single pumpkin fruit in the field, counting fruits and the prediction of fruit size and weight. Oblitas-Cruz et al. [15] and Oblitas et al. [16] applied image processing combined with neural networks for the discrimination of microstructural elements of the pumpkin tissue. Image

analysis was also used for the research related to the drying of pumpkin. Zenoozian et al. [17, 18] applied a computer vision system for the prediction of the surface color, percentage of shrinkage and Heywood shape factor of pumpkin cubes subjected to osmotic dehydration and hot-air drying. The analysis of X-ray pumpkin seed images was used for the evaluation of the internal morphology of seeds [19]. The above-mentioned examples for pumpkin indicated several different applications of image processing. The usefulness of imaging techniques was also confirmed in our research for a new direction for research aimed at pumpkin cultivar discrimination.

## Conclusion

The applied digital image processing proved to be effective for the classification of the pumpkin flesh with high probability. The developed cultivar discriminative models built based on texture parameters of pumpkin flesh images provided very satisfactory accuracy. The distinguishing of the pumpkin cultivars was the most correct in the case of model built for the textures selected from the color space Lab and the total accuracy was equal up to 98%, and for individual cultivars even reached 100%. In the case of models built for color channels, the results were the highest for color channel *b* and reached 91% for the discrimination of flesh of all four cultivars. The developed procedures may be used in the

**Table 3** Discrimination accuracy of pumpkin flesh based on textures selected from color space XYZ and channels *X*, *Y*, *Z*

Classifier	Accuracy (%)				Total
	‘Bambino’	‘Butternut’	‘Uchiki Kuri’	‘Orange’	
Color space XYZ					
Multilayer perceptron (functions)	96	96	100	98	97.5
IBk (lazy)	94	94	98	96	95.5
PART (rules)	80	94	94	90	89.5
Random forest (decision trees)	94	94	94	96	94.5
Color channel <i>X</i>					
Multilayer perceptron (functions)	92	94	84	82	88
IBk (lazy)	84	90	82	84	85
PART (rules)	84	92	88	86	87.5
Random forest (decision trees)	90	92	86	90	89.5
Color channel <i>Y</i>					
Multilayer perceptron (functions)	78	92	84	82	84
IBk (lazy)	76	92	72	84	81
PART (rules)	78	98	72	74	80.5
Random forest (decision trees)	84	96	80	84	86
Color channel <i>Z</i>					
Multilayer perceptron (functions)	82	82	94	92	87.5
IBk (lazy)	78	84	92	86	85
PART (rules)	70	82	96	74	80.5
Random forest (decision trees)	86	86	94	90	89

practice to distinguish the pumpkin belonging to different cultivars.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Compliance with ethics requirements** This article does not contain any studies with human or animal subjects.

**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

- Ahmad G, Khan AA (2019) Pumpkin: horticultural importance and its roles in various forms; a review. *Int J Hort Agric* 4(1):1–6
- Ramjan Md, Pandey AK, Raghavan M, Lian HN, Ansari MT (2018) Variations in pumpkin (*Cucurbita moschata* Duch. ex Poir.) landraces for earliness, yield and quality attributes in North East India. *Int J Curr Microbiol App Sci* 7(7):741–748
- Roshanianfard A, Noguchi N (2018) Characterization of pumpkin for a harvesting robot. *IFAC PapersOnLine* 51–17:23–30
- Dhiman AK, Sharma KD, Surekha A (2009) Functional constituents and processing of pumpkin: a review. *J Food Sci Technol* 46(5):411–417
- Adubofuor J, Amoah I, Agyekum PB (2016) Physicochemical properties of pumpkin fruit pulp and sensory evaluation of pumpkin-pineapple juice blends. *Am J Food Sci Tech* 4(4):89–96
- Sharma S, Rao TVR (2013) Nutritional quality characteristics of pumpkin fruit as revealed by its biochemical analysis. *Int Food Res J* 20:2309–2316
- Montesano D, Blasi F, Simonetti MS, Santini A, Cossignani L (2018) Chemical and nutritional characterization of seed oil from *Cucurbita maxima* L. (var. Berrettina) pumpkin. *Foods* 7:30
- Zdunić GM, Menković NR, Jadranin MB, Novaković MM, Šavikin KP, Živković JČ (2016) Phenolic compounds and carotenoids in pumpkin fruit and related traditional products. *Hem Ind* 70(4):429–433
- Kulczyński B, Gramza-Michałowska A (2019) The profile of secondary metabolites and other bioactive compounds in *Cucurbita pepo* L. and *Cucurbita moschata* pumpkin cultivars. *Molecules* 24:2945
- Sim S-C, Hong J-H, Kwon Y-S (2015) DNA profiling of commercial pumpkin cultivars using simple sequence repeat polymorphisms. *Hortic Environ Biotechnol* 56(6):811–820
- Szczypliński PM, Strzelecki M, Materka A, Klepaczko A (2009) MaZda—a software package for image texture analysis. *Comput Methods Programs 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. University of Waikato, Hamilton
- Witten IH, Frank E (2005) Data mining. in practical machine learning tools and techniques, 2nd edn. Elsevier, San Francisco
- Wittstruck L, Kühling I, Trautz D, Kohlbrecher M, Jarmer T (2021) UAV-based RGB imagery for Hokkaido pumpkin (*Cucurbita max.*) detection and yield estimation. *Sensors* 21:118
- Oblitas-Cruz JF, Castro-Silupu WM, Mayor-López L (2016) Effect of different combinations of size and shape parameters in the percentage error of classification of structural elements in vegetal tissue of the pumpkin *Cucurbita pepo* L. using probabilistic neural networks. *Rev Fac Ing* 78:30–37
- Oblitas J, Mejia J, De-la-Torre M, Avila-George H, Seguí Gil L, Mayor López L, Ibarz A, Castro W (2021) Classification of the microstructural elements of the vegetal tissue of the pumpkin (*Cucurbita pepo* L.) using convolutional neural networks. *Appl Sci* 11(1581):10–13
- Zenoozian MS, Devahastin S, Razavi MA, Shahidi F, Poreza HR (2007) Use of artificial neural network and image analysis to predict physical properties of osmotically dehydrated pumpkin. *Drying Technol* 26:132–144
- Zenoozian MS, Feng H, Razavi SMA, Shahidi F, Pourreza HR (2008) Image analysis and dynamic modeling of thin-layer drying of osmotically dehydrated pumpkin. *J Food Process Preserv* 32:88–102
- Silva PP, Freitas RA, Cícero SM, Marcos-Filho J, Nascimento WM (2014) Análise de imagens no estudo morfológico e fisiológico de sementes de abóbora [Image analysis to associate morphological and physiological characteristics in pumpkin seeds]. *Hortic Bras* 32:210–214

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