Acta Sci. Pol. Hortorum Cultus, online first, 1–12

https://czasopisma.up.lublin.pl/index.php/asphc ISSN 1644-0692 e-ISSN 2545-1405 https://doi.org/10.24326/asphc.2025.5370

ORIGINAL PAPER

Received: 8.04.2024
Accepted: 2.03.2025
First published online: 27.03.2025

NONDESTRUCTIVE DISCRIMINATION OF ADVANCED CLONES AND CULTIVARS OF STRAWBERRY USING AN INNOVATIVE APPROACH INVOLVING IMAGE ANALYSIS AND MACHINE LEARNING

Ewa Ropelewska[1] ✉, Agnieszka Masny[2]

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

[2] Horticultural Crop Breeding Department, The National Institute of Horticultural Research, Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland

ABSTRACT

Different clones and cultivars of strawberry can differ in morphological and chemical properties, as well as productivity, adaptation to cultivation conditions, and post-harvest quality during storage and processing. Due to differences in the quality of raw materials and final products depending on the strawberry clone/cultivar, correct distinguishing clones and cultivars is important for growers, consumers and processors. This study was aimed at distinguishing advanced clones and cultivars of strawberry using an innovative approach involving image processing and artificial intelligence. The raw material included the advanced clones and cultivars of strawberry, such as clone with the breeding code T-201457-16 (Grandarosa × Elsanta), clone T-201536-06 (Clery × Grandarosa), clone T-201567-01 (Patty × Panvik), as well as the cultivars Fibion, Grandarosa, and Markat. The fruit image acquisition was performed using a digital camera. As many as 2172 image parameters were extracted from the image of each fruit converted to different color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, and *S* and textures with the highest discriminative power were selected to develop models using various machine learning algorithms, such as Multilayer Perceptron, MultiClass Classifier, IBk, and LMT, Linear Discriminant, Quadratic SVM, Subspace Discriminant, and Wide Neural Network. The most accurate classifications were obtained for a model built using Subspace Discriminant (96.30%) and Multilayer Perceptron (95.83%). For the model developed using Subspace Discriminant, clone T-201567-01 and cultivar Markat were completely correctly classified with the highest accuracy of 100%. Whereas in the case of the model built using Multilayer Perceptron clone T-201567-01 was characterized by the highest classification metrics, such as Precision and F-measure equal to 0.983, MCC of 0.980, PRC Area and ROC Area of 1.000. The developed approach can be used in practice to discriminate advanced clones and cultivars of strawberry in an objective and nondestructive manner.

Keywords: image textures, classification models, fruit, Subspace Discriminant, Multilayer Perceptron

INTRODUCTION

Strawberry (*Fragaria* × *ananassa* Duch.) is a herbaceous plant cultivated and consumed worldwide [Sun et al. 2023, Şener et al. 2023]. Strawberries are planted due to their red, aromatic, and sweet fruit [Patel et al. 2023]. The red color and unique flavor resulting from the combination of taste, aroma, and mouthfeel sensations especially attract the consumers. Thus, strawberries can be intended for the fresh market or

✉ ewa.ropelewska@inhort.pl

used as food ingredients in bakery and dairy products, as well as in beverages [Teribia et al. 2021]. Strawberry can be considered as a functional food source with benefits to human health, since it's rich with numerous nutritional compounds such as polyphenols, antioxidants, vitamins, fiber, minerals, and trace elements. Due to the presence of these compounds, strawberry is characterized by antioxidant, anti-aging, pro-digestive, anti-inflammatory, antihypertensive, antiproliferative, and antihyperlipidemic activities [Şener et al. 2023, Ladika et al. 2024, Tang et al. 2024].

The color, shape and chemical properties of strawberries can differ depending on the cultivar [Sturm et al. 2003, Boonyakiat et al. 2016, Parra-Palma et al. 2020, Lee et al. 2022]. Furthermore, strawberry can be characterized by different volatile and taste profiles with specific and balanced stability during processing depending on the cultivar. For this reason, final products belonging to different cultivars are qualitatively different after processing and behave dissimilarly during storage, leading food manufacturers to pay more attention to raw material selection [Teribia et al. 2021].

Moreover, many strawberry breeding programs were carried out to develop new cultivars or improved clones better adapted to environmental conditions, more productive, with higher disease resistance and with better post-harvest quality [Galvão et al. 2017]. The resilience and adaptation to cultivation conditions can also depend on cultivars and clones. Therefore, clone selection with improved resistance is with importance in strawberry stress tolerance breeding [Dziadczyk et al. 2003]. Additionally, there may be differences in morphological characteristics in terms of color, brightness, shape, and size of individual clones and cultivars [de Souzam et al. 2021]. The distinguishing or identification of cultivars is essential to protect the breeder's rights, select improved cultivars for breeding programs and meet consumer needs. Initially, the strawberry cultivar identification was performed using morphological features and then biochemical markers such as isozymes [Jung et al. 2017]. For the effective assessment of strawberry genetic diversity and identification of cultivars and clones, molecular markers can also be applied [Tyrka et al. 2002, Whitaker 2011, Jung et al. 2017]. Both morphological and molecular analyses are destructive and require tedious and time-consuming work as well as hard technicity to achieve cultival discrimination. Therefore, the added value of artificial intelligence technologies can be explored.

Thus, the objective of this study was to distinguish advanced clones and cultivars of strawberry using a nondestructive, objective, and inexpensive approach involving image analysis and artificial intelligence. The innovative models were built based on selected textures extracted from images in different color channels R, G, B, L, a, b, X, Y, Z, U, V, and S.

MATERIALS AND METHODS

Materials

The research included fruit of advanced clones and cultivars of strawberry (Tab. 1), collected during the fully-ripening period of the plants in a field trial of the National Institute of Horticultural Research in Skierniewice, Poland. Each genotype was represented by 60 plants, grown in the soil and managed in accordance with recommendations for commercial plantations (mechanical and manual removal of weeds and runners, plant irrigation using self-propelled sprinkler, fertilization with YaraMila™ Complex multi-component fertilizer, integrated protection against diseases and pests in accordance with the current Strawberry Plant Protection Program). All the tested genotypes were characterized by very high productivity, large, attractive and firm fruits as well as low plant susceptibility to fungal leaf diseases. Fruits selected for the studies were uniform in shape and color, typical for

Table 1. Clones and cultivars of strawberry used in the experiment

Clones (breeding code and pedigree)	Cultivars
T-201536-06, pedigree Clery × Grandarosa	Fibion
T-201567-01, pedigree Patty × Panvik	Grandarosa
T-201457-16, pedigree Grandarosa × Elsanta	Markat

each of the genotype. Strawberries were washed and cleaned directly after harvesting, and then subjected to image acquisition.

Image acquisition and processing

The strawberry images were acquired using a digital camera (Canon Inc., Tokyo, Japan) on a black background using light-emitting diode illumination. Fruit images were obtained in one hundred repetitions for each clone of T-201536-06, T-201567-01, and T-201457-16, and each cultivar of Fibion, Grandarosa, and Markat. The acquired images were processed using MaZda software (Łódź University of Technology, Institute of Electronics, Łódź, Poland) [Szczypiński et al. 2007, Szczypiński et al. 2009, Strzelecki et al. 2013]. The image processing included image conversion to color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, and *S*, image segmentation based on the intensity of pixel brightness, and ROI (region of interest) determination. The last step was the texture extraction from images. For each fruit considered as one ROI, 2172 image texture parameters were computed based on the run-length matrix, co-occurrence matrix, autoregressive model, histogram, Haar wavelet transform, and gradient map.

Distinguishing advanced clones and cultivars of strawberry using machine learning models

Strawberry advanced clones and cultivars, such as T-201536-06, T-201567-01, T-201457-16, Fibion, Grandarosa, Markat were distinguished using models built based on selected image texture parameters. Various models were developed using WEKA machine learning software (Machine Learning Group, University of Waikato, Hamilton, New Zealand) [Witten and Frank 2005, Bouckaert et al. 2016, Frank et al. 2016] and MATLAB (MathWorks, Inc., Natick, MA, USA). Before building classification models, image textures with the highest discriminative power were selected by Best First using WEKA. The same set of selected textures was used to build models using both WEKA and MATLAB. A test mode of 10-fold cross-validation was applied in both cases. For machine learning models developed using WEKA, different algorithms from groups of Functions, Bayes, Meta, Lazy, Trees, and Rules were used. It was observed that the classifiers providing the highest correctness were Multilayer Perceptron from Functions, MultiClass Classifier from Meta, IBk from Lazy, and LMT from Trees. The parameters of classifiers used to build models are presented in Table 2.

For a model developed using each classifier, confusion matrix with accuracies for each strawberry class, average accuracy, Kappa statistic, Precision, Recall, MCC (Matthews Correlation Coefficient), F-measure, PRC Area (Precision-Recall Area), and ROC Area (Receiver Operating Characteristic Area) were determined using the Equations 1–10 [Ropelewska 2022, Ropelewska et al. 2022, Unlersen et al. 2022, Ropelewska et al. 2023]. Additionally, the PRC curves and ROC curves for each clone and cultivar were determined in the case of the model characterized by the highest average accuracy.

Table 2. The model parameters of classifiers applied to distinguish strawberry advanced clones and cultivars using WEKA

Classifier	Parameters
Multilayer Perceptron	autoBuild: True; batchSize: 100; decay: False; debug: False; doNotCheckCapabilities: False; momentum: 0.2; learningRate: 0.3; hiddenLayers: a; nominalToBinaryFilter: True; validationTreshold: 20; normalizeNumericClass: True; normalizeAttributes: True; reset: True; trainingTime: 500; resume: False; seed: 0
MultiClass Classifier	batchSize: 100; classifier: Logistic, ridge: 1.0E–8, maxIts: –1, numDecimalPlaces: 4; debug: False; logLossDecoding: False; doNotCheckCapabilities: False; method: one-against-all; randomWidthFactor: 2.0; seed: 1; use PairwiseCoupling: False
IBk	batchSize: 100; KNN: 1; debug: False; distanceWeighting: No distance weighting; doNotCheckCapabilities: False; meanSquared: False; nearestNeighbourSearchAlgorithm: LinearNNSearch
LMT	batchSize: 100; debug: False; fastRegression: True; doNotCheckCapabilities: False; errorOnProbabilities: False; minNumInstances: 15; numBoostingIterations: –1; splitOnResiduals: False

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \tag{1}$$

$$\text{Kappa} = \frac{\frac{(\text{TP} + \text{FP})(\text{TP} + \text{FN})}{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})} + \frac{(\text{TN} + \text{FP})(\text{TN} + \text{FN})}{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{MCC} = \frac{(\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN})}{\sqrt{((\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN}))}} \tag{5}$$

$$\text{F-measure} = \frac{2 \cdot \text{Precsion} \cdot \text{Recall}}{(\text{Precision} + \text{Recall})} \tag{6}$$

$$\text{PRC Area} = \text{area under Precision vs. Recall curve} \tag{7}$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{9}$$

$$\text{ROC Area} = \text{area under TPR vs. FPR curve} \tag{10}$$

where TP is true positive; TN is true negative; FP is false positive; FN is false negative; TPR is true positive rate; FPR is false positive rate.

In the case of models developed using MATLAB, also four most effective models were selected. The hy-perparameters of classifiers are presented in Table 3. For the evaluation of classification, the confusion matrices and average accuracies were determined.

RESULTS

Results generated using WEKA software. The confusion matrices and average accuracies of the classifica-

Table 3. The hyperparameters of classifiers used to distinguish strawberry advanced clones and cultivars using MATLAB

Classifier type	Model Hyperparameters
Linear Discriminant	Preset: Linear Discriminant; Covariance structure: Full
SVM	Preset: Quadratic SVM; Model function: Quadratic; Kernel scale: Automatic; Box constraint level: 1; Multiclass method: One-vs-one; Standardize data: Yes
Ensemble	Preset: Subspace Discriminant; Ensemble method: Subspace; Learner type: Discriminant; Number of learners: 30; Subspace dimension: 42
Neural Network	Preset: Wide Neural Network; Number of fully connected layers: 1; Iteration limit: 1000; First layer size: 100; Activation: ReLU; Regularization strength (Lambda): 0; Standardize data: Yes

tion of strawberry advanced clones, such as T-201536-06, T-201567-01, T-201457-16, and cultivars Fibion, Grandarosa, Markat built based on selected texture parameters from images in color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, and *S* using machine learning models are presented in Table 4. The most correct classification was obtained for a model developed using Multilayer Perceptron. The average accuracy reached 95.83%. Also, the Kappa statistic of 0.9500 was the highest for a model built using Multilayer Perceptron. In the case of other algorithms, the average accuracies and values of Kappa statistic were 93.89% and 0.9267 for LMT, 92.78% and 0.9133 for MultiClass Classifier, and 92.22% and 0.9067 for IBk, respectively. In the case of all models, the clone T-201567-01 was characterized by very high classification accuracy, reaching 100% for a model built using LMT, 98% for models developed using Multilayer Perceptron and IBk, and 97% in the case of a model built using MultiClass Classifier. Generally, the highest number of misclassified cases occurred between the clone T-201536-06 and Grandarosa. The high misclassification was also observed between clones T-201536-06 and T-201457-16. The number of cases belonging to Grandarosa and misclassified as T-201536-06 reached 15 for a model built using IBk. For the same model, as many as 7 cases from the actual class T-201536-06 were incorrectly classified as Grandarosa. The highest number of cases belonging to T-201457-16, which were incorrectly included in the predicted class T-201536-06 were equal to 7 and it was observed for a model developed using MultiClass Classifier.

The other classification performance metrics are shown in Table 5. It was found that strawberry clone

Table 4. The accuracies of classification of strawberry advanced clones and cultivars using machine learning models built based on selected image texture parameters

Algorithm	Predicted class (%)						Actual class	Average accuracy (%)
	T-201536-06	T-201567-01	T-201457-16	Fibion	Grandarosa	Markat		
Multilayer Perceptron	93	0	2	2	2	1	T-201536-06	95.83
	0	98	0	0	0	2	T-201567-01	
	2	0	98	0	0	0	T-201457-16	
	0	0	2	95	3	0	Fibion	
	3	2	0	3	92	0	Grandarosa	
	0	0	0	2	0	98	Markat	
MultiClass Classifier	93	2	2	3	0	0	T-201536-06	92.78
	1	97	0	0	0	2	T-201567-01	
	7	0	93	0	0	0	T-201457-16	
	3	0	2	95	0	0	Fibion	
	5	0	0	3	90	2	Grandarosa	
	5	3	0	2	2	88	Markat	
IBk	87	1	3	2	7	0	T-201536-06	92.22
	0	98	0	0	0	2	T-201567-01	
	3	2	93	0	2	0	T-201457-16	
	2	0	1	95	2	0	Fibion	
	15	0	0	2	82	1	Grandarosa	
	0	2	0	0	0	98	Markat	
LMT	88	1	2	2	5	2	T-201536-06	93.89
	0	100	0	0	0	0	T-201567-01	
	3	2	95	0	0	0	T-201457-16	
	2	0	0	95	1	2	Fibion	
	3	2	0	3	92	0	Grandarosa	
	0	0	0	5	2	93	Markat	

Table 5. The performance metrics of distinguishing strawberry advanced clones and cultivars based on selected texture parameters of images

Algorithm	Class	Precision	Recall	MCC	F-measure	PRC Area	ROC Area
Multilayer Perceptron	T-201536-06	0.949	0.933	0.930	0.941	0.977	0.983
	T-201567-01	0.983	0.983	0.980	0.983	1.000	1.000
	T-201457-16	0.967	0.983	0.970	0.975	0.991	0.998
	Fibion	0.934	0.950	0.931	0.942	0.977	0.989
	Grandarosa	0.948	0.917	0.919	0.932	0.994	0.999
	Markat	0.967	0.983	0.970	0.975	0.998	1.000
	Weighted average	0.958	0.958	0.950	0.958	0.989	0.995
MultiClass Classifier	T-201536-06	0.812	0.933	0.843	0.868	0.894	0.966
	T-201567-01	0.951	0.967	0.950	0.959	0.986	0.995
	T-201457-16	0.966	0.933	0.939	0.949	0.972	0.990
	Fibion	0.919	0.950	0.921	0.934	0.979	0.996
	Grandarosa	0.982	0.900	0.929	0.939	0.976	0.993
	Markat	0.964	0.883	0.908	0.922	0.983	0.997
	Weighted average	0.932	0.928	0.915	0.929	0.965	0.989
IBk	T-201536-06	0.813	0.867	0.806	0.839	0.704	0.911
	T-201567-01	0.952	0.983	0.961	0.967	0.929	0.978
	T-201457-16	0.949	0.933	0.930	0.941	0.901	0.957
	Fibion	0.966	0.950	0.950	0.958	0.915	0.964
	Grandarosa	0.891	0.817	0.825	0.852	0.758	0.858
	Markat	0.967	0.983	0.970	0.975	0.932	0.978
	Weighted average	0.923	0.922	0.907	0.922	0.857	0.941
LMT	T-201536-06	0.914	0.883	0.879	0.898	0.967	0.991
	T-201567-01	0.952	1.000	0.971	0.976	0.995	0.999
	T-201457-16	0.983	0.950	0.960	0.966	0.994	0.999
	Fibion	0.905	0.950	0.912	0.927	0.990	0.998
	Grandarosa	0.917	0.917	0.900	0.917	0.972	0.994
	Markat	0.966	0.933	0.939	0.949	0.995	0.999
	Weighted average	0.939	0.939	0.927	0.939	0.986	0.997

MCC – Matthews Correlation Coefficient, PRC Area – Precision-Recall Area, ROC Area – Receiver Operating Characteristic Area

T-201567-01 was distinguished by the highest values of Precision of 0.983, MCC of 0.980, F-measure of 0.983, PRC Area of 1.000, and ROC Area of 1.000 for a model developed using Multilayer Perceptron and the highest Recall of 1.000 for a model built using LMT. These values indicated high classification correctness. The lowest values of Precision of 0.812 (MultiClass Classifier), MCC of 0.806, F-measure of 0.839, and PRC Area of 0.704 (IBk) were determined for T-201536-06. Whereas the lowest Recall of 0.817 and ROC Area of 0.858 (IBk) were obtained for Grandarosa. It indicated low classification accuracies.

In addition to numeric values of classification performance metrics, the PRC (Precision–Recall) curves and ROC curves were determined for the model built using Multilayer Perceptron, which provided the highest average accuracy. The PRC curves for each clone and cultivar are presented in Figure 1 and the ROC curves are shown in Figure 2. The graphs confirmed very high correctness of distinguishing the clone T-201567-01. Both PRC Area and ROC Area were equal to 1.000, which is visible in Figures 1b and 2b, respectively. Whereas curves presented in Figures 1a and 1d indicate the lowest PRC Area of 0.977 for T-201536-06 and Fibion. The lowest ROC Area of 0.983 for T-201536-06 is confirmed by curves in Figure 1a.

Results generated using MATLAB software

Average accuracies of the classification of strawberry advanced clones and cultivars performed using

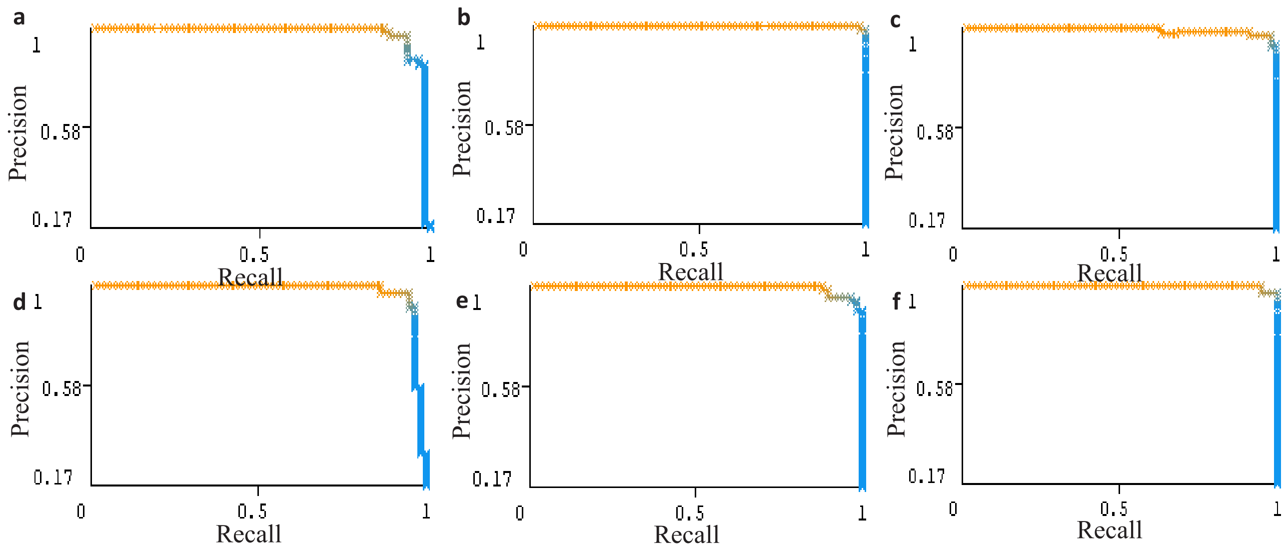

Fig. 1. The PRC (Precision–Recall) curves for distinguishing strawberry advanced clones T-201536-06 (a), T-201567-01 (b), T-201457-16 (c), and cultivars Fibion (d), Grandarosa (e), Markat (f) using Multilayer Perceptron

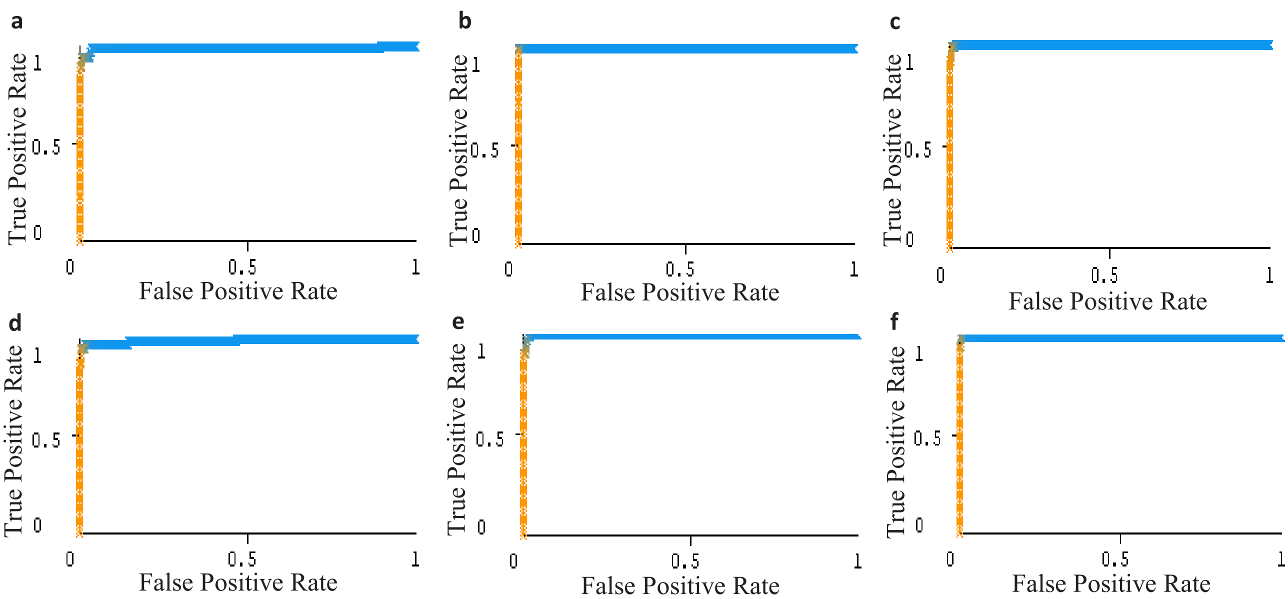

Fig. 2. The ROC (Receiver Operating Characteristic) curves for the classification of strawberry advanced clones T-201536-06 (a), T-201567-01 (b), T-201457-16 (c), and cultivars Fibion (d), Grandarosa (e), Markat (f) using Multilayer Perceptron

a Predicted class (%)

T-201536-06	T-201567-01	T-201457-16	Fibion	Grandarosa	Markat	Actual class
93	0	0	2	5	0	T-201536-06
0	98	0	0	0	2	T-201567-01
5	0	95	0	0	0	T-201457-16
0	0	1	97	2	0	Fibion
3	0	0	5	92	0	Grandarosa
0	0	0	0	0	100	Markat

b Predicted class (%)

T-201536-06	T-201567-01	T-201457-16	Fibion	Grandarosa	Markat	Actual class
88	0	3	2	7	0	T-201536-06
0	100	0	0	0	0	T-201567-01
7	0	93	0	0	0	T-201457-16
0	0	1	97	2	0	Fibion
2	0	0	1	97	0	Grandarosa
1	0	0	2	0	97	Markat

c Predicted class (%)

T-201536-06	T-201567-01	T-201457-16	Fibion	Grandarosa	Markat	Actual class
95	0	2	1	2	0	T-201536-06
0	100	0	0	0	0	T-201567-01
2	0	98	0	0	0	T-201457-16
0	0	2	93	5	0	Fibion
2	1	2	3	92	0	Grandarosa
0	0	0	0	0	100	Markat

d Predicted class (%)

T-201536-06	T-201567-01	T-201457-16	Fibion	Grandarosa	Markat	Actual class
90	0	5	0	3	2	T-201536-06
0	100	0	0	0	0	T-201567-01
5	0	95	0	0	0	T-201457-16
0	0	1	97	2	0	Fibion
5	1	0	2	92	0	Grandarosa
2	0	0	0	0	98	Markat

Fig. 3. The confusion matrices of classification of strawberry advanced clones and cultivars using models built based on selected image texture parameters using Linear Discriminant (a), Quadratic SVM (b), Subspace Discriminant (c), and Wide Neural Network (d)

MATLAB were equal to 95.80% for Linear Discriminant, 95.30% for Quadratic SVM, 96.30% for Subspace Discriminant, and 95.30% for Wide Neural Network. The confusion matrices in Figure 3 present the highest classification accuracy of 100.00% for T-201567-01 for three out of four applied models built using Quadratic SVM (Fig. 3b), Subspace Discriminant (Fig. 3c), and Wide Neural Network (Fig. 3d). The lowest accuracy of 88.00% was determined for T-201536-06 in the case of a model built using Quadratic SVM (Fig. 3b). The high number of misclassified cases were between T-201536-06 and Grandarosa and between T-201536-06 and T-201457-16. These observations were similar to the classification results obtained using WEKA (Tab. 4).

DISCUSSION

The approach involving image analysis and artificial intelligence innovative proved to be useful for the discrimination of advanced clones T-201536-06, T-201567-01, and T-201457-16, and cultivars Fibion, Grandarosa, and Markat of strawberry. Models built based on selected texture parameters from images in color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, and *S* using machine learning algorithms were successful and provided a high average accuracy of up to 96.30% for the model developed using MATLAB (Subspace Discriminant) and 95.83% for the model built by WEKA algorithm (Multilayer Perceptron). It showed that the obtained results were very similar, regardless of the applied software. The high discrimination results revealed the great usefulness of image texture parameters for distinguishing strawberry clones and cultivars in a nondestructive, objective, and effective manner.

In the case of models developed using WEKA and MATLAB, the high misclassification of cases was observed between the clone T-201536-06 and Grandarosa and between clones T-201536-06 and T-201457-16. It may be due to the fact that both clones had the Grandarosa cultivar in their pedigree. Generally, for the models built using WEKA and MATLAB, the clone T-201567-01 was discriminated with very high accuracy. It was confirmed by other performance metrics and graphs presented PRC (Precision–Recall) curves and ROC (Receiver Operating Characteristic) curves. It meant that the clone T-201567-01 was very different in terms of image textures from other clones and cultivars.

In the previous literature, there are also reports concerning the application of imaging and artificial intelligence for studies of strawberry. Yamamoto et al. [2015] used the image analysis system for the strawberry quality evaluation and cultivar identification based on appearance characteristics. The classification models built using linear discriminant analysis (LDA) based on a single feature type such as shape, size, or color classified 14 strawberry cultivars with an accuracy of less than 42%. However, the accuracy increased to 68% for a model combining shape, size, and color features. Nevertheless, the obtained accuracy of 68% was lower than the accuracies determined in our study for models developed based on image textures. Whereas Amoriello et al. [2022] predicted the internal quality features of strawberry based on color parameters, such as L^*, a^*, and b^* using the artificial neural network (ANN) and multiple linear regression models (MLR). The application of ANN allowed for obtaining high prediction results, such as $R^2 = 0.906$, and $R^2 = 0.943$ for antioxidant activity and the total monomeric anthocyanin, respectively. Color images and neural networks were also used by Choi et al. [2021] for the evaluation of the strawberry external quality. The recognition models developed based on RGB images using convolutional neural networks (CNNs) allowed for the distinguishing fresh, moldy, and bruised strawberries with the correctness reaching 97%. Computer vision combined with deep learning was also used by Patel et al. [2021] to detect the strawberry plant wetness.

In addition to color images, also hyperspectral imaging combined with artificial intelligence was used for strawberry quality evaluation. For example, real-time hyperspectral imaging and deep learning were applied for the in-field strawberry ripeness estimation providing a classification accuracy of 98.6% for the early ripe and ripe samples [Gao et al. 2020]. Hyperspectral imaging combined with deep learning was also used for strawberry maturity determination and soluble solids content estimation [Su et al. 2021], combined with support vector machine (SVM) for strawberry ripeness evaluation [Zhang et al. 2016], and with SVM and back propagation neural network (BPNN) for the identification of healthy, and bruise and fungi infected strawberries [Liu et al. 2018].

The above-mentioned literature data confirmed the usefulness of imaging and artificial intelligence for the determination of external and internal quality of strawberries and for strawberry classification. Our study expanded knowledge about the application of color imaging and machine learning and set new directions in nondestructive strawberry quality evaluation. It was revealed that image texture parameters can be useful for distinguishing strawberries based on external appearance. Undertaken studies can be continued by involving more clones and cultivars, internal features, and deep learning. The approach combining image analysis and artificial intelligence can be useful in practice to distinguish clones and cultivars of strawberry in a nondestructive and objective manner.

CONCLUSIONS

In this study, an objective, nondestructive, and inexpensive approach involving image analysis and artificial intelligence was developed to classify advanced clones and cultivars of strawberry. The innovative models developed based on selected textures from images in color channels *R*, *G*, *B*, *L*, *a*, *b*, *X*, *Y*, *Z*, *U*, *V*, and *S* proved to be useful for distinguishing advanced clones T-201536-06, T-201567-01, and T-201457-16, and cultivars Fibion, Grandarosa, and Markat with an average accuracy reaching 95.83% and 96.30% for models built using Multilayer Perceptron and Subspace Discriminant, respectively. The applied approach is a novelty in strawberry clone and cultivar classification. The procedure is characterized by great practical applications. Including image textures selected from a set of as many as 2172 parameters in the classification models allowed for very accurate discrimination of advanced clones and cultivar of strawberry. This approach based on image analysis and artificial intelligence may be useful for strawberry growers and processors. However, further research can be carried out, involving a larger number of clones and cultivars, fruit internal features, and deep learning.

FUNDING

The research was carried out in the frame of subsidy of the Ministry of Agriculture and Rural Development special-purpose – Task 3.4: “Developing strawberry (*Fragaria* × *ananassa* Duch.) breeding material, characterized by plant tolerance to verticillium, resistance to low temperatures and drought, and high content of health-promoting ingredients in fruits”.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Amoriello, T., Ciccoritti, R., Ferrante, P. (2022). Prediction of strawberries’ quality parameters using artificial neural networks. Agronomy, 12, 963. https://doi.org/10.3390/agronomy12040963

Boonyakiat, D., Chuamuangphan, C., Maniwara, P., Seehanam, P. (2016). Comparison of physico-chemical quality of different strawberry cultivars at three maturity stages. Int. Food Res. J., 23, 2405–2412.

Bouckaert, R.R., Frank, E., Hall, M., Kirkb, R., Reutemann, P., Seewald, A., Scuse, D. (2016) WEKA manual for version 3-9-1. University of Waikato, Hamilton, New Zealand.

Choi, J.Y., Seo, K., Cho, J.S., Moon, K.D. (2021). Applying convolutional neural networks to assess the external quality of strawberries. J. Food Compos. Anal., 102, 104071. https://doi.org/10.1016/j.jfca.2021.104071

de Souzam D.C., Ossani, P.C., Costa, A.S., Guerra, T.S., Araújo, A.L., Resende, F.V., Resende, L.V. (2021). Selection of experimental strawberry clones for fruit appearance attributes. Pesqi. Agropecu. Bras., 56, e02560. https://doi.org/10.1590/S1678-3921.pab2021.v56.02560

Dziadczyk, P., Bolibok, H., Tyrka, M., Hortyński, J.A. (2003). In vitro selection of strawberry (*Fragaria* × *ananassa* Duch.) clones tolerant to salt stress. Euphytica, 132(1), 49–55. https://doi.org/10.1023/A:1024647600516

Frank, E., Hall, M.A., Witten, I.H. (2016). The WEKA Workbench. Online appendix for data mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington.

Galvão, A.G., Resende, L.V., Maluf, W.R., de Resende, J.T.V., Ferraz, A.K.L., Marodin, J.C. (2017). Breeding new improved clones for strawberry production in Brazil. Acta Sci. Agron., 39, 149–155. https://doi.org/10.4025/actasciagron.v39i2.30773

Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y., Han, X. (2020) Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. Artif. Intell. Agric., 4, 31–38. https://doi.org/10.1016/j.aiia.2020.04.003

Jung, H.J., Veerappan, K., Natarajan, S., Jeong, N., Hwang, I., Nagano, S., Shirasawa, K., Isobe, S., Nou, I.S. (2017). A system for distinguishing octoploid strawberry cultivars using high-throughput SNP genotyping. Tropical Plant Biol., 10, 68–76. https://doi.org/10.1007/s12042-017-9185-8

Ladika, G., Strati, I.F., Tsiaka, T., Cavouras, D., Sinanoglou, V.J. (2024). On the assessment of strawberries' shelf-life and quality, based on image analysis, physicochemical methods, and chemometrics. Foods, 13, 234. https://doi.org/10.3390/foods13020234

Lee, C., Lee, J., Lee, J. (2022). Relationship of fruit color and anthocyanin content with related gene expression differ in strawberry cultivars during shelf life. Sci. Hortic., 301, 111109. https://doi.org/10.1016/j.scienta.2022.111109

Liu, Q., Sun, K., Peng, J., Xing, M., Pan, L., Tu, K. (2018). Identification of bruise and fungi contamination in strawberries using hyperspectral imaging technology and multivariate analysis. Food Anal. Methods, 11, 1518–1527. https://doi.org/10.1007/s12161-017-1136-3

Parra-Palma, C., Morales-Quintana, L., Ramos, P. (2020). Phenolic content, color development, and pigment‒related gene expression: a comparative analysis in different cultivars of strawberry during the ripening process. Agronomy, 10, 588. https://doi.org/10.3390/agronomy10040588

Patel, A., Lee, W.S., Peres, N.A. (2021). Strawberry plant wetness detection using computer vision and deep learning. Smart Agric. Technol., 1, 100013. https://doi.org/10.1016/j.atech.2021.100013

Patel, H., Taghavi, T., Samtani, J.B. (2023). Fruit quality of several strawberry cultivars during the harvest season under high tunnel and open field environments. Horticulturae, 9, 1084. https://doi.org/10.3390/horticulturae9101084

Ropelewska, E. (2022). Diversity of plum stones based on image texture parameters and machine learning algorithms. Agronomy, 12, 762. https://doi.org/10.3390/agronomy12040762

Ropelewska, E., Cai, X., Zhang, Z., Sabanci, K., Aslan, M.F. (2022). Benchmarking machine learning approaches to evaluate the cultivar differentiation of plum (*Prunus domestica* L.) kernels. Agriculture, 12, 285. https://doi.org/10.3390/agriculture12020285

Ropelewska, E., Rady, A.M., Watson, N.J. (2023). Apricot stone classification using image analysis and machine learning. Sustainability, 15, 9259. https://doi.org/10.3390/su15129259

Strzelecki, M., Szczypiński, P., Materka, A., Klepaczko, A. (2013) A software tool for automatic classification and segmentation of 2D/3D medical images. Nucl. Instrum. Methods Phys. Res., sec. A, Accel. Spectrom. Detect. Assoc. Equip., 702, 137–140. https://doi.org/10.1016/j.nima.2012.09.006

Sturm, K., Koron, D., Stampar, F. (2003). The composition of fruit of different strawberry varieties depending on maturity stage. Food Chem., 83, 417–422. https://doi.org/10.1016/S0308-8146(03)00124-9

Su, Z., Zhang, C., Yan, T., Zhu, J., Zeng, Y., Lu. X. et al. (2021) Application of hyperspectral imaging for maturity and soluble solids content determination of strawberry with deep learning approaches. Front. Plant Sci., 12,736334. https://doi.org/10.3389/fpls.2021.736334

Sun, C., Yang, X., Gu, Q., Jiang, G., Shen, L., Zhou, J., Li, L., Chen, H., Zhang, G., Zhang, Y. (2023). Comprehensive analysis of nanoplastic effects on growth phenotype, nanoplastic accumulation, oxidative stress response, gene expression, and metabolite accumulation in multiple strawberry cultivars. Science of The Total Environment 897, 165432. https://doi.org/10.1016/j.scitotenv.2023.165432

Szczypiński, P.M., Strzelecki, M., Materka, A. (2007) Mazda-a software for texture analysis. In Proceedings of the 2007 International Symposium on Information Technology Convergence (ISITC 2007), Jeonju, Korea, 23–24 November 2007, pp. 245–249.

Szczypiński, P.M., Strzelecki, M., Materka, A., Klepaczko, A. (2009). MaZda – A software package for image texture analysis. Comp. Meth. Progr. Biomed., 94, 66–76. https://doi.org/10.1016/j.cmpb.2008.08.005

Şener, S., Sayğı, H., Duran, C.N. (2023). Responses of in vitro strawberry plants to drought stress under the influence of nano-silicon dioxide. Sustainability, 15, 15569. https://doi.org/10.3390/su152115569

Tang, X., Li, Y., Fang, M., Li, W., Hong, Y., Li, Y. (2024). Effects of different water storage and fertilizer retention substrates on growth, yield and quality of strawberry. Agronomy, 14, 205. https://doi.org/10.3390/agronomy14010205

Teribia, N., Buvé, C., Bonerz, D., Aschoff, J., Hendrickx, M., Van Loey, A. (2021). Effect of cultivar, pasteurization and storage on the volatile and taste compounds of strawberry puree. LWT (Lebensm. Wiss. Technol.), 150, 112007. https://doi.org/10.1016/j.lwt.2021.112007

Unlersen, M.F., Sonmez, M.E., Aslan, M.F., Demir, B., Aydin, N., Sabanci, K., Ropelewska, E. (2022). CNN–

SVM hybrid model for varietal classification of wheat based on bulk samples. Eur. Food Res. Technol., 248, 2043–2052. https://doi.org/10.1007/s00217-022-04029-4

Whitaker, V.M. (2011). Applications of molecular markers in strawberry. J. Berry Res., 1, 115–127.

Witten, I.H., Frank, E. (2005). Data mining: practical machine learning tools and techniques. Elsevier, San Francisco.

Yamamoto, K., Ninomiya, S., Kimura, Y., Hashimoto, A., Yoshioka, Y., Kameoka, T. (2015). Strawberry cultivar identification and quality evaluation on the basis of multiple fruit appearance features. Comput. Electron. Agric., 110, 233–240. https://doi.org/10.1016/j.compag.2014.11.018

Zhang, C., Guo, C., Liu, F., Kong, W., He, Y., Lou, B. (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. J. Food Eng., 179, 11–18. https://doi.org/10.1016/j.jfoodeng.2016.01.002