Journal of Agricultural Machinery

Homepage: https://jame.um.ac.ir

Research Article
Vol. 15, No. 1, Spring 2025, p. 129-144

Hyperparameter Optimization of ANN, SVM, and KNN Models for Classification of Hazelnuts Images Based on Shell Cracks and Feature Selection Method

H. Bagherpour [1*], **F. Fatehi** [1], **A. Shojaeian** [1], **R. Bagherpour** [2]

1- Department of Biosystems Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran
2- Department of Computer Engineering, School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran
(*- Corresponding Author Email: h.bagherpour@basu.ac.ir)

Received: 01 May 2024 Revised: 02 July 2024 Accepted: 11 July 2024 Available Online: 11 February 2025	**How to cite this article:** Bagherpour, H., Fatehi, F., Shojaeian, A., & Bagherpour, R. (2025). Hyperparameter Optimization of ANN, SVM, and KNN Models for Classification of Hazelnuts Images Based on Shell Cracks and Feature Selection Method. *Journal of Agricultural Machinery*, *15*(1), 129-144. https://doi.org/10.22067/jam.2024.87830.1244

Abstract

In some countries, people commonly consume hazelnuts in their shells to extend shelf life or due to technological limitations. Therefore, open-shell hazelnuts are more marketable. At the semi-industrial scale, open-shell and closed-shell hazelnuts are currently separated from each other through visual inspection. This study aims to develop a new algorithm to separate open-shell hazelnuts from cracked or closed-shell hazelnuts. In the first approach, dimension reduction techniques such as Sequential Forward Feature Selection (SFFS) and Principal Component Analysis (PCA) were used to select or extract a combination of color, texture, and grayscale features for the model's input. In the second approach, individual features were used directly as inputs. In this study, three famous machine learning models, including Support Vector Machine (SVM), K-nearest neighbors (KNN), and Multi-Layer Perceptron (MLP) were employed. The results indicated that the SFFS method had a greater effect on improving the performance of the models than the PCA method. However, there was no significant difference between the performance of the models developed with combined features (98.00%) and that of the models using individual features (98.67%). The overall results of this study indicated that the MLP model, with one hidden layer, a dropout of 0.3, and 10 neurons using Histogram of Oriented Gradients (HOG) features as input, is a good choice for classifying hazelnuts into two classes of open-shell and closed-shell.

Keywords: Closed-shell, Dimension reduction, Machine learning, Open-shell, PCA

Introduction

Hazelnut is one of the garden products with the highest nutritional value for humans. It is utilized as snack, in baking and desserts, and in breakfast cereals like muesli. In confectionery, it is used for making pralines and are combined with chocolate for truffles, alongside other popular treats like chocolate bars and hazelnut cocoa spreads like Nutella. It is also used in the cosmetics industry (FAOSTAT, 2021).

Hazelnuts are available in the market both in-shell and shelled. Although in many industrialized countries, hazelnuts are sold in the form of kernels, in many countries, including the Third World countries, a large amount of hazelnut is marketed in the form of open-shell. Shelled hazelnuts account for 5 to

https://doi.org/10.22067/jam.2024.87830.1244

10% of the global hazelnut market (FAOSTAT, 2021). During the cracking process undertaken to increase the marketability of hazelnuts, three different classes are produced after cracking: open-shell, cracked, and closed-shell. Among these, only the open-shell hazelnuts can be sold in the market. As a result, separating the cracked and closed-shell hazelnuts and making them open-shell is necessary. Since the cracks are very small, manual separation of closed-shell from open-shell hazelnuts is a tedious and time-consuming task. In commercial scale production, having a fast, non-destructive method and reliable classification is crucial.

Commercial hazelnut processing generally includes drying, sizing, cracking, and separating impurities (Menesatti *et al.*, 2008; Wang, Jung, McGorrin, & Zhao, 2018). By reviewing previous studies, few studies have been found in the field of hazelnut classification. In a study, sound signal was used to classify hazelnuts into two classes of underdeveloped and fully developed hazelnuts. The sound signals were obtained by dropping hazelnuts from a certain height onto a steel plate (Kalkan & Yardimci, 2006). In another study, a morphological method based on elliptic Fourier approximation to closed contours in a two-dimensional plane was applied to the RGB images to classify four local hazelnut cultivars in Italy. The coefficients of harmonic equations were obtained by PLS-DA. Menesatti *et al.* (2008) evaluated the potential use and efficacy of shape-based techniques in order to discriminate four traditional Italian hazelnut cultivars. The higher percentage of correct classification accuracy was reported between 77.5%- 98.8%. Seventeen hazelnut cultivars were classified using a developed convolutional neural network. This network had the highest accuracy (98.63%) as compared to other pre-trained models (Taner, Öztekin, & Duran, 2021).

A significant number of studies have presented the use of machine learning (ML) techniques for classification or qualitative evaluation of nuts and fruits. ML methods have been widely used for classification of various agricultural products, such as grading hazelnut kernels (Giraudo *et al.*, 2018), detection of hazelnut cultivars (Taner *et al.*, 2021), grading almond kernels (Vidyarthi , Singh, Xiao, & Tiwari, 2021), orange (Komal & Sonia, 2019), cucumber (Pourdarbani & Sabzi, 2022), apple (Lashgari, Imanmehr, & Tavakoli, 2020), classification of weed seeds (Luo *et al.*, 2023), and detection of abnormal lettuce leaves (Yang *et al.*, 2023). In a latest study on hazelnut classification based on shell crack detection, a deep convolutional neural network (DCNN) algorithm was employed (Shojaeian *et al.*, 2023). Although the results of their study were satisfactory, they did not assess the features individually, without providing any insights regarding the importance of the specific features.

To the best of our knowledge, there is currently no intelligent system available for the classification of hazelnuts based on the presence of shell cracks. Therefore, this research aims to classify the hazelnuts based on cracks in their shells, utilizing color and texture features extracted from RGB images, employing models such as MLP, SVM, and KNN.

Materials and Methods

Fig. 1a illustrates the schematic diagram of steps involved in modeling machine learning methods. In the first approach, images of the hazelnut samples were captured, and subsequently some preprocessing operations were performed. After extracting the color, grayscale, and texture features, their dimensions were reduced using Principal Component Analysis (PCA) technique, and Sequential Forward Feature Selection (SFFS) was employed for feature selection. As shown in Fig. 1b in the second approach, four investigated features were used individually as inputs to three classifiers. In this approach, the same optimized hyperparameters obtained in the first approach were utilized.

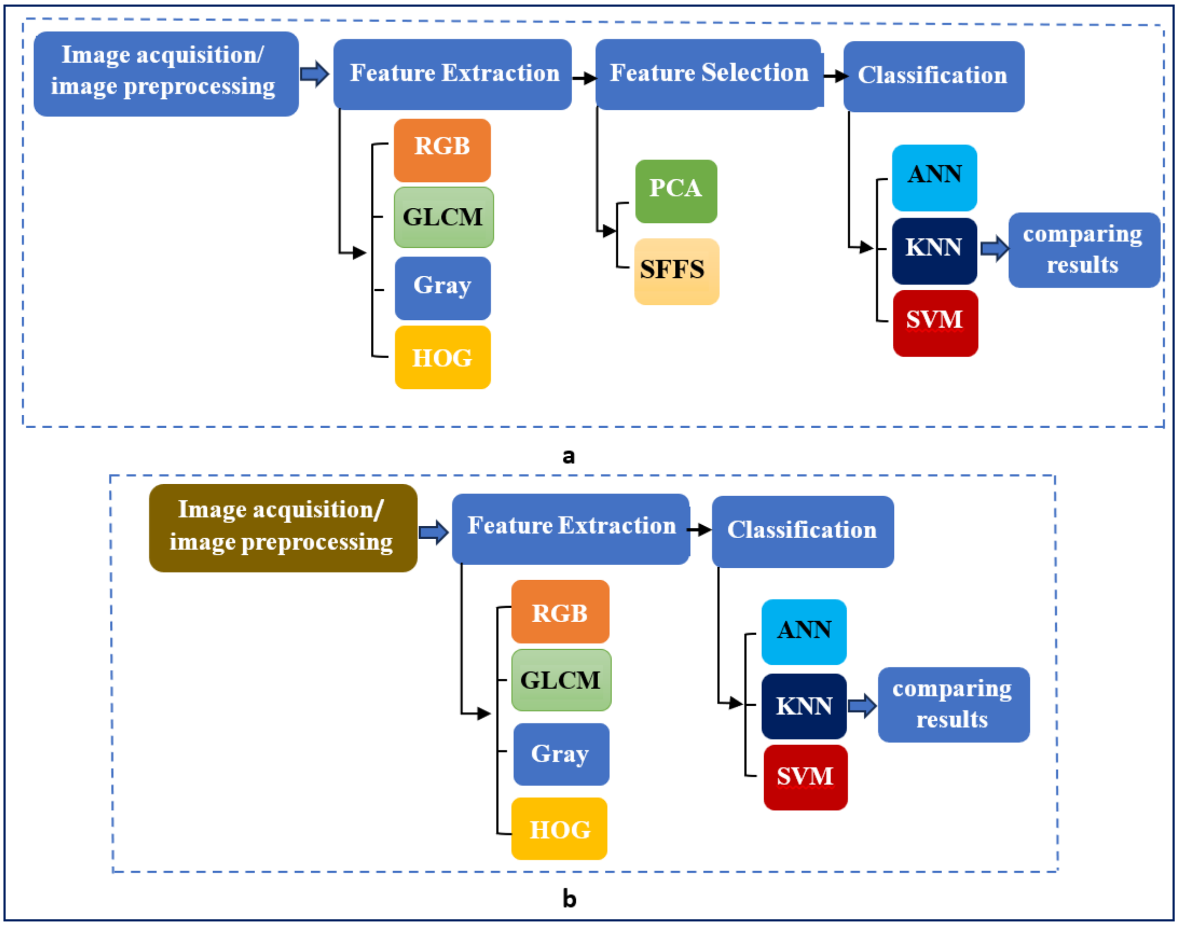

Fig. 1. Flowchart of hazelnut classification using machine learning algorithms. Two approaches were used: a) incorporating feature selection algorithms (first approach) and b) individual features used as input to three classifiers (second approach)

Sample preparation

Hazelnut samples were purchased during the summer of 2022 from Rahim Abad, located in Rudsar city, Gilan province, Iran. Five hundred samples were randomly selected for each class. The classes were as follows: 1) open-shell and 2) closed-shell hazelnuts (without cracks or with tiny cracks). Among these samples, 48% were open-shell, 32% were closed-shell, and 20% had tiny cracks.

To prepare images under consistent conditions and eliminate ambient effects, an imaging box was used. A camera (Samsung J5 smartphone) with a resolution of 2448 × 2448 pixels was positioned at the top of the box. Additionally, a 6-watt circular LED panel provided uniform illumination on the sample. The inner side walls of the box were covered with white cardboard, while blue cardboard was used as the background to increase the contrast between the hazelnuts and the background. Examples of captured hazelnut images from two different classes are shown in Fig. 2.

Feature Extraction

Crack Size

Five steps were carried out to identify cracks on the shell surface (Fig. 3). These steps include removing the background and converting the image to grayscale, implementing thresholding to create a mask, applying the mask to the original image using the concatenate function (cat (a, c)), and finally, applying a threshold to the R component of the RGB and the S component of the HSV to reveal the cracks in the hazelnuts (Fig. 3 f). An area threshold was then applied to separate open and cracked shell samples.

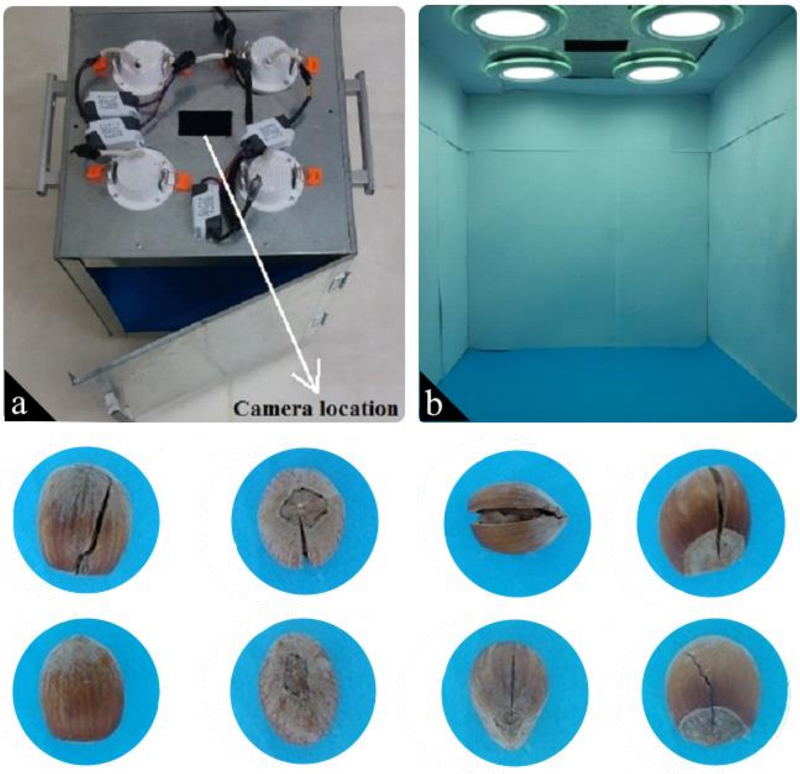

Fig. 2. The (a) exterior and (b) interior views of the imaging box. (c) The images in the first row and the second row show the open-shell (class 1) and closed-shell (class 2) hazelnuts, respectively

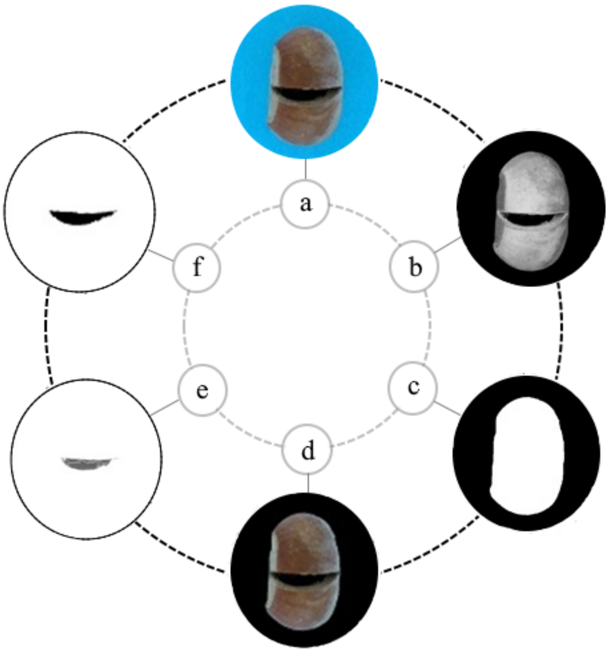

Fig. 3. Image processing for crack detection. a) Original RGB image, b) gray-scale image, c) binary image, d) concatenation of the original image and the corresponding masks, e) crack detection through a linear combination of the R component of the RGB color space and the S component of the HSV color space, and f) thresholding on image "e"

Color and Texture Features

The mean, standard deviation, skewness, and elongation of the color components were calculated using the image shown in Fig. 3 d. Table 1 shows these features, with R, G, and B representing the red, green, and blue components of the RGB image, respectively. Additionally, p, n, and i are the normalized color histogram, intensity, and number of color component levels, respectively.

To extract textural features, Fig. 3d was converted to a gray-scale image and the Gray-Level Co-Occurrence Matrix (GLCM) was derived from each image. Furthermore, all textural features were extracted from the gray-scale image (Pourreza, Pourreza, Abbaspour-Fard, & Sadrnia, 2012). The gray features include the histograms of gray images and the aforementioned matrices as well as those mentioned in Table 1. The Gray-Level Co-Occurrence Matrix (GLCM) is a statistical method for analyzing the texture of an image. It considers the spatial relationship between pixels with specific intensity values. The GLCM functions characterize the texture by calculating how often pairs of pixels with certain values occur in a specified spatial relationship within the image.

Table 1- The features extracted from RGB, GLCM, and gray matrices

Features	Equation
Color Features	
Mean R	$\mu_R = \sum_i ip_R(i)$
Mean G	$\mu_G = \sum_i ip_G(i)$
Mean B	$\mu_B = \sum_i ip_B(i)$
Standard deviation R	$\sigma_R = \sqrt{\sum_i (i-\mu_R)^2 p_R(i)}$
Standard deviation G	$\sigma_G = \sqrt{\sum_i (i-\mu_G)^2 p_G(i)}$
Standard deviation B	$\sigma_B = \sqrt{\sum_i (i-\mu_B)^2 p_B(i)}$
Skewness R	$(\sum_{i=1}^{n}(i-\mu_R)^3)/(n-1){\sigma_R}^3$
Skewness G	$(\sum_{i=1}^{n}(i-\mu_G)^3)/(n-1){\sigma_G}^3$
Skewness B	$(\sum_{i=1}^{n}(i-\mu_B)^3)/(n-1){\sigma_B}^3$
Kurtosis R	$(n\sum_I (i-\mu_R)^4/\sum_I (i-{\mu_R}^2)^2)-3$
Kurtosis G	$(n\sum_I (i-\mu_G)^4/\sum_I (i-{\mu_G}^2)^2)-3$
Extracted features from GLCM matrix	
Mean	$\mu = \sum_i ip(i)$
Standard deviation	$\sigma = \sqrt{\sum_i (i-\mu)^2 p(i)}$
Smoothness	$1\text{-}1/(1+\sigma^2)$
Third moment	$\sum_i (i-\mu)^3 p(i)$
Uniformity	$\sum_i p(i)^2$
Entropy	$-\sum_{i,j} p(i,j) log\,(p(i,j))$
Uniformity	$\sum_{i,j} p(i,j)^2$
Homogeneity	$\sum_{i,j} p(i,j)/(1+(i-j)^2)$
Inertia	$\sum_{i,j} (i-j)^2 p(i,j)$
Cluster shade	$\sum_{i,j} (i+j-2\mu)^3 p(i,j)$
Cluster prominence	$\sum_{i,j} (i+j-2\mu)^4 p(i,j)$
Maximum probability	$Max\,(p(i,j))$
Correlation	$\sum_{i,j} (i-\mu)(j-\mu)\sigma^2 p(i,j)$
Extracted features from gray matrices	
Mean	$\mu = \sum_i \mathrm{ip(i)}$
Standard deviation	$\sigma = \sqrt{\sum_i (i-\mu)^2 \mathrm{p(i)}}$
Third moment	$\sum (i-\mu)^3\, p(i)$
Smoothness	$1\text{-}1/(1+\sigma^2)$
Uniformity	$\sum_i \mathrm{p(i)}^2$
Entropy	$-\sum_i \mathrm{p(i)log\,(p(i))}$
Crack area	$\sum i_b \mid (i_b \mathrm{i} = 1)$

In addition to the above features, the Histogram of Oriented Gradients (HOG) feature was also used as input of the proposed models to classify the hazelnuts according to their cracks. For this purpose, image sizes of 128×128 pixels were examined. HOG was calculated using 8×8 cell sizes and spread across 9 bins, resulting in an 8100-dimensional feature vector for each image.

Feature Selection

Feature selection is an important step in the process of building classifiers. It is a process that chooses a subset of features from the original set of features so that the features space is optimally reduced according to a certain criterion (Tan, Hoon, Yong, Kong, & Lin, 2005). Using the first approach in this study, a large number of features were initially extracted from the samples to identify the optimal features. The performance of the classifiers was then evaluated based on each category of input features. On the other hand, the extracted features may contain noise and irrelevant information, so the number of features should be reduced by employing feature conditioning methods (Garcia-Allende, Mirapeix, Conde, Cobo, & Lopez-Higuera, 2009). For this purpose, the PCA and SFFS algorithms were applied separately on the features to reduce the number of features based on their approach. In this research, six features were selected by SFFS for MLP, and eleven features were selected for SVM and KNN. In the PCA method, the six components that could explain 98% of variances were selected as inputs for the models.

Machine Learning Models

To achieve a simple structure, with the least complexity and the best performance without underfitting and overfitting, several MLP architectures were evaluated by changing the number of layers (one and two layers) and the number of neurons (3-12 neurons) in each hidden layer. As Fig. 4 shows, in the proposed network, six selected features by the SFFS method were considered as input of the network. The sigmoid active function was considered in the hidden layer neurons and the linear activation function was considered in the output layer neurons of the network. The Levenberg–Marquardt algorithm was used to train the network and the MSE criterion was also used to stop the training (Heaton, 2008).

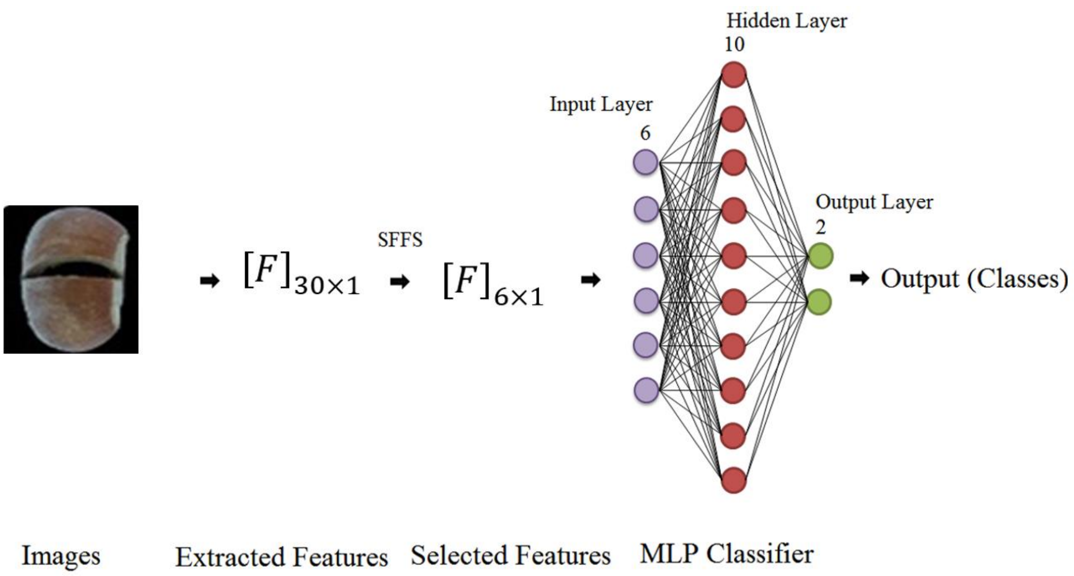

Fig. 4. The architecture of MLP model with one hidden layer containing 10 neurons

For each experiment, the initial learning rate was set as 0.001 and the number of iterations was 300. In data segmentation, 70%, 15%, and 15% of the data were used for training, validation, and testing of the network, respectively.

The KNN rule is one of the well-known supervised learning models in classification tasks. This rule simply retains all training sets during learning and assigns a class to each query represented by the majority label of its k-nearest neighbors in the training dataset (Gou, Du, Zhang, & Xiong, 2012). The main problem is that the behavior of this model is affected by many parameters, including distance criteria, weights of neighborhoods (Table 2), and the number of neighbors (K) (Geler, Kurbalija, Radovanović, & Ivanović, 2016). Therefore, the effect of these factors was evaluated in this study. In these models as well as SVM, 80% of the dataset was considered for training and 20% of the dataset for testing. Note that the values of the neighborhood size k in the experiments vary from 3 to 11 by Step 2.

Table 2- Different weights of KNN model

Model	Weight (Sigma and C are constant)
KNN	-----
WKNN1	$1/D$
WKNN2	$1/D^2$
WKNN3	$1/(D^2 + C)$
WKNN4	exp (D^2/Sigma)

The SVM was another model investigated in this study. This model is a binary classifier which gives better performance in the classification tasks. SVM classifies two classes by constructing a hyperplane in high-dimensional feature space. A decision hyperplane is constructed in this higher dimension such that the distance between hyperplane and the support vectors of both classes is maximized (Way, Sahiner, Hadjiiski, & Chan, 2010). We evaluated the SVM model using the suggested RBF for the classification models (Manekar & Waghmare, 2014). There are two parameters in the RBF Kernel type of SVM: C (Cost) and g (gamma). The accuracy of the SVM for RBF type depends on these two parameters (Gopi, Jyothi, Narayana, & Sandeep, 2023).

Evaluation Metrics

The performance of the classifiers was evaluated considering the results obtained from the confusion matrix, along with key statistical metrics: accuracy (Eq. 1), sensitivity (Eq. 2), specificity (Eq. 3), precision (Eq. 4), and F1-Score (Eq. 5). MATLAB R2019a was used to extract the features and implement the models.

$$Accuracy = \frac{TP+TN}{N} \quad (1)$$

$$Sensitivity\ (Recall) = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$F1\text{-}Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (5)$$

where N is the total number of samples. TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives. The F1-score can have values between 0 and 1, with 1 being the best score.

Results and Discussion

Effect of dimension reduction methods on the model's performance

In this study, PCA and SFFS methods were used to assess the effect of dimension reduction methods. The results in Table 3 illustrates the confusion matrix obtained from the MLP results related to the proposed method and PCA. These results indicated that the feature vectors obtained by SFFS outperform PCA. In the SFFS method, the F1-score for open-shell and closed-shell was 98.67 and 98.67%, respectively. While in the PCA method, this index was 78.67 and 80.00%, respectively. In a study to recognize facial expressions using RGB images, the feature selection method of SFFS and the ML (Machine Learning) approach suggested that the selected subset of features not only enhances the classification performance, but also reduces computational complexity, making the system more practical for real-time applications (Li, Lu, & Liu, 2014). Furthermore, the SFFS method demonstrated superior performance in detecting stems and calyxes (SC) in apple stems using support vector classifiers (Unay, Gosselin, & Debeir, 2006).

Table 3- Confusion matrix of MLP model using SFFS and PCA method

		Predicted	
	Class	Open-Shell	Cracked or Closed-shell
		SFFS	
Actual	Open-Shell	74	1
	Cracked or Closed-shell	1	74
		PCA	
	Open-Shell	59	16
	Cracked or Closed-shell	15	60

In examining the performance of SVM and KNN classifiers with the feature subsets selected from SFFS, these models showed the classification accuracies of 96.67% and 98%, respectively. On the other hand, like the MLP classifier, in the SVM and KNN classifiers, using the features mapped by PCA, the accuracy of these models was less than 79% (Table 4). The low accuracy of the PCA method suggests that using linear transformation to map features on the orthogonal directions can complicate the feature space and may not always be beneficial (Jolliffe, 2002). In the SFFS algorithm, the feedback of the desired classifier is considered to select the feature during feature selection (Lu, Wang, Wu, & Xie, 2016).

Table 4- Effect of dimension reduction methods on the performance of MLP, SVM, and WKNN2 models in the classification of hazelnut (WKNN2 results was obtained with k=7, criteria distance of Cityblock)

		Test data			
Method	**Model**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**Accuracy (%)**
PCA	MLP	79.03	79.03	79.03	79.03
	SVM	50.00	100	66.67	50.00
	WKNN2	62.94	71.33	66.83	64.67
SFFS	MLP	98.67	98.67	98.67	98.67
	SVM	96.05	97.33	96.69	96.67
	WKNN2	96.15	100	98.04	98.00

Number of Neurons of the MLP Structure

In the MLP classifier, the number of neurons in the hidden layer has the highest impact on the performance of the network. Therefore, finding its optimal value is important (Heaton, 2008). In examining the effect of the number of neurons, the artificial neural network (ANN) model with 10 neurons in the hidden layer had the highest accuracy (98.67%). In this selected network, the lowest mean squared error (MSE = 0.08379) for validation data was obtained in the epoch of 17 (Fig. 5). Similar results have been published in studies that investigated the effect of the number of neurons in the hidden layer on the performance of artificial neural networks (Çolak, 2021; Liu, Starzyk, & Zhu, 2007). As the results of table 5 show, using a dropout of 0.3 between input and hidden layers significantly improved the network accuracy. The decrease in accuracy with a dropout rate of 0.5 can be attributed to removing too many neurons during the training process.

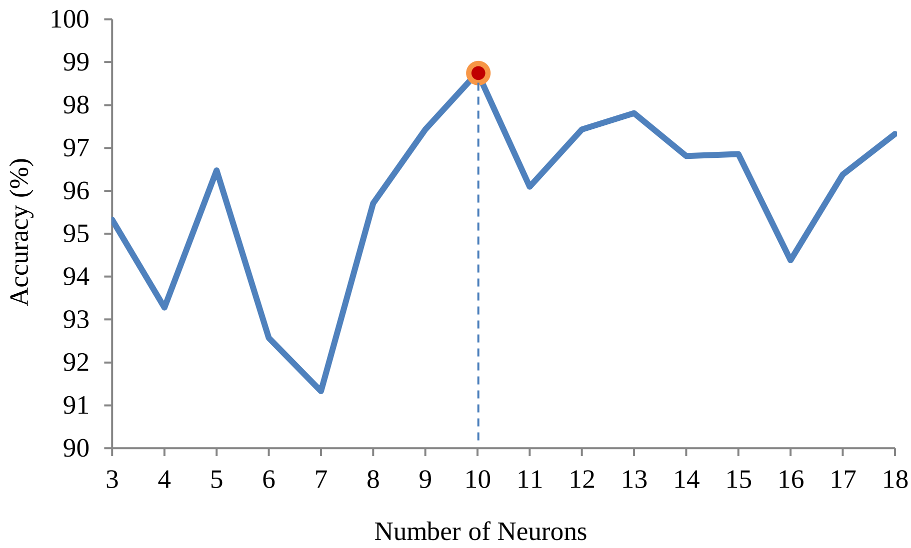

Fig. 5. Accuracy of MLP with different neurons in hidden layer

Table 5- Effect of dropout and the number of hidden layers on the accuracy of ANN model with HOG feature

Model	Number of layers	Dropout	Accuracy (%)
ANN	1	-	0.955
	1	0.3	0.986
	1	0.5	0.930
	2	-	0.940
	2	0.3	0.958
	2	0.5	0.942

KNN Classifier

The performance of various KNN classifier configurations was evaluated by considering different distance metrics (D), different neighborhood weighting schemes (w), and varying numbers of neighbors (k). The best average accuracy of the test data for each classifier was obtained with k=7 (Fig. 6) and the Cityblock distance metric (Table 6). In general, the weighted KNN models outperformed the unweighted model for different values of k. Although the accuracy of most weighted KNN configurations was above 95%, the classification accuracy of WKNN2 (98.00%) was the highest among the weighted KNNs. Therefore, the WKNN2 classifier was selected for further analysis. In the similar study to compare the performance of KNN and WKNN, the results of their comparison showed that the WKNN had higher performance than KNN (Tarakci & Ozkan, 2021). Evaluating the performance of KNN and WKNN in the classification of the UCI database revealed that the highest and lowest classification accuracy was related to WKNN and KNN, respectively (Gou *et al.*, 2012).

Table 6- Effect of distance criteria and weight of distance on the performance of the KNN model

Model	Distance criteria (with SFFS method and k=7)					
	Chebychev	**Cityblock**	**Correlation**	**Cosine**	**Euclidean**	**Mincowski**
KNN	89.67	95.31	94.33	93.67	92.42	89.33
WKNN1	93.33	97.33	96.33	97.33	95.23	92.67
WKNN2	**93.33**	**98.00**	**96.43**	**97.40**	**95.67**	**94.20**
WKNN3	90.07	96.33	94.33	93.67	93.14	89.67
WKNN4	90.15	97.33	94.33	93.67	95.33	90.33

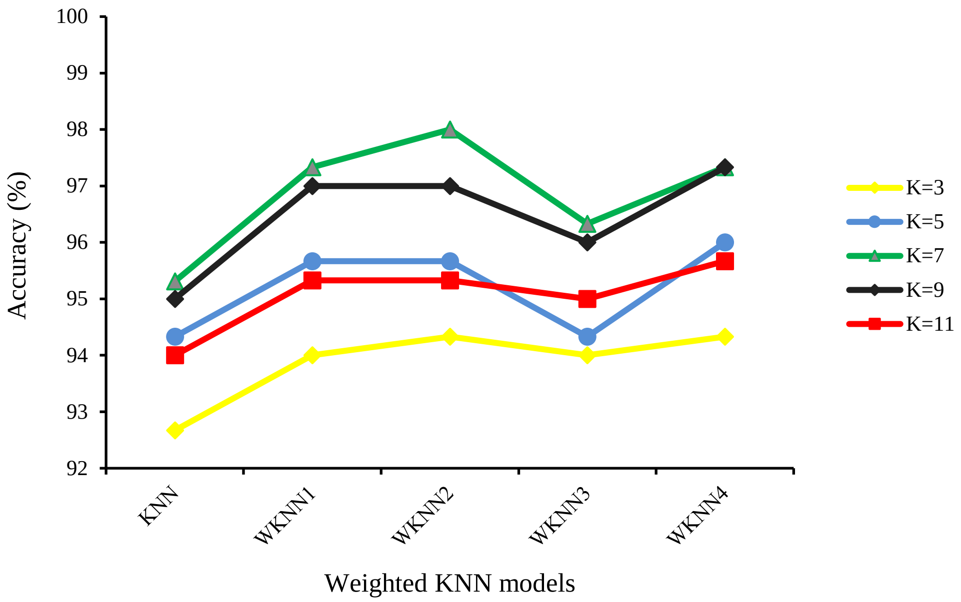

Fig. 6. Effect of number of neighborhood and weight of distance on the accuracy of the KNN model with the distance criteria of Cityblock and reduction method of SFFS

Effect of different individual features on the classifiers' accuracy

Fig. 7 shows the accuracy of MLP, SVM, and KNN classifiers based on different individual features. The results shown in this chart indicate that the color features (mean R, mean G, and mean B) and grayscale features performed well in the classification of hazelnuts. Conversely, the GLCM features yielded poor results. The high performance of the Color feature can be attributed to the presence of cracks on the Hazelnut surfaces. The larger the cracks, the greater the effect on the average value of the color indices. It should be mentioned that for all three feature types, the MLP model outperformed the SVM and KNN models. However, by comparing the results, although the MLP model achieved the highest accuracy (98.67%) using the HOG feature, it shows little difference with color and gray features, and it can be said that these three methods exhibited similar performance. Additionally, in the overall comparison between the classifiers, the KNN classifier exhibited lower performance than the other classifiers. In a similar study to compare ANN, Fuzzy, EDT, and KNN models with the aim of developing a cherry fruit packing system, the ANN model with HOG feature showed the higher accuracy of 95% (Momeny, Jahanbakhshi, Jafarnezhad, & Zhang, 2020).

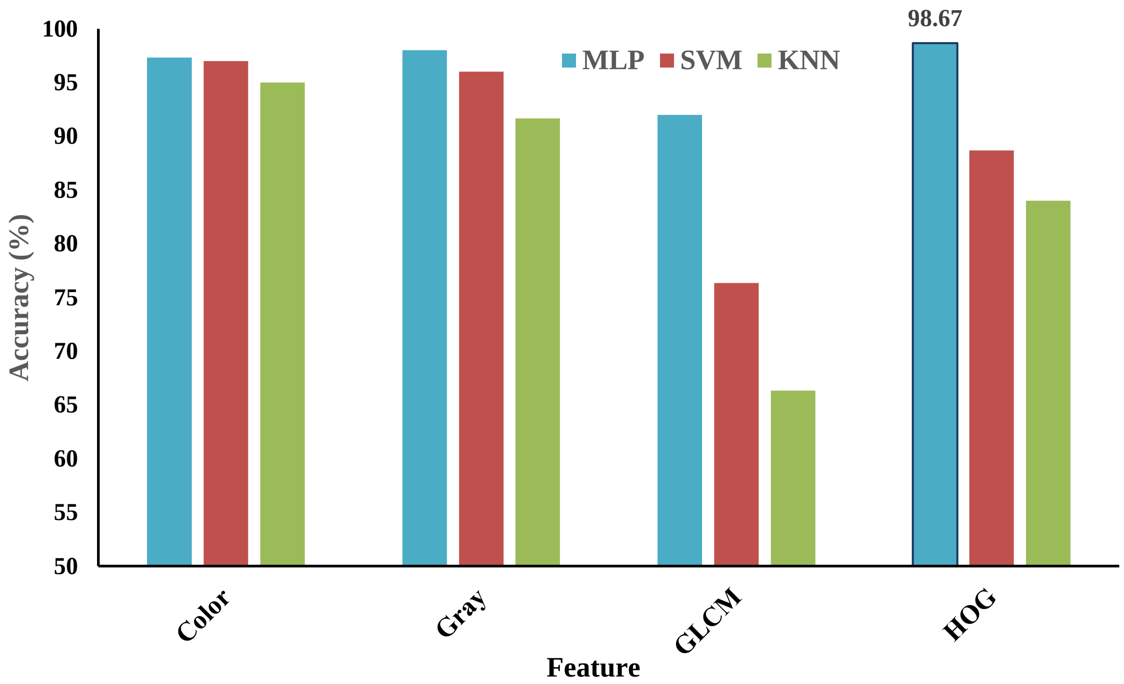

Fig. 7. Classification accuracy of MLP, SVM, and KNN using different features

The results of model evaluation are shown in Table 7. According to the F1-score measure, among the three features (HOG, Color, and Gray), the HOG is the best feature for the MLP model, while color features are recommended for the SVM and KNN models. Although all three models demonstrated satisfactory accuracy, the MLP showed better predictive capabilities for hazelnut classification based on surface cracks.

In the similar study that aimed to classify strawberry fruit into two classes of ripe and unripe, six classifiers including MLP, SVM, KNN, DT, NBC, and LR were investigated using bioimpedance data and surface color features. The classification results highlighted that, among all the tested models, MLP networks had the best performances (Ibba *et al.*, 2021). Four methods of SVM, KNN, and LDA (Linear Discriminant Analysis) were investigated to distinguish healthy and defective apples from each other. For this purpose, HOG and GLCM features were extracted. The SVM classifier was able to achieve 98.9% accuracy using these features. Additionally, applying PCA to these features did not affect the accuracy of the SVM and KNN classifiers (Singh & Singh, 2019). In a study, different classifiers including MLP and SVM were used to detect cracks in the walls using features extracted from the grayscale images. The MLP classifier exhibited the best performance in detecting cracked walls (Hallee, Napolitano, Reinhart, & Glisic, 2021).

Compared to previously studies, there have been hardly any studies in the literature performing classification of nuts using machine learning models to compare our results. However, we found some similar research in literature on smart sorting of pistachio nuts and almonds based on acoustic signals and deep learning approaches. Omid (2011) proposed an expert system based on acoustic emission signal and fuzzy logic classifier for sorting open and closed-shell pistachio nuts and the overall accuracy of the sorting system was 95.56 % for test datasets. In the other study, the performance of feature learning from frequency spectrum was tested for sorting pistachio nuts. The accuracy of the MLP classifier with features extracted from wavelet domain data was 96.1% (Hosseinpour-

Zarnaq, Omid, Taheri-Garavand, Nasiri, & Mahmoudi, 2022). The results of our proposed ANN model are similar to those reported in these studies. It is worth noting that in the similar study, authors detected hazelnut based on their crack using deep convolutional neural network (DCNN) algorithm (Shojaeian *et al.*, 2023). While their approach demonstrated superior detection accuracy compared to ours, our study has transparently disclosed the specific features utilized, which was not the case in their work. Additionally, their model is highly elaborate and computationally intensive.

Table 7- Classification performance of MLP, SVM, and KNN models at different features

			MLP		
Feature	Color	Gray	GLCM	HOG	Crack
Sensitivity	0.997	0.977	0.899	0.998	0.957
Specificity	0.944	0.984	0.775	0.976	0.763
precision	0.952	0.988	0.816	0.971	0.779
F1-Score	0.975	0.982	0.855	0.986	0.859
			SVM		
Sensitivity	0.987	0.947	0.640	0.833	0.933
Specificity	0.953	0.980	0.613	0.940	0.807
precision	0.955	0.979	0.623	0.933	0.828
F1-Score	0.970	0.963	0.631	0.880	0.878
			KNN		
Sensitivity	0.987	0.960	0.900	0.813	0.893
Specificity	0.913	0.973	0.727	0.867	0.333
precision	0.919	0.983	0.767	0.859	0.573
F1-Score	0.952	0.920	0.828	0.836	0.698

Conclusion

In countries where hazelnuts are sold in shell form, creating open-shell hazelnuts can increase the value of the product and the proportion of satisfied customers. The results of this study revealed that the well-known machine learning methods such as MLP, SVM, and KNN have great potential for the classification of hazelnuts. Although many features showed strong correlations with the hazelnut cracks, a greater number of them, especially HOG, exhibited higher accuracy. Meanwhile, the MLP model using the HOG feature achieved the highest accuracy, while GLCM features yielded low accuracy. The higher accuracy of the models using HOG features can be attributed to the fact that HOG can detect the object’s edge and provide the outline of a shape, which can be effective features for representing different types of cracks. Additionally, SFFS as a feature selection method showed better results than PCA. The overall results of this study clearly indicate that it is feasible to monitor and classify hazelnuts based on shell cracks. While the developed machine learning models demonstrated a good ability in classifying nuts, the main drawback of this study is the lack of information about situations where the crack is on the side of the hazelnut, which should be considered in future studies. It is suggested to employ two cameras to capture images of the falling hazelnuts.

Conflict of Interest

The authors declare no competing interests.

Author Contributions

H. Bagherpour: Supervision, Conceptualization, Methodology, Software, Reviewing.

F. Fatehi: Software, Methodology, Data pre and post processing, Writing, Validation.

A. Shojaeian: Data curation, Methodology.

R. Bagherpour: Software, Validation.

References

1. Çolak, A. B. (2021). A novel comparative investigation of the effect of the number of neurons on the predictive performance of the artificial neural network: An experimental study on the thermal conductivity of ZrO_2 nanofluid. *International Journal of Energy Research*, *45*(13), 18944-18956. https://doi.org/10.1002/er.6989
2. FAOSTAT. (2021). Crops production data. http://wwwfaoorg/faostat/en/#data/QC. Accessed 20 March 2021
3. Garcia-Allende, P. B., Mirapeix, J., Conde, O. M., Cobo, A., & Lopez-Higuera, J. M. (2009). Spectral processing technique based on feature selection and artificial neural networks for arc-welding quality monitoring. *Ndt & E International*, *42*(1), 56-63. https://doi.org/10.1016/j.ndteint.2008.07.004
4. Geler, Z., Kurbalija, V., Radovanović, M., & Ivanović, M. (2016). Comparison of different weighting schemes for the k NN classifier on time-series data. *Knowledge and Information Systems*, *48*, 331-378. https://doi.org/10.1007/s10115-015-0881-0
5. Giraudo, A., Calvini, R., Orlandi, G., Ulrici, A., Geobaldo, F., & Savorani, F. (2018). Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams. *Food Control*, *94*, 233-240. https://doi.org/10.1016/j.foodcont.2018.07.018
6. Gopi, A. P., Jyothi, R. N. S., Narayana, V. L., & Sandeep, K. S. (2023). Classification of tweets data based on polarity using improved RBF kernel of SVM. *International Journal of Information Technology*, *15*(2), 965-980. https://doi.org/10.1007/s41870-019-00409-4
7. Gou, J., Du, L., Zhang, Y., & Xiong, T. (2012). A new distance-weighted k-nearest neighbor classifier. *J. Inf. Comput. Sci*, *9*(6), 1429-1436.
8. Hallee, M. J., Napolitano, R. K., Reinhart, W. F., & Glisic, B. (2021). Crack detection in images of masonry using cnns. *Sensors*, *21*(14), 4929. https://doi.org/10.3390/s21144929
9. Heaton, J. (2008). *Introduction to Neural Networks with Java*. Heaton Research, Inc.
10. Hosseinpour-Zarnaq, M., Omid, M., Taheri-Garavand, A., Nasiri, A., & Mahmoudi, A. (2022). Acoustic signal-based deep learning approach for smart sorting of pistachio nuts. *Postharvest Biology and Technology*, *185*, 111778. https://doi.org/10.1016/j.postharvbio.2021.111778
11. Ibba, P., Tronstad, C., Moscetti, R., Mimmo, T., Cantarella, G., Petti, L., ... & Lugli, P. (2021). Supervised binary classification methods for strawberry ripeness discrimination from bioimpedance data. *Scientific Reports*, *11*(1), 11202. https://doi.org/10.1038/s41598-021-90471-5
12. Jolliffe, I. T. (2002). *Principal component analysis for special types of data* (pp. 338-372). Springer New York. https://doi.org/10.1007/0-387-22440-8_13
13. Kalkan, H., & Yardimci, Y. (2006, September). *Classification of hazelnut kernels by impact acoustics*. In 2006 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing (pp. 325-330). IEEE. https://doi.org/10.1109/mlsp.2006.275569
14. Komal, K., & Sonia, D. (2019). GLCM algorithm and SVM classification method for Orange fruit quality assessment. *International Journal of Engineering Research & Technology (IJERT)*, *8*(9), 697-703.
15. Lashgari, M., Imanmehr, A., & Tavakoli, H. (2020). Fusion of acoustic sensing and deep learning techniques for apple mealiness detection. *Journal of Food Science and Technology*, *57*, 2233-2240. https://doi.org/10.1007/s13197-020-04259-y
16. Li, J., Lu, H., & Liu, X. (2014). Feature selection method based on SFFS and SVM for facial expression recognition. In 2014 *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE.
17. Liu, Y., Starzyk, J. A., & Zhu, Z. (2007). Optimizing number of hidden neurons in neural networks. *EeC*, *1*(1), 6.

18. Lu, F., Wang, D., Wu, H., & Xie, W. (2016). *A multi-classifier combination method using sffs algorithm for recognition of 19 human activities.* In Computational Science and Its Applications–ICCSA 2016: 16th International Conference, Beijing, China, July 4-7, 2016, Proceedings, Part II 16 (pp. 519-529). Springer International Publishing. https://doi.org/10.1007/978-3-319-42108-7_40
19. Luo, T., Zhao, J., Gu, Y., Zhang, S., Qiao, X., Tian, W., & Han, Y. (2023). Classification of weed seeds based on visual images and deep learning. *Information Processing in Agriculture*, *10*(1), 40-51. https://doi.org/10.1016/j.inpa.2021.10.002
20. Manekar, V., & Waghmare, K. (2014). Intrusion detection system using support vector machine (SVM) and particle swarm optimization (PSO). *International Journal of Advanced Computer Research*, *4*(3), 808.
21. Menesatti, P., Costa, C., Paglia, G., Pallottino, F., D'Andrea, S., Rimatori, V., & Aguzzi, J. (2008). Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystems Engineering*, *101*(4), 417-424. https://doi.org/10.1016/j.biosystemseng.2008.09.013
22. Momeny, M., Jahanbakhshi, A., Jafarnezhad, K., & Zhang, Y. D. (2020). Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach. *Postharvest Biology and Technology*, *166*, 111204. https://doi.org/10.1016/j.postharvbio.2020.111204
23. Omid, M. (2011). Design of an expert system for sorting pistachio nuts through decision tree and fuzzy logic classifier. *Expert Systems with Applications*, *38*(4), 4339-4347. https://doi.org/10.1016/j.eswa.2010.09.103
24. Pourdarbani, R., Sabzi, S. (2022). Detection of Cucumber Fruits with Excessive Consumption of Nitrogen using Hyperspectral imaging (With Emphasis on Sustainable Agriculture). *Journal of Environmental Sciences Studies*, *7*(4), 5485-5492.
25. Pourreza, A., Pourreza, H., Abbaspour-Fard, M. H., & Sadrnia, H. (2012). Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Computers and Electronics in Agriculture*, *83*, 102-108. https://doi.org/10.1016/j.compag.2012.02.005
26. Shojaeian, A., Bagherpour, H., Bagherpour, R., Parian, J. A., Fatehi, F., & Taghinezhad, E. (2023). The Potential Application of Innovative Methods in Neural Networks for Surface Crack Recognition of Unshelled Hazelnut. *Journal of Food Processing and Preservation*, 2023(1), 2177724. https://doi.org/10.1155/2023/2177724
27. Singh, S., & Singh, N. P. (2019). Machine learning-based classification of good and rotten apple. In *Recent Trends in Communication, Computing, and Electronics: Select Proceedings of IC3E 2018* (pp. 377-386). Springer Singapore. https://doi.org/10.1007/978-981-13-2685-1_36
28. Tan, S. S., Hoon, G. K., Yong, C. H., Kong, T. E., & Lin, C. S. (2005). *Mapping search results into self-customized category hierarchy*. In Intelligent Information Processing II: IFIP TC12/WG12. 3 International Conference on Intelligent Information Processing (IIP2004) October 21–23, 2004, Beijing, China 2 (pp. 311-323). Springer US. https://doi.org/10.1007/0-387-23152-8_41
29. Taner, A., Öztekin, Y. B., & Duran, H. (2021). Performance analysis of deep learning CNN models for variety classification in hazelnut. *Sustainability*, *13*(12), 6527. https://doi.org/10.3390/su13126527
30. Tarakci, F., & Ozkan, I. A. (2021). Comparison of classification performance of kNN and WKNN algorithms. *Selcuk University Journal of Engineering Sciences*, *20*(2), 32-37.
31. Unay, D., Gosselin, B., & Debeir, O. (2006, January). *Apple stem and calyx recognition by decision trees*. In Proceedings of the 6th IASTED International Conference on Visualization, Imaging, and Image Processing, VIIP (pp. 549-552).
32. Vidyarthi, S. K., Singh, S. K., Xiao, H. W., & Tiwari, R. (2021). Deep learnt grading of almond kernels. *Journal of Food Process Engineering*, *44*(4), e13662.

https://doi.org/10.1111/jfpe.13662

33. Wang, W., Jung, J., McGorrin, R. J., & Zhao, Y. (2018). Investigation of the mechanisms and strategies for reducing shell cracks of hazelnut (*Corylus avellana* L.) in hot-air drying. *Lwt*, *98*, 252-259. https://doi.org/10.1016/j.lwt.2018.08.053
34. Way, T. W., Sahiner, B., Hadjiiski, L. M., & Chan, H. P. (2010). Effect of finite sample size on feature selection and classification: a simulation study. *Medical Physics*, *37*(2), 907-920. https://doi.org/10.1118/1.3284974
35. Yang, R., Wu, Z., Fang, W., Zhang, H., Wang, W., Fu, L., ... & Cui, Y. (2023). Detection of abnormal hydroponic lettuce leaves based on image processing and machine learning. *Information Processing in Agriculture*, *10*(1), 1-10. https://doi.org/10.1016/j.inpa.2021.11.001

نشریه ماشین‌های کشاورزی
https://jame.um.ac.ir

مقاله پژوهشی
جلد ۱۵، شماره ۱، بهار ۱٤۰٤، ص ۱٤٤-۱۲۹

بهینه‌سازی هایپرپارامترهای مدل‌های ماشین بردار پشتیبان، k نزدیک‌ترین همسایه و شبکه عصبی مصنوعی برای طبقه‌بندی تصاویر فندق‌ها بر اساس روش انتخاب ویژگی‌ها

حسین باقرپور [۱*]، فرهاد فاتحی [۱]، علیرضا شجاعیان [۱]، رضا باقرپور [۲]

تاریخ دریافت: ۱۴۰۳/۰۲/۱۲
تاریخ پذیرش: ۱۴۰۳/۰۴/۲۱

چکیده

در برخی کشورها، فندق‌ها به دلیل محدودیت‌های فناوری موجود و افزایش طول عمر نگهداری‌شان، معمولاً با پوسته مصرف می‌شوند. بنابراین، فندق‌های خندان مشتری پسندی بالاتری دارند. در مقیاس نیمه‌صنعتی، فندق‌های خندان و دهان بسته در حال حاضر از طریق بازرسی بصری از یکدیگر جدا می‌شوند. این مطالعه به‌منظور توسعه یک الگوریتم جدید برای جداسازی فندق‌های خندان از فندق‌های ترک‌خورده یا دهان بسته انجام شده است. در رویکرد اول، تکنیک‌های کاهش بعد مانند روش‌های مبتنی بر انتخاب ویژگی (SFFS) و تحلیل مؤلفه اصلی (PCA) برای انتخاب یا استخراج ترکیبی از ویژگی‌های رنگ، بافت و خاکستری به‌عنوان ورودی مدل استفاده شدند. در رویکرد دوم، ویژگی‌های به شکل انفرادی مستقیماً به‌عنوان ورودی‌ها استفاده شدند. در این مطالعه، سه مدل معروف یادگیری ماشین، شامل ماشین بردار پشتیبان (SVM)، نزدیک‌ترین همسایه‌ها (KNN) و پرسپترون چندلایه (MLP) مورد استفاده قرار گرفتند. نتایج نشان داد که روش SFFS تأثیر بیشتری در بهبود عملکرد مدل‌ها نسبت به روش PCA دارد. با این حال، تفاوت معنی‌داری بین عملکرد مدل‌های توسعه‌یافته با ویژگی‌های ترکیبی (۹۸/۰۰٪) و عملکرد مدل‌های با استفاده از ویژگی‌های انفرادی (۹۸/۶۷٪) وجود نداشت. نتایج کلی این مطالعه نشان داد که مدل MLP با یک لایه پنهان، دراپ اوت برابر با ۰/۳ و ۱۰ نورون، با استفاده از ویژگی HOG به‌عنوان ورودی، انتخاب خوبی برای طبقه‌بندی فندق‌ها به دو دسته خندان و دهان بسته می‌باشد.

واژه‌های کلیدی: PCA، خندان، کاهش بعد، دهان بسته، یادگیری ماشین

۱- گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه بو علی سینا، همدان، ایران
۲- گروه مهندسی کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت، تهران، ایران
(*- نویسنده مسئول: Email: h.bagherpour@basu.ac.ir)

https://doi.org/10.22067/jam.2024.87830.1244