European Food Research and Technology (2024) 250:97–110
https://doi.org/10.1007/s00217-023-04369-9

ORIGINAL PAPER

Detection of hazelnut varieties and development of mobile application with CNN data fusion feature reduction-based models

Bunyamin Gencturk[1] · Sadiye Arsoy[1] · Yavuz Selim Taspinar[2] · Ilkay Cinar[3] · Ramazan Kursun[4] · Elham Tahsin Yasin[5] · Murat Koklu[3]

Received: 26 June 2023 / Revised: 31 August 2023 / Accepted: 2 September 2023 / Published online: 28 September 2023

Abstract
In many crops worldwide, including hazelnuts, the majority of stages in production and delivery to end-users are conducted either manually or with machine equipment lacking the advancements brought by technology. Non-destructive, fast, and reliable methods, particularly deep learning algorithms, have emerged as prominent techniques for determining product quality and classification in fruits, vegetables, and cereal products in recent years. This study aims to classify hazelnuts using deep learning algorithms, thereby minimizing the labor, time, and cost expended during the sorting process. Hazelnut images were obtained from Giresun, Ordu, and Van hazelnut varieties. The dataset consists of 1165 images of Giresun, 1324 images of Ordu, and 1138 images of Van hazelnut varieties. The classification was performed using deep learning models such as InceptionV3 and ResNet50. To combine the classification capabilities of the models, an InceptionV3 + ResNet50 data fusion model was created using the data fusion method. In addition, feature reduction processes were conducted by adding a convolutional layer to the data fusion model to decrease the number of features. The classification was conducted using a total of 3627 images, resulting in a 100% classification accuracy. Furthermore, the classification times of all models were analyzed. Based on these analyses, the 1024 reduced features data fusion model with 100% classification accuracy exhibited the shortest classification time. This model was selected, and a mobile application was developed for easy on-field hazelnut classification. The hazelnut classification performed using deep learning algorithms in the application will facilitate the work of both non-experts and professionals in industrial and personal domains. Through these methods, patents for products and devices developed for use in different industries can be obtained, thereby increasing the economic value added of our country.

Keywords Artificial intelligence · Computer vision · Convolutional neural networks · Deep learning · Hazelnut classification

✉ Yavuz Selim Taspinar
ytaspinar@selcuk.edu.tr

Bunyamin Gencturk
bunyamin.gencturk.58@gmail.com

Sadiye Arsoy
sadiyearsoy21@gmail.com

Ilkay Cinar
ilkay.cinar@selcuk.edu.tr

Ramazan Kursun
rkursun@selcuk.edu.tr

Elham Tahsin Yasin
ilham.tahsen@gmail.com

Murat Koklu
mkoklu@selcuk.edu.tr

[1] Technology Faculty, Selcuk University, Konya, Türkiye

[2] Doganhisar Vocational School, Selcuk University, 42930 Doganhisar, Konya, Türkiye

[3] Department of Computer Engineering, Selcuk University, 42250 Selcuklu, Konya, Türkiye

[4] Guneysinir Vocational School, Selcuk University, 42490 Guneysinir, Konya, Türkiye

[5] Graduate School of Natural and Applied Sciences, Selcuk University, Konya, Türkiye

Introduction

Hazelnut (*Corylus avellana L.*) is a widely popular product worldwide. It is mentioned that the hazelnut plant, grown along the Eastern Black Sea coasts, consists of more than 15 species [1]. In recent years, along with Turkey, which is the leading global producer of hazelnuts, many countries have seen an increase in hazelnut production due to advancements in agricultural techniques and the emergence of new varieties [2, 3]

Since 2017, the global hazelnut production has reached 1 million tons. Turkey holds the first position in the world hazelnut production with a share of 65.3%, followed by Italy with 13.1%, Azerbaijan with 4.3%, and the United States with 2.9%. When considering Turkey's leading position in hazelnut production, the yield obtained in Turkey is reported to be 154 kg/ha as of 2017. Hazelnuts have a significant place among agricultural products in Turkey, being the most exported and foreign currency-earning agricultural commodity [4]. Turkey plays a significant role in the market and pricing determinations by ensuring the production of high-quality hazelnuts [5]. Therefore, Turkey continues to maintain its important role in the hazelnut industry, preserving the significance it holds even in the present day [6, 7].

Hazelnuts generate employment opportunities and contribute to the effective utilization of resources. They are utilized as raw materials, semi-finished products, and finished products (such as hazelnut paste/spread) in various industries. This utilization enhances the added value of Turkey's agricultural economy [8]. In addition to their economic value, hazelnuts have applications as raw materials in the confectionery, chocolate, bakery, dairy [2], cosmetic, and pharmaceutical industries [9].

In hazelnut production, it is important to cultivate high-quality, uniform, and well-filled nuts that meet market standards. Hazelnut size, shape, presence of defects, flavor, as well as oil content, oil properties, and protein content are essential factors in evaluating hazelnut varieties. Since these characteristics are influenced by environmental and cultural factors, the relationship between variety and quality becomes significant [2, 10]. These factors play a crucial role in enabling producers to gain market recognition. In Turkey, similar to many other crops, the majority of stages in hazelnut production and reaching the end-user are carried out either manually or with mechanical equipment that lacks the innovations brought by technology [11]. Therefore, producers, consumers, and regulatory authorities require new techniques and methods to determine the geographical origin and authenticity of different products [12]. Non-destructive, fast, and reliable methods have become increasingly important in determining product quality or type. Deep learning algorithms have emerged as a frequently used approach in recent years for fruit [13, 14], vegetable [15, 16], and cereal [17, 18] products.

Deep learning is one of the evolving research areas in machine learning [19]. Deep learning, in general, considers two factors: the supervised or unsupervised learning of nonlinear operations in multiple layers or stages.

Deep learning can be utilized for the automation of quality control, production, and distribution processes in the food industry. Food classification enables the categorization of products based on their types, maturity levels, quality, sizes, and other attributes. This allows for the rapid and accurate classification of food products, facilitating early separation of unwanted or low-quality items and enhancing the efficiency of production processes. Furthermore, food classification can improve quality control processes in the food industry, contributing to the enhancement of product quality. As a result, consumers can be provided with higher quality and safer food products [20]. The aim of this article is to automatically and accurately classify hazelnut varieties using their visual features. In addition, the intention is to develop an application that can be used in the field by utilizing the obtained classification models. Accordingly, the procedures conducted in this study are outlined below:

- InceptionV3 and ResNet50 models were used for image classification.
- Data fusion was performed by combining the features of InceptionV3 and ResNet50 models to increase classification accuracy and reduce computation time.
- The classification performance of models with different feature numbers was compared by reducing the features of the obtained data fusion model.
- The most efficient and successful classification model that can perform classification in the shortest time was used in the mobile application.
- The mobile application made the classification models accessible to everyone.

The contributions of the study based on the conducted procedures are as follows:

- *Increase in efficiency:* Image processing techniques can increase efficiency in hazelnut harvesting. Identification of hazelnut varieties can accelerate and automate the harvesting process, reducing time and labor costs.
- *Quality control:* The detection of hazelnut varieties using image processing techniques can play a significant role in quality control. Determining the quality and homogeneity of hazelnut varieties can provide a competitive advantage in export.

- *Disease detection:* Image processing techniques can be used to detect diseases in hazelnut trees. Early diagnosis of hazelnut tree diseases can help prevent their spread.
- *Automatic classification*: Automatic classification of hazelnut varieties using image processing techniques can speed up the classification process and reduce human errors.
- *Data collection:* Image processing techniques can provide hazelnut farmers with data collection and analysis capabilities. This enables hazelnut farmers to make better decisions and manage their products more efficiently.
- *Precision agriculture:* Image processing techniques can enable precision agriculture practices in hazelnut orchards. This allows for determining the needs of each tree and intervening when necessary, increasing product productivity and making agricultural activities more sustainable.

The rest of the study is sectioned as follows. The second section reviews other relevant studies in the literature. The third section presents the creation of the dataset, the dataset itself, the classification methods and architectures used in the study, and performance evaluation methods. The experimental results are presented in the fourth section, and the discussion and conclusions are provided in the final section.

Related works

When the literature is examined, it can be observed that hazelnut varieties have been classified in recent years using deep learning methods, machine learning methods, and computer vision. The following studies are provided below.

Sayinci et al. [21] aimed to compare the shape characteristics of 17 different hazelnut varieties grown in Turkey and differentiate them based on shape features using image processing, particularly Elliptic Fourier analysis (EFA). The study reported highly successful results in terms of shape features among hazelnut varieties using EFA. Differences between hazelnut types were extracted using principal component analysis (PCA). They stated that the highest classification result among the varieties was 93.9%.

Guvenc et al. [22] aimed to classify damaged, undamaged (clean), and rotten hazelnut samples obtained through the processing of in-shell hazelnuts using computer vision. They captured images of the damaged, clean, and rotten in-shell hazelnut samples and classified them using Support Vector Machine-Radial Basis Function (SVM-RBF), Bayes, Artificial Neural Networks (ANN), Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA). They reported achieving the highest accuracy value of 93.57% using Artificial Neural Networks (ANN).

Bayrakdar et al. [23] aimed to determine the type and quality of in-shell hazelnuts using image processing techniques based on the size and shape features of the hazelnut shells. They used the area, geodesic diameter, and radius of the largest circle inside the hazelnut from the obtained images and classified the hazelnut quality and type with 84% accuracy using ImageJ software.

Manfredi et al. [24] conducted a study on a portable Fourier Transform Infrared (FTIR) spectrometer coupled with multivariate statistical analysis for the classification of hazelnut varieties. They used PCA followed by LDA and Partial Least Squares Discriminant Analysis (PLS-DA) as classification methods. They reported that the highest accuracy value was achieved with PLS-DA in the classification results.

Giraudo et al. [25] utilized RGB image analysis for the detection of defective hazelnuts. A total of 2000 images, half of which depicted defective hazelnuts and the other half depicted defect-free hazelnuts, were obtained. The images were transformed into a signal matrix of size 2000×4900 for the development of multivariate classification models. They employed PLS-DA as the classification model and achieved approximately 97% accuracy in detecting defective hazelnuts.

Solak and Altinsik [26] aimed to classify hazelnut fruits into small, medium, and large sizes using image processing techniques and clustering methods. As a result of the study, they mentioned that both mean-based and K-means clustering methods yielded classification results with a similarity of 90% or higher.

Dheir et al. [27] proposed a model using a Convolutional Neural Network (CNN) to classify 5 different types of nuts (Peanuts, Almonds, Pecan, Hazelnut, and Chestnuts). They reported achieving a classification accuracy of 98% in their study.

Farhadi et al. [28] proposed a method for classifying nuts as large/small, hollow, or damaged using audio signal processing and artificial neural network techniques. They designed a system that records the sound produced when a nut collides with a steel disk using a microphone placed beneath the disk. The extracted features from the recorded sound signals were used as input to a neural network, achieving a classification success rate of 89.3% or higher.

Koc et al. [29] utilized DL4J (Deep Learning for Java) and ensemble learning algorithms to classify three different hazelnut varieties (Sivri, Kara, Tombul). The study reported a classification accuracy of 95% for hazelnut varieties using Gradient Boost and DL4J while achieving 100% accuracy using the Random Forest (RF) algorithm.

Keles and Taner [30] aimed to classify 11 different hazelnut varieties by determining physical, mechanical, and optical properties along three main axes. They created separate models for each axis. Through the use of ANN and

Discriminant Analysis (DA) for classification, they achieved success rates of 89.1% and above for different axes.

Taner et al. [31] proposed a new model using Convolutional Neural Networks (CNN) to classify 17 different hazelnut varieties. They compared the results of the proposed CNN model, named Lprttnr1, with four pre-trained CNN models (VGG16, VGG19, ResNet50, and InceptionV3). According to their findings, the proposed model achieved an accuracy of 98.63%.

Ayvaz et al. [32] used a total of 280 hazelnut samples belonging to seven different hazelnut varieties grown in Turkey. These samples were collected in three physical conditions: in-shell, shelled, and white interior of split hazelnut kernels. They trained the collected spectra using PLS-DA and Logistic Models. As a result, they mentioned that both PLS-DA and Logistic Models based on all NIR spectra, with the preprocessing technique, were able to correctly determine unknown hazelnut varieties with 100% accuracy in all three physical conditions. Summary information from relevant studies in the literature is presented in Table 1.

Upon examining Table 1, it can be observed that the datasets utilized in the referenced studies contain a smaller number of data compared to the dataset employed in the present study. Moreover, it is noteworthy that the studies did not incorporate data fusion techniques. Furthermore, the combination of data fusion and feature reduction methods was not explored concurrently in the previous studies. Consequently, this study distinguishes itself from the existing literature by addressing these aforementioned aspects.

Table 1 Summary of studies found in the literature

No	Class	Data count	Method	Accuracy (%)	References
1	17	2040	EFA-LDA	6.7–83.2	[21]
2	3	900	YSA	93.57	[22]
3	3	30	ImageJ	84	[23]
4	3	60	PLS- DA	98	[24]
5	2	2000	PLS-DA	~97	[25]
6	3	150	Clustering	90 and above	[26]
7	5	2868	CNN	98	[27]
8	3	2400	MLP	89.3 and above	[28]
9	3	150	RF	100	[29]
10	17	510	CNN	98.63	[31]
11	11	550	ANN and DA	89.1 and above	[30]
12	3	280	PLS-DA	100	[32]

Material and methods

The implementation of deep learning methods in this research was conducted on the MATLAB platform. The operational procedure of the study is illustrated in Fig. 1, depicting the flow model. Furthermore, comprehensive details regarding the original dataset and the theoretical underpinnings of the employed system are provided. In addition, a comprehensive exposition is given regarding the methods, architectures, and other pertinent particulars utilized in the classification of the original dataset, as outlined below.

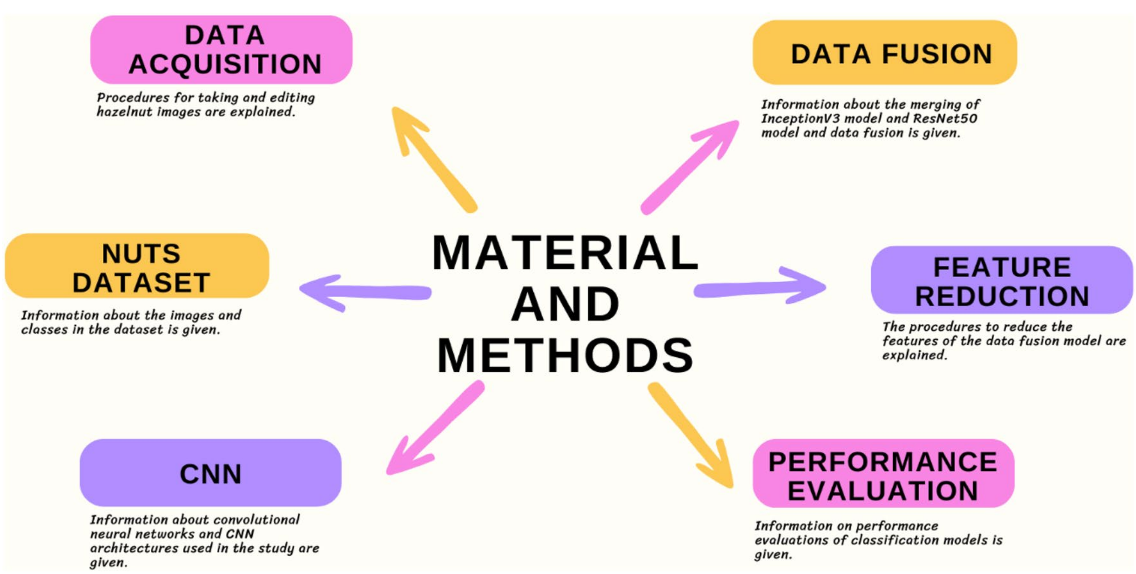


Fig. 1 Block diagram depicting the operational process of the study

Data acquisition

In this study, images of Giresun, Ordu, and Van hazelnut varieties were used. The hazelnut varieties used in the study were obtained from their respective regions in Turkey. Due to the unripe harvest time of these hazelnut varieties in Turkey, a limited diversity of varieties was ensured. The hazelnut images used in the study were obtained using the image acquisition system set up as shown in Fig. 2. To minimize variations such as angle, resolution, lighting adjustment, and background shadows in the generated images, a fixed lighting and camera system were preferred. The hazelnuts placed inside the box in an appropriate lighting environment were transferred to the computer system via a camera, as depicted in Fig. 2. This method not only saved time but also provided a cost-effective solution.

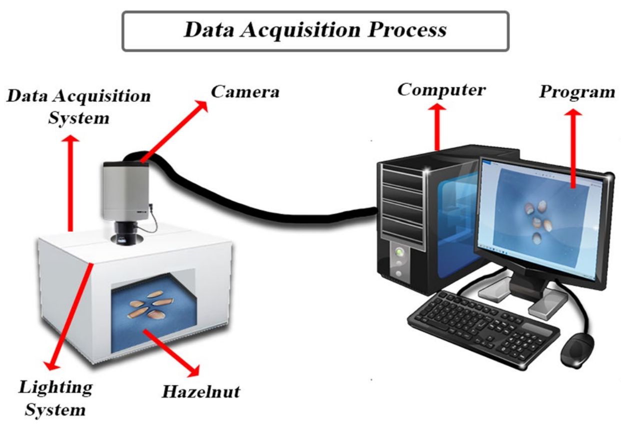


Fig. 2 Data acquisition process

Hazelnut dataset

The dataset consists of images of three different hazelnut varieties (Giresun, Ordu, and Van) obtained from regions where hazelnuts are grown in Turkey. The dataset contains a total of 3627 hazelnut images, with 1165 images for Giresun, 1324 images for Ordu, and 1138 images for Van. The backgrounds of the images have been converted to black. The images in this dataset have RGB color space, with dimensions of 2000 × 2000 pixels, and are in "jpg" format. As shown in Table 2, the dataset has been divided into random 80% training data and 20% testing data for each class, which are used as input for the models.

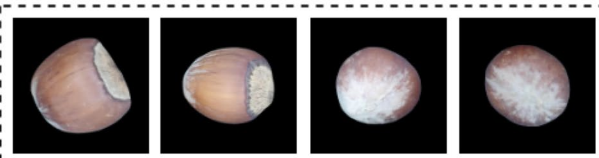
Convolutional neural networks

There are numerous types of CNN architectures in the literature. These architectures are developed by taking inspiration from the natural visual perception mechanisms of living organisms [33]. CNN architectures extract features through layers formed as a result of the given data inputs and learn and classify based on these features [34]. In a general overview, CNN architectures consist of five layers: the convolutional layer, pooling layer, activation layer, fully connected layer, and softmax layer [35]. When looking at the historical context, a multilayer artificial neural network called LeNet-5 was developed for classifying handwritten digits. This architecture has served as a precursor for many studies and has paved the way for new methods in the field [36].

ResNet50

Residual Networks (ResNet), developed by He Kaiming and his colleagues in 2015, is a structure with the highest

Table 2 Information for hazelnut dataset

HAZELNUT VARIETIES	EXAMPLE IMAGES	NUMBER OF IMAGES
GIRESUN		1165
ORDU		1324
VAN		1138

number of layers and consists of 152 layers in its initial implementation. When compared to ResNet-152, ResNet-50 has a smaller layer architecture. It has been trained on over one million images from the ImageNet database. ResNet-50 is a CNN and has a depth of 50 layers. It transfers the values required to reach the outcome through predictions from one layer to another. To improve the accuracy of these predictions, a block structure has been preferred, which enhances the quality of results generated within a network [37].

InceptionV3

In 2014, Szegedy and his colleagues proposed a new architecture called Inception based on the concept of "GoogLeNet" using the idea of poli depth. The aim was to increase the efficiency of computational resources. Inception consists of four sub-paths. Each convolutional layer and path have different operation layers. These (1×1 convolutions, 3×3 convolutions, 5×5 convolutions, and 3×3 max pooling) in InceptionV3, despite having 42 layers, are deeper and computationally more expensive compared to GoogLeNet, but they are much more efficient [38].

Data fusion

Data fusion is used to combine data from different sources to obtain more accurate results. This method is employed in various fields such as military, medical, environmental, agriculture, transportation, energy, security, nuclear, aviation, and many more. Data fusion can be used to achieve more accurate and comprehensive results by combining data from multiple diverse sources. Therefore, data fusion has become an important topic in areas such as data analytics, artificial intelligence, and machine learning. The data fusion approach can provide higher success rates than results obtained from a single data source. For example, in an image processing application, instead of using only one data source to classify an object, merging data from multiple sources (e.g., different sensors) can lead to more accurate results. This can improve the accuracy of object classification. The data fusion method can be employed especially when working with large and complex datasets to obtain more precise results by combining data from different sources. However, if not properly implemented, the results of data fusion can be misleading and lead to failures [39]. The main purpose of reclassification by reducing the features obtained from the data fusion method to a different number of features is to increase the classification success and to have the speed factor at the highest level necessary for using the models in real-world problems. In order to find the optimum model, a different number of features were used in the classification. The number of features at the exit of both models is 2048. In order to perform fusion in the data fusion method, the number of image features taken from the architectures must be the same. In addition, the reason why these models are preferred is that they are used continuously in the literature and their classification speed is fast because they are not very large models. While deciding on these models, many models were analyzed and the selection of these models was carried out in this way. In this study, the InceptionV3 and ResNet50 models were merged using the data fusion method and utilized in classification tasks.

CNN feature reduction

In CNNs, feature reduction is a type of feature extraction. This process transforms the network into a smaller-sized feature vector, enhancing efficiency and enabling better results by filtering out unnecessary information. Some benefits of feature reduction in CNNs include:

Reducing computational load: Feature reduction reduces the size of the network and decreases the computational load. This allows the network to operate faster and requires less processing power.

Preventing overfitting: A smaller-sized feature vector allows the network to focus on fewer details and can prevent overfitting. This helps the network to create a more generalized model and achieve better results.

Robustness to noise: Feature reduction makes the data more resilient to noise. This enables the network to perform better and produce more accurate results.

Dimensionality reduction: Feature reduction reduces the dimensionality of the data, resulting in a smaller storage requirement. This facilitates faster data loading and can lead to lower storage costs.

Feature reduction is particularly beneficial when working with large datasets. This method enables more efficient data processing and improves the quality of results. The filter size depends on how the network performs feature extraction. The filter size determines the dimensions of the matrix where the filters are applied by sliding over the feature map. A larger filter size will be applied to a wider area and require fewer filters, while a smaller filter size will be applied to a narrower area and require more filters. Changing the filter size in convolutional layers can also affect the number of extracted features [20]. In this study, filter numbers 32, 64, 128, 256, and 512 were used to reduce the network's features.

Performance evaluations

Confusion matrix

Confusion matrix is a tool used to evaluate the performance of classification models [40]. Classification models can sometimes make incorrect predictions while trying to predict the classes of examples. The confusion matrix is a matrix that visualizes the relationship between the true class and the class predicted by the model [41]. The confusion matrix represents the relationship between the true class and the class predicted by the model in four fundamental categories:

True positive (TP): The true class is positive, and the model also predicted it as positive.

False positive (FP): The true class is negative, but the model predicted it as positive incorrectly.

True negative (TN): The true class is negative, and the model correctly predicted it as negative.

False negative (FN): The true class is positive, but the model predicted it as negative incorrectly.

The confusion matrix is used to evaluate the performance of a classification model and calculate various performance metrics [42]. Performance metrics such as precision, recall, F1 score, and accuracy are calculated based on the TP, FP, TN, and FN values in the confusion matrix [43]. A three-class confusion matrix with its values is shown in Fig. 3.

Performance metrics

The basic performance metrics used for evaluating classification performance with machine learning methods are as follows:

Accuracy: Accuracy is a performance metric used to measure the performance of a model in classification problems using machine learning methods. This metric measures the ratio of correctly classified samples by the model. Accuracy is one of the fundamental performance metrics in classification problems. However, this metric can be misleading in some cases. For example, if the class distribution in the dataset to be classified is imbalanced, the accuracy performance metric can provide misleading results. In this case, the model can classify every sample as the negative class, resulting in a high accuracy value, while misclassifying samples from the positive class would be a significant error. Equation 1 shows the formula for accuracy [44].

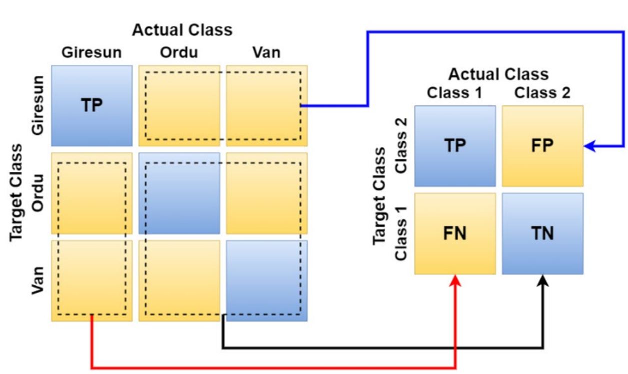


Fig. 3 Three classed confusion matrix

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

Precision: Precision is a performance metric used to evaluate the performance of a model in classification problems using machine learning methods. This metric measures the proportion of the samples that the model classifies as positive that are actually positive. Precision is particularly important in applications where the cost of false positives is high. When the cost of false-positive predictions is low and the cost of false-negative predictions is high, the precision performance metric can misleadingly give high results. Therefore, the precision performance metric should be used in conjunction with the recall performance metric. Recall measures how well the model can correctly predict the proportion of true positive examples out of the total positive examples. When these two metrics are used together, it is possible to measure the model's success in finding true positives and the number of false-positive predictions. Equation 2 shows the formula for precision [13].

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{2}$$

Recall: Recall is a performance metric used to evaluate the performance of a model in classification problems using machine learning methods. It measures how well the model can correctly predict the proportion of true positive examples out of the total positive examples [45]. Equation 3 shows the formula for recall.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{3}$$

F1 score: The F1 score metric is used to evaluate the performance of a model in classification problems using machine learning methods. This metric combines the measurements of precision and recall obtaining a single performance measure. Precision measures how well the model can correctly predict the proportion of true positive examples out of the total positive examples, while recall measures how well the model can correctly predict the proportion of true positive examples out of all the actual positive examples. The F1 score combines these two measurements to obtain a performance measure. To achieve the best result, both precision and recall should be high. The F1 score is particularly useful in imbalanced classification problems where there is

Table 3 Parameters used in model training

Parameter	Value
Validation frequency	5
Max epochs	8
Mini batch size	11
Initial learn rate	0.0001
Solver	sgdm
L2 regularization	0.0001
Execution environment	GPU

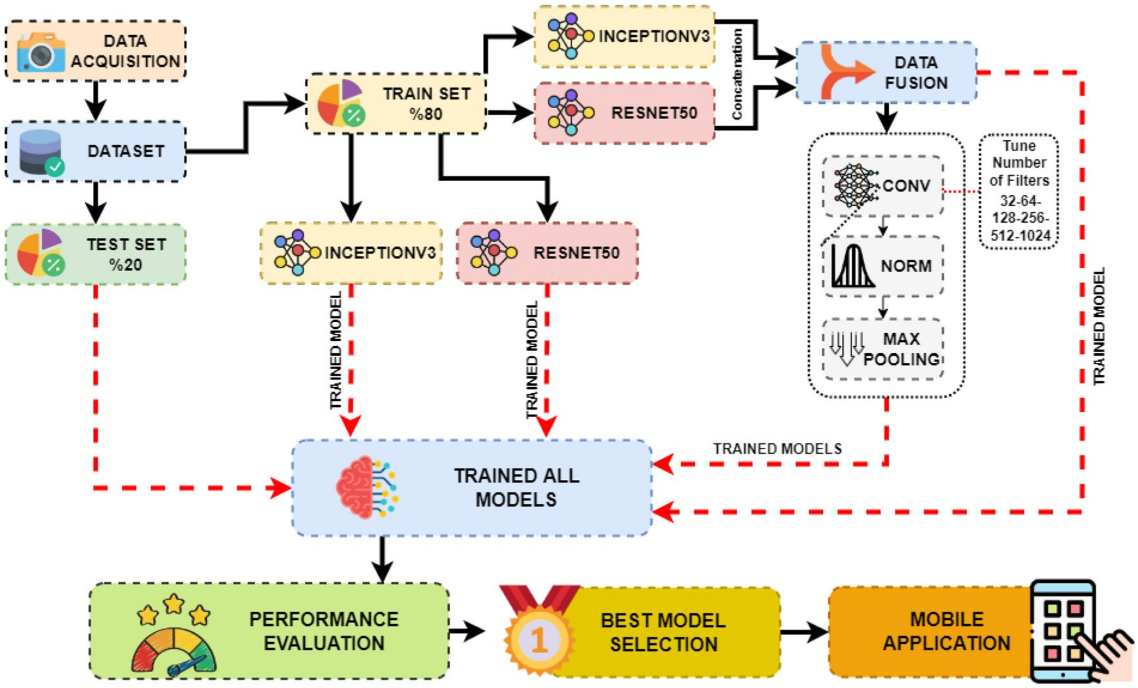


Fig. 4 Flow chart of the operations performed in the study

Fig. 5 Confusion matrix of InceptionV3, ResNet50, and data fusion models

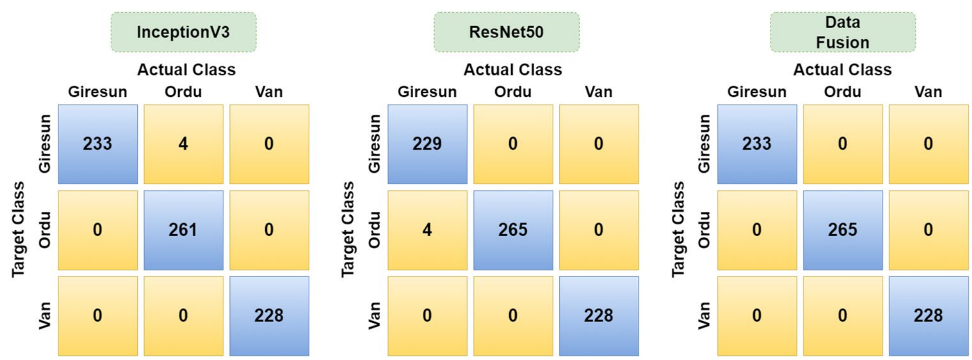


Table 4 Performance metrics of InceptionV3, ResNet50, and data fusion models by class

	InceptionV3			ResNet50			Data fusion		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Giresun	0.9931	1	0.9915	1	0.9828	0.9913	1	1	1
Ordu	1	0.9849	0.9924	0.9851	1	0.9925	1	1	1
Van	1	1	1	1	1	1	1	1	1

a noticeable difference in the number of examples between classes. In such cases, accurately classifying the minority class examples is more important, and the F1 score is a more suitable measure for evaluating performance, taking this imbalance into account. The F1 score formula is shown in Eq. 4.

Table 5 Inception V3, ResNet50, and data fusion models general performance metrics

	InceptionV3	ResNet50	Data fusion
Accuracy	0.9945	0.9945	1
Misclassification rate	0.0055	0.0055	1
Macro F1 score	0.9946	0.9946	1
Weighted F1 score	0.9945	0.9945	1

$$F1\ \text{Score} = 2 * (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall}) \tag{4}$$

These performance metrics can be used in different ways depending on the characteristics of the classification problem. For example, in a health diagnosis problem where precision is crucial, recall may be more important. In addition, these performance metrics can be used to compare different classification models.

The Macro F1 score is a metric where the F1 scores for each class are calculated separately, and they have equal weights across the classes. The F1 score is calculated for each class, and then the arithmetic mean of these scores is taken. The Macro F1 score evaluates the performance of each class independently and is a metric to consider in the case of class imbalance. On the other hand, the Weighted F1 score calculates a weighted F1 score considering class imbalance. The F1 score for each class is weighted based on the proportion of samples in that class. This ensures that larger classes have a greater impact in the case of class imbalance. The Weighted F1 score is commonly used to evaluate performance on datasets with class imbalance. In summary, while the Macro F1 score gives equal weight to all classes, the Weighted F1 score takes into account class imbalance and performs weighting based on the classes [46]. The misclassification rate is used to measure how correct or incorrect the predictions of the classification model are. It is usually expressed as a percentage and is a measure of how well or poorly the model is performing. A low rate of misclassification may mean that the model performs better, while a high rate may indicate that the model is of lower quality [47]. The misclassification rate formula is shown in Eq. 5.

$$\text{Misclassification rate} = (\text{FP} + \text{FN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{5}$$

Experimental results

In this section, a series of data fusion methods based on two different CNN architectures have been analyzed for the classification of hazelnut varieties from hazelnut images. The aim was to determine the optimal classification accuracy and classification time for the classification of images and use them in practical applications. The models used in the study were implemented in MATLAB 2021b environment. The specifications of the computer used were Intel® Core i7™ 12700 K 3.61 GHz, NVIDIA GeForce RTX 3080Ti, and 64 GB RAM. The parameters used in the training of the models are given in Table 3. The dataset was divided into training and testing sets with a ratio of 80% for training and 20% for testing. In datasets where the number of data in the dataset is sufficient and the number of classes is balanced, dividing at this rate is a frequently used ratio to measure the success of the models. The flowchart of the operations performed in the study is shown in Fig. 4.

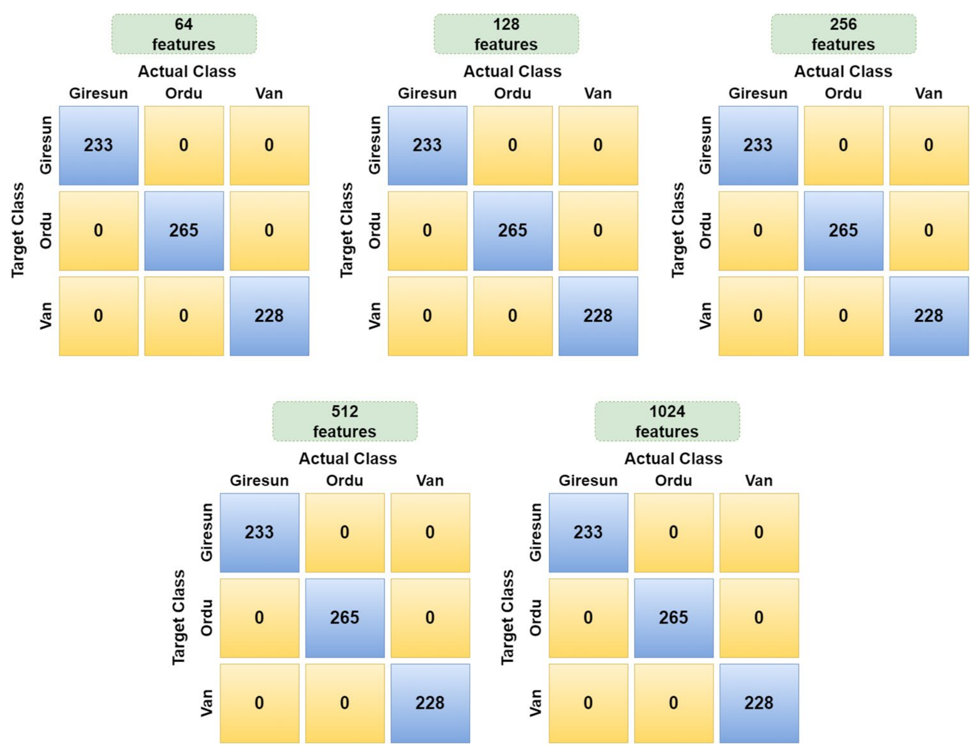


Fig. 6 Confusion matrices of 64, 128, 256, 512, and 1024 reduced features data fusion models

Table 6 Performance metrics of Inceptionv3, ResNet50, and data fusion models by class

	64 features			128 features			256 features			512 features			1024 features		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Giresun	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ordu	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Van	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

After the data acquisition processes, as seen in Fig. 4, images were preprocessed to create a dataset with three classes. The dataset consists of a total of 3627 images, with 1165 images from the Giresun class, 1324 images from the Ordu class, and 1138 images from the Van class. In the study, eight different classification processes were performed. The data were classified using the InceptionV3 and ResNet50 models. The data fusion method was employed to combine the ability of InceptionV3 and ResNet50 models to extract different features, aiming to obtain a more powerful classification model. The features of the models were concatenated at the layer just before the final fully connected layer. Each model extracts 2048 features from each image. These concatenated features, resulting from the data fusion method, were fed into the fully connected layer as input for the classification process. The confusion matrices obtained from InceptionV3, ResNet50, and data fusion models are shown in Fig. 5.

According to Fig. 5, the InceptionV3 model misclassified 4 hazelnut images from the Ordu class as Giresun. On the other hand, the ResNet50 model misclassified 4 hazelnut images from the Giresun class as Ordu. However, the data fusion model obtained by concatenating these two models correctly classified all hazelnut images. The precision, recall, and F1 Score performance metrics calculated based on the confusion matrix data from Fig. 5 are shown in Table 4 for the respective classes.

According to Table 4, the data fusion model achieved a perfect score of 1 for all metrics. This is because it did not misclassify any of the images. Similarly, both the InceptionV3 and ResNet50 models achieved a score of 1 for the performance metrics of the Van class. This is because the images from the Van class were correctly classified without any confusion with images from other classes. The accuracy, misclassification rate, macro F1 score, and Weighted F1 score values for the InceptionV3, ResNet50, and data fusion models are shown in Table 5.

According to Table 5, the data fusion model has higher values compared to the InceptionV3 and ResNet50 models. Concatenating the features of the two models has improved the classification performance.

In this study, the reduction of features obtained through data fusion was performed using a convolutional layer, batch normalization layer, and max pooling layer. By reducing the number of filters in the convolutional layer, different numbers of features were obtained. With 32 filters, the 2048 features can be reduced to 64 features. The result obtained from this is that twice the number of features can be obtained as the number of filters. By setting the values of 32, 64, 128, 256, and 512 as parameters, 5 different classifications were performed with 64, 128, 256, 512, and 1024 features. The confusion matrices obtained from all classification processes are given in Fig. 6.

Table 7 InceptionV3, ResNet50, and data fusion models general performance metrics

	64 features	128 features	256 features	512 features	1024 features
Accuracy	1	1	1	1	1
Misclassification rate	0	0	0	0	0
Macro F1 score	1	1	1	1	1
Weighted F1 score	1	1	1	1	1

Table 8 Classification times of all models used in the study

Model	Features	Accuracy (%)	Classification time
InceptionV3	2048 features	99.45	116 min 17 s
ResNet50	2048 features	99.45	102 min 23 s
Data fusion	2048 features	100	120 min 38 s
Data fusion	64 features	100	93 min 33 s
	128 features	100	93 min 49 s
	256 features	100	115 min 4 s
	512 features	100	117 min 31 s
	1024 features	100	119 min 26 s

According to Fig. 6, the classification results obtained from reducing the features of the data fusion model using a convolutional layer with different filter sizes for the 2048 features are the same. In the classifications performed with 64, 128, 256, 512, and 1024 features, all models have the same classification accuracy and performance metrics. Table 6 shows the performance metrics for the reduced features data fusion models according to the classes, and Table 6 shows the accuracy, misclassification rate, macro F1 score, and weighted F1 score values for these models.

According to Table 7, the classifications performed by reducing the 2048 features of the data fusion model have resulted in the same outcome for all cases. At first glance, one might question why these models were run repeatedly. However, for a detailed analysis of the models, classification processes were conducted using all features. The classification times and performance of all models used in the study are provided in Table 8.

According to Table 8, all models except InceptionV3 and ResNet50 have achieved a classification success rate of 100%. However, the classification times vary depending on the number of features used for classification. The data fusion model has the highest classification time, while the 64 reduced features data fusion model has the lowest classification time. Fast operation of classification models is essential for their use in mobile applications. Once the optimal model is identified, it should be embedded within the mobile application to perform the necessary operations. Upon careful examination of all models, the 64 reduced features data fusion model stands out as the most optimal model. This model was selected and used for the mobile application developed in the study.

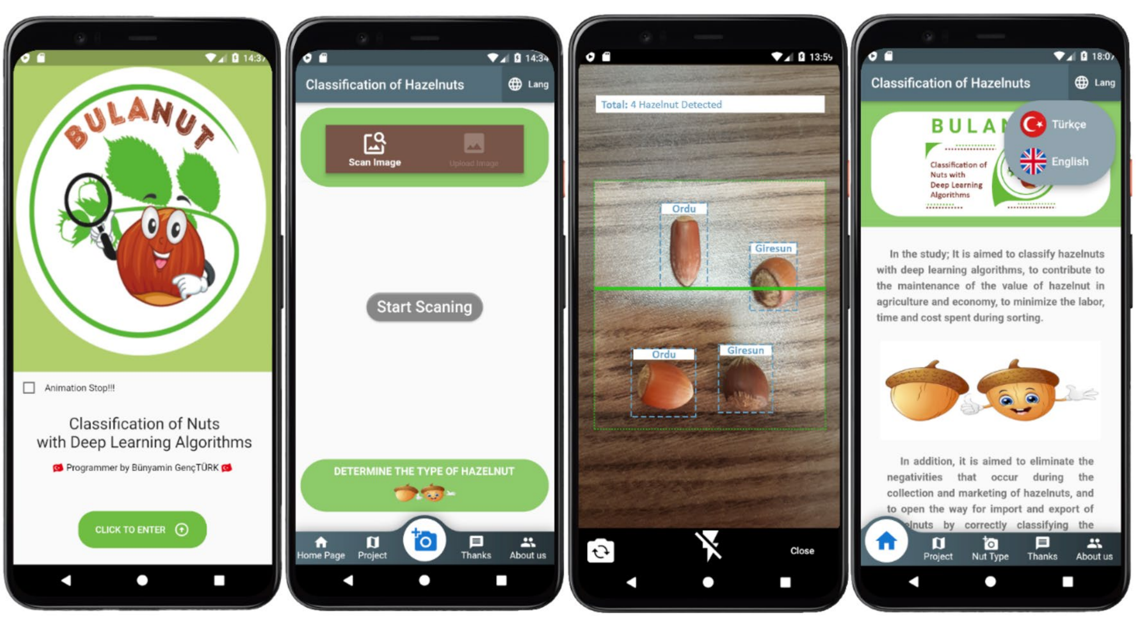


Fig. 7 The mobile application developed for the classification of hazelnuts: BULANUT

Mobile applications are software programs that run on smartphones, tablets, or other mobile devices. The main purpose of mobile applications is to provide users with a specific functionality or service through their mobile devices. Mobile applications are designed to enhance user experience and functionality. The primary goals of mobile applications are functionality, user experience, and efficiency. Due to these advantages, a mobile application has been developed in order to make the models in the study accessible to users. The mobile application is named BULANUT. In the application, users can classify hazelnut varieties using the BULANUT application installed on their Android smartphones. Figure 7 displays screenshots of the mobile application.

Conclusions

Hazelnut varieties generally have similar appearances, making manual classification a difficult and time-consuming task. In addition, the hazelnut industry is rapidly growing, requiring more effective and efficient classification methods. To address these problems, this study created a dataset containing images of three different hazelnut varieties. Various image classification methods were analyzed to classify the images in the dataset. InceptionV3 and ResNet50 architectures were used as the basis for hazelnut image classification in this study. Classification processes were performed using these models, and a data fusion model was created by combining them to classify the images based on their 2048 features. Feature reduction layers were added to the Data Fusion model to decrease classification time, and classification processes were performed using the data fusion model at five different feature sizes. The analysis of all classifications resulted in the highest classification accuracies being achieved by models other than InceptionV3 and ResNet50. 100% classification accuracy was obtained from these models. The data fusion model based on 64 reduced features had the lowest classification time. Therefore, this model was selected for use in the mobile application. The developed mobile application allows for the classification of hazelnut images belonging to three different classes. As a result, the feasibility of the proposed models and mobile applications in the field has been established.

The proposed models and mobile applications can contribute to the development of more effective classification and quality control processes in the hazelnut industry. They can also provide more information about the characteristics and qualities of different hazelnut varieties and help improve production processes. The identification of hazelnut varieties enables faster and more automated harvesting processes, reducing time, and labor costs. As a result, productivity can be increased. The identification of hazelnut varieties can play an important role in quality control. Determining the quality and homogeneity of hazelnut varieties can provide a competitive advantage in exports. These models can be revised to be used in the detection of diseases in hazelnut trees. Early diagnosis of hazelnut diseases can help prevent the spread of diseases.

These methods can also provide hazelnut farmers with the opportunity to collect and analyze data. As a result, hazelnut farmers can make better decisions and manage their products more efficiently. It enables precision agriculture applications in hazelnut orchards, allowing the determination of each tree's needs and intervention when necessary. This increases product efficiency and makes agricultural activities more sustainable.

Data availability The dataset in this link can be used by citing the source https://www.muratkoklu.com/datasets/Hazelnut_Images_Dataset.zip.

Declarations

Conflict of interest The authors declare that they have no competing interests.

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

References

1. Tercan E, Dengiz O, Özkan B, Dereli MA, Öztekin YB (2022) Geographic information system–assisted site quality assessment for hazelnut cultivation using multi-criteria decision analysis in the Black Sea region. Turkey Environ Sci Pollut Res 29(24):35908–35933
2. Król K, Gantner M (2020) Morphological traits and chemical composition of hazelnut from different geographical origins: a review. Agriculture 10(9):375
3. Köksal Aİ, Artik N, Şimşek A, Güneş N (2006) Nutrient composition of hazelnut (*Corylus avellana* L.) varieties cultivated in Turkey. Food Chem 99(3):509–515
4. Yıldırım Ç, Türkten H, Boz İ (2022) Assessing the sustainability index of part-time and full-time hazelnut farms in Giresun and Ordu Province. Turkey Environ Sci Pollut Res 29(52):79225–79240
5. Secer A (2008) An investigation on Turkish hazelnut export concentration. J Appl Sci Res 4(11):1557–1560
6. Şahinli MA, Özçelik A (2016) Fındık arz fonksiyonu tahmin modeli: Türkiye üzerine ekonometrik bir uygulama. Nevşehir Bilim ve Teknoloji Dergisi 5:79–88
7. Yenisu E (2017) TÜRK FINDIK SEKTÖRÜNÜN REKABET GÜCÜ: BALASSA ENDEKSİ YAKLAŞIMI. Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi 3(5):22–37
8. Battistoni C, Dominici L, Barbero S, Comino E (2020) Systemic design methodology applied to hazelnut processing. Int J Des Sci Technol 24(1):39–65
9. Islam A (2019) Fındık ıslahında gelişmeler. Akademik Ziraat Dergisi 8:167–174

10. Hosseinpour A, Seifi E, Javadi D, Ramezanpour SS, Molnar TJ (2013) Nut and kernel characteristics of twelve hazelnut cultivars grown in Iran. Sci Hortic 150:410–413
11. Ozdemir F, Akinci I (2004) Physical and nutritional properties of four major commercial Turkish hazelnut varieties. J Food Eng 63(3):341–347
12. Ghisoni S, Lucini L, Rocchetti G, Chiodelli G, Farinelli D, Tombesi S, Trevisan M (2020) Untargeted metabolomics with multivariate analysis to discriminate hazelnut (*Corylus avellana* L.) cultivars and their geographical origin. J Sci Food Agric 100(2):500–508
13. Koklu M, Kursun R, Taspinar YS, Cinar I (2021) Classification of date fruits into genetic varieties using image analysis. Math Probl Eng 2021:1–13
14. Sabanci K, Koklu M, Unlersen MF (2015) Classification of Siirt and long type pistachios (*Pistacia vera* l.) by artificial neural networks. Int J Intell Syst Appl Eng 3(2):86–89
15. Örnek MN, Kahramanlı Örnek H (2021) Developing a deep neural network model for predicting carrots volume. J Food Measurement Charact 15(4):3471–3479
16. Mulla NA, Ravichandran S, Naik PK, Balappa B (2022) Computer Vision System to detect maturity of tomatoes in real time using deep learning. Advances in distributed computing and machine learning learning: Proceedings of ICADCML 2021. Springer
17. Koklu M, Ozkan IA (2020) Multiclass classification of dry beans using computer vision and machine learning techniques. Comput Electron Agric 174:105507
18. Koklu M, Cinar I, Taspinar YS (2021) Classification of rice varieties with deep learning methods. Comput Electron Agric 187:106285
19. Taspinar YS (2023) Light weight convolutional neural network and low-dimensional images transformation approach for classification of thermal images. Case Stud Therm Eng 41:102670
20. Singh D, Taspinar YS, Kursun R, Cinar I, Koklu M, Ozkan IA, Lee H-N (2022) Classification and analysis of pistachio species with pre-trained deep learning models. Electronics 11(7):981
21. Sayinci B, Kara M, Ercişli S, Duyar Ö, Ertürk Y (2015) Elliptic Fourier analysis for shape distinction of Turkish hazelnut cultivars. Erwerbs-obstbau 57(1):1–11
22. Guvenc SA, Senel FA, Cetisli B (2015) Classification of processed hazelnuts with computer vision. In: 2015 23nd Signal Processing and Communications Applications Conference (SIU). IEEE.
23. Bayrakdar S, Çomak B, Başol D, Yücedag İ (2015) Determination of type and quality of hazelnut using image processing techniques. In: 2015 23nd Signal Processing and Communications Applications Conference (SIU). IEEE.
24. Manfredi M, Robotti E, Quasso F, Mazzucco E, Calabrese G, Marengo E (2018) Fast classification of hazelnut cultivars through portable infrared spectroscopy and chemometrics. Spectrochim Acta Part A Mol Biomol Spectrosc 189:427–435
25. 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
26. Solak S, Altinişik U (2018) Detection and classification of hazelnut fruit by using image processing techniques and clustering methods. Sakarya Univ J Sci 22(1):56–65
27. Campmajó G, Saez-Vigo R, Saurina J, Núñez O (2020) High-performance liquid chromatography with fluorescence detection fingerprinting combined with chemometrics for nut classification and the detection and quantitation of almond-based product adulterations. Food Control 114:107265
28. Farhadi M, Abbaspour-Gilandeh Y, Mahmoudi A, Mari Maja J (2020) An integrated system of artificial intelligence and signal processing techniques for the sorting and grading of nuts. Appl Sci 10(9):3315
29. Caner K, Gerdan D, EmİNoĞLu MB, YegüL U, Bulent K, VatandaŞ M (2020) Classification of hazelnut cultivars: comparison of DL4J and ensemble learning algorithms. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 48(4):2316–2327
30. Keles O, Taner A (2021) Classification of hazelnut varieties by using artificial neural network and discriminant analysis. Span J Agric Res 19(4):e0211–e0211
31. Taner A, Öztekin YB, Duran H (2021) Performance analysis of deep learning CNN models for variety classification in hazelnut. Sustainability 13(12):6527
32. Ayvaz H, Temizkan R, Genis HE, Mortas M, Genis DO, Dogan MA, Nazlim BA (2022) Rapid discrimination of Turkish commercial hazelnut (*Corylus avellana* L.) varieties using near-infrared spectroscopy and chemometrics. Vib Spectrosc 119:103353
33. Karasulu B, Yücalar F, Borandağ E (2022) A hybrid approach based on deep learning for gender recognition using human ear images. J Faculty Eng Archit Gazi Univ 37(3):1579–1594
34. Taspinar YS, Selek M (2020) Object recognition with hybrid deep learning methods and testing on embedded systems. Int J Intell Syst Appl Eng 8(2):71–77
35. Butuner R, Cinar I, Taspinar YS, Kursun R, Calp MH, Koklu M (2023) Classification of deep image features of lentil varieties with machine learning techniques. Eur Food Res Technol 249(5):1303–1316
36. Dogan M, Taspinar YS, Cinar I, Kursun R, Ozkan IA, Koklu M (2023) Dry bean cultivars classification using deep cnn features and salp swarm algorithm based extreme learning machine. Comput Electron Agric 204:107575
37. Ozaltin O, Yeniay O, Subasi A (2023) OzNet: a new deep learning approach for automated classification of COVID-19 computed tomography scans. Big Data. https://doi.org/10.1089/big.2022.0042
38. Ramaneswaran S, Srinivasan K, Vincent PDR, Chang C-Y (2021) Hybrid inception v3 XGBoost model for acute lymphoblastic leukemia classification. Comput Math Methods Med 2021:1–10
39. Gong W, Chen H, Zhang Z, Zhang M, Wang R, Guan C, Wang Q (2019) A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-SVM and multichannel data fusion. Sensors 19(7):1693
40. Feyzioglu A, Taspinar YS (2023) Beef quality classification with reduced E-nose data features according to beef cut types. Sensors 23(4):2222
41. Ozaltin O, Coskun O, Yeniay O, Subasi A (2022) A deep learning approach for detecting stroke from brain CT images using OzNet. Bioengineering 9(12):783
42. Yucalar F, Ozcift A, Borandag E, Kilinc D (2020) Multiple-classifiers in software quality engineering: combining predictors to improve software fault prediction ability. Eng Sci Technol Int J 23(4):938–950
43. Kishore B, Yasar A, Taspinar YS, Kursun R, Cinar I, Shankar VG, Koklu M, Ofori I (2022) Computer-aided multiclass classification of corn from corn images integrating deep feature extraction. Comput Intell Neurosci. https://doi.org/10.1155/2022/2062944
44. Borandağ E, Özçift A, Kaygusuz Y (2021) Development of majority vote ensemble feature selection algorithm augmentedwith rank allocation to enhance Turkish text categorization. Turkish J Electr Eng Comput Sci 29(2):514–530
45. Sofaer HR, Hoeting JA, Jarnevich CS (2019) The area under the precision-recall curve as a performance metric for rare binary events. Methods Ecol Evol 10(4):565–577
46. Guibon G, Ermakova L, Seffih A, Firsov, Noe-Bienvenu GL. Multilingual fake news detection with satire. In: Computational Linguistics and Intelligent Text Processing: 20th International Conference, CICLing 2019, La Rochelle, France, April 7–13, 2019, Revised Selected Papers, Part II. 2023. Springer.

47. Naseer I, Akram S, Masood T, Jaffar A, Khan MA, Mosavi A (2022) Performance analysis of state-of-the-art cnn architectures for luna16. Sensors 22(12):4426