Volume 48 | Number 1 Article 9

2-1-2024

Classification of hybrid chestnut cultivars (Castanea sativa) registered in Türkiye with artificial neural networks, based on some physical properties of their nuts

TANER YILDIZ
tyildiz@omu.edu.tr

Turkish Journal of Agriculture and Forestry
http://journals.tubitak.gov.tr/agriculture/
Research Article
Turk J Agric For
(2024) 48: 95-101
© TÜBİTAK
doi:10.55730/1300-011X.3165

Classification of hybrid chestnut cultivars *(Castanea sativa)* registered in Türkiye with artificial neural networks, based on some physical properties of their nuts

Taner YILDIZ*
Department of Agricultural Machinery and Technologies Engineering, Faculty of Agriculture,
Ondokuz Mayıs University, Samsun, Turkiye

Received: 25.06.2023 • **Accepted/Published Online:** 07.11.2023 • **Final Version:** 01.02.2024

Abstract: Understanding how the classification and identification of biological species can evaluate improvements in newly developed cultivars, including chestnuts (*Castanea sativa*), is crucial for product processing and equipment design. To evaluate this, in the present study artificial neural networks (ANNs) were used to characterize four hybrid chestnut cultivars (Macit 55, Akyüz, and Ali Nihat registered in Türkiye and Bouche de Betizac registered in France). A backpropagation neural networks algorithm was used in the ANN approach based on nine physical properties. These properties included shelled nut weight and volume, sphericity, geometric mean diameter, bulk density, surface area, true density, porosity, and length, which can be deemed for classifying the cultivars. The ANN model was composed of input (9), hidden (6-5), and output (1) layers. In the hidden layers and output layer, tansig transfer and linear transfer functions were used, respectively. The R^2 value for the test and training data was 0.99999 (RMSE = 0.000083 and 0.0023, respectively). The relative error (ε) between the real values and the estimated values was 0.079%. In conclusion, the ANN approach is able to discriminate among Macit 55, Akyüz, Ali Nihat, and Bouche de Betizac accessions based on the values of R^2 and ε.

Key words: Backpropagation, biological species classification, crop properties, statistical pattern technique

1. Introduction

Chestnut (*Castanea sativa* Mill.) production has increased dramatically in Türkiye (i.e., the Aegean, Marmara, and Black Sea regions), as in many countries around the world (Serdar et al., 2018; Yıldız and Cevher, 2022). The features such as kernel size, taste, texture, and starch and sugar contents, which vary according to chestnut genotypes, affect the technological quality of chestnut fruits consumed in fresh or dried form (Serdar et al., 2018; Massantini et al., 2021). Thus, it is essential to reliably and accurately determine and classify wild, hybrid, and cultivated chestnut varieties or genotypes.

The artificial neural network (ANN) approach, inspired by the biological nervous systems of humans (Keskenler and Keskenler, 2017), is used as a reliable tool to classify chestnut varieties based on the physical and mechanical properties of fruits (or nuts) or leaves (Mancuso et al., 1999; Öztekin et al., 2020). Moreover, this approach helps construct a mathematical function from the relationship between inputs and output for classifying and describing crops or biological species (Visen et al., 2002; Pandolfi et al., 2009; Keskenler and Keskenler, 2017).

In Türkiye, the many hybrid chestnut cultivars developed, mostly heterogeneous and of complex characterization (Mancuso et al., 1999), have been selected for further studies and breeding (Serdar et al., 2018). There are 17 varieties in Türkiye (Serdar et al., 2018; Öztekin et al., 2020), all of which are registered by the Ministry of Agriculture and Forestry, Variety Registration and Seed Certification Center. In this selection, while sometimes cultivars for roasting and producing chestnut candy are preferred, other times seedlings and clonal trees have been used for chestnut production or tolerating chestnut diseases (Serdar et al., 2018). Understanding how the classification and identification of biological species can evaluate improvements in newly developed cultivars, including chestnuts, is crucial for product processing and equipment design. However, there is little information on the classification of nuts from hybrid chestnut cultivars registered in Türkiye regarding product processing and equipment design. Therefore, the present study aimed to evaluate the classification regarding product processing and equipment design of three hybrid cultivars (Macit 55, Akyüz, and Ali Nihat) registered in Türkiye and one standard cultivar (Bouche de Betizac registered in France) using the ANN approach.

* Correspondence: tyildiz@omu.edu.tr

2. Materials and methods

2.1. Materials

Nuts from four hybrid chestnut cultivars were used: (1) Macit 55, owned by Limited Company of Academic Agriculture registered by Samsun Black Sea Agricultural Research Institute, (2) Akyüz, registered by Ondokuz Mayıs University, (3) Ali Nihat, registered by Technology Transfer Office of Ondokuz Mayıs University, and (4) Bouche de Betizac, a French chestnut cultivar developed by the Institut National de la Recherche Agronomique (INRA) at the station of Malemort-sur-Corrèze (Figure 1).

2.2. Methods

2.2.1. Collection of samples and measurements

The chestnut nut samples were collected from five trees belonging to each cultivar during the harvest season (October to November) at a research station located in the Black Sea Region (41°21´55″N, 36°11´14″E; 190 m above sea level) in Türkiye. All nut samples that were free from broken, damaged, and immature nuts as well as foreign matter (Öztekin et al., 2020) were stored (at 0 °C and 75%–85% humidity) in perforated nylon bags until measurements. The physical properties of the samples were measured in the Biological Materials Testing Laboratory (Department of Agricultural Machinery and Technologies Engineering, Ondokuz Mayıs University, Faculty of Agriculture). The physical properties regarding product processing and equipment design belonged to four cultivars that exhibited differences in shelled nut weight (SNW) and volume (SNV), sphericity, geometric mean diameter (GMD), bulk density (BD), true density (TD), surface area (SA), true density (TD), porosity, and length.

To determine SNW, each nut was weighed using an electronic weighing scale with a precision of 0.01 g (Mettler Toledo). The nut's length, width, and thickness were measured with a digital caliper (150 mm range, linear tools) with a precision of 0.01 mm. The GMD (1),

(A)

(B)

(C)

(D)

Figure 1. (A) The nuts from hybrid Macit 55, (B) Akyüz, (C) Ali Nihat, and (D) Bouche de Betizac cultivars.

sphericity (2), V (3), SNV (4), and BD (5) of the nuts were calculated using the following equations (Öztekin et al., 2020):

$$\mathrm{GMD} = (LWT)^{1/3} \tag{1}$$

$$\emptyset = \frac{(\mathrm{GMD})^{1/3}}{\mathrm{L}} 100 \tag{2}$$

$$V = \left[\frac{\Pi B^2 L^2}{6\,(2\mathrm{L} - \mathrm{B})}\right] \tag{3}$$

$$\mathrm{SNV} = \left[\frac{\Pi B\, L^2}{2\mathrm{L} - \mathrm{B}}\right] \tag{4}$$

$$BD = (WT)^{1/2} \tag{5}$$

2.2.2. Artificial neural networks approach

The ANN computer programs use some classified or learned knowledge with the help of neural sensors to make decisions (Keskenler and Keskenler, 2017). Thus, the ANN approach, trained with specific examples, produces an output (the level where it can generalize and make decisions) in response to the information given to the program as an input set (Figure 2).

MATLAB NN Toolbox software was used for developing the ANN approach using 400 item of data that were normalized (Equation 6) between 0 and 1 (Purushothaman and Srinivasa, 1994).

$$y_{nor} = \frac{y - y_{min}}{y_{max} - y_{min}} \tag{6}$$

To obtain the actual values from the normalized values, the 'y_{nor}' value calculated was used. The data used in the study were divided into two sets: training (n = 32) and test (n = 8) datasets.

In the developed ANN approach with feed-forward backpropagation, a multilayer perceptron network structure (Jacobs, 1988; Bekesiene et al., 2021), while the physical properties were used as the input parameters, the cultivar was used as the output parameter. In the approach, the Levenberg–Marquardt algorithm was used as the training algorithm, as explained previously (Kalogirou, 2001; Öztekin et al., 2020). The performance of the approach in terms of training was determined by root squared mean squared error (RMSE) and determination coefficient (R^2), calculated with Equations 7 and 8, respectively (Bechtler et al., 2001). Moreover, the relative error (ε) between the real and estimated values was determined by Equation 9.

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(z_{1i} - z_i)^2\right)^{1/2} \tag{7}$$

$$R^2 = 1 - \left(\sum_{i=1}^{n}(z_{1i} - z_i)^2\right) \Big/ \left(\sum_{i=1}^{n}(z_{1i})^2\right) \tag{8}$$

$$\mathcal{E} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{(z_i - z_{1i})}{z_{1i}}\right| \tag{9}$$

Here n is the number of data items, z_i is the real value, and z_{1i} is the estimated value.

The construction, training, and testing stages of the ANN approach were performed as explained by Öztekin

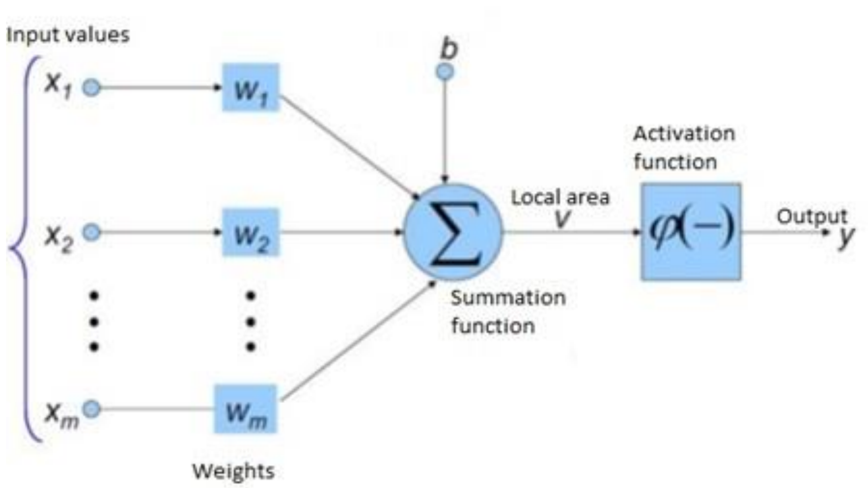

Figure 2. An exemplary artificial neural networks approach structure adapted from Bekesiene et al. (2021).

et al. (2020) and Bekesiene et al. (2021). To construct the ANN approach and examine their precision, IBM SPSS Statistics v21 was used.

3. Results and discussion

There were differences among the GMD (between 25.29 and 32.58 mm), length (between 32.41 and 38.27 mm), nut weight (between 7.27 and 17.57 g), sphericity (between 0.78 and 0.86%), SA (between 2010 and 3336 mm^2), BD (between 552.8 and 620.7 g cm^{-3}), and TD (between 1380 and 1808 g cm^{-3}) of the studied chestnut cultivars ($p < 0.05$). Macit 55 had the lowest length value, followed by Ali Nihat, Bouche de Betizac, and Akyüz. The lowest SNW, GMD, sphericity, and SA values were detected in Macit 55 and Bouche de Betizac, while the highest BD and TD values were in Ali Nihat-Akyüz and Macit 55-Akyüz.

The structure of the ANN approach was designed as [9-(6-5)-1], that is, 9 input layers, 2 hidden layers (6-5), and 1 output layer (Figure 3).

The purlin transfer and logsig transfer functions were used in the hidden layers, while the purlin transfer function was in the output layer. The training error of the approach, obtained at 100 epoch numbers, was the lowest level.

To calculate SNW, SNV, GMD, sphericity, SA, BD, TD, porosity, and length, the mathematical equations of the ANN approach were developed and are given below. The linear transfer function (Equation 10) for the second hidden layer (F_k) and the tansig transfer function (Equation 11) for the first hidden layer (F_j) were used.

$$y_m = \sum_k (W_3)_{k,m}.F_k + b_k \quad (10)$$

Linear transfer function for the second hidden layer (F_k),

$$F_k = NET_k \quad (11)$$

$$NET_k = \sum_j (W_2)_{j,k}.F_j + b_j \quad (12)$$

Tansig transfer function for the first hidden layer (*Fj*),

$$F_j = \frac{2}{(1 + e^{(-2NET_j)})} - 1 \quad (13)$$

$$NET_j = \sum_i (W_1)_{i,j}.x_1 + b_i \quad (14)$$

Here i is the number of inputs; j is the number of neurons in the first hidden layer; k is the number of neurons in the second hidden layer; m is the number of outputs; W_1, W_2, and W_3 are the connection weights; χ is the input parameter; y_m is the output parameter; and b is bias.

The weights (W) are presented in Tables 1–3. The bias values are given in Table 4. In the ANN approach, the R^2 for training and testing values was 0.9999 (RMSE = 0.00083 and 0.0023, respectively).

The R^2 of the association between the actual data and the test data calculated by ANN was 99.99% (Figure 4).

Although the mechanical and physical traits of the chestnuts have been widely explored in the literature (Mancuso et al., 1999; Feng et al., 2018; Öztekin et al., 2020), in the present study the aim was to predict which the studied physical traits were the most important in parameters of fundamentally changing complex hybrid

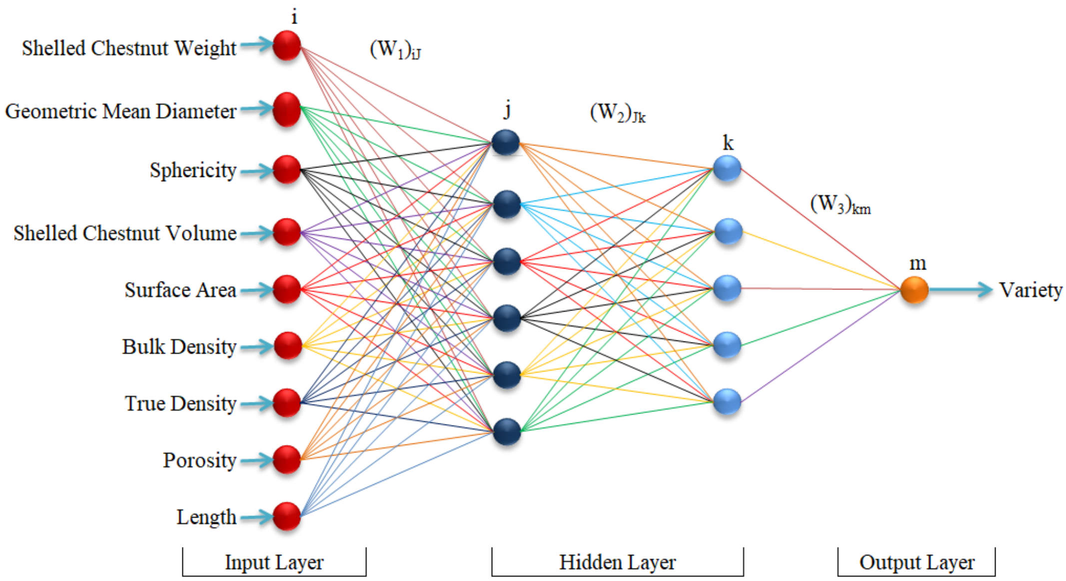

Figure 3. Scheme for the structure description of the artificial neural networks.

Table 1. Connection weight (W_3) values for Equation (10).

m	k_1	k_2	k_3	k_4	k_5
1	−0.04	0.2807	−0.0902	0.0818	−1.9312

Table 2. Connection weight (W_2) values for Equation (12).

k	j_1	j_2	j_3	j_4	j_5	j_6
1	−0.0633	0.4515	0.4806	0.6755	0.9323	1.2698
2	0.5551	−0.9632	−0.3051	0.2303	1.1396	−0.5596
3	1.2638	−0.1868	1.3037	−0.432	−0.1408	−0.3283
4	0.5108	−0.4077	0.2061	0.5882	1.1744	1.0564
5	0.8215	0.1295	−0.4486	0.4271	−1.0577	0.2797

Table 3. Connection weight (W_1) values for Equation (14).

j	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9
1	0.9911	0.0363	0.9598	0.2271	−0.6162	−0.5813	−0.0343	0.5028	−0.2513
2	−0.643	−0.7593	0.5336	1.0751	−0.3129	0.3739	0.5341	−0.1131	−0.7222
3	−0.4358	0.8984	0.6269	−0.5077	−0.4248	0.1912	−0.5641	1.0872	−0.1822
4	0.0239	−0.1105	−0.4378	−0.2211	0.2497	0.9143	0.4357	0.0384	0.2166
5	1.0946	0.1881	0.6222	0.3124	−0.6952	−0.0866	0.3773	0.0585	−0.1743
6	0.139	0.6963	0.8287	−0.8086	−0.3934	−0.2182	−0.2937	0.1587	−0.2009

Table 4. Bias values.

The number of neurons	b_i	b_j	b_k
1	−0.8567	1.8814	0.9318
2	0.157	−0.6724	
3	−0.6825	−0.3435	
4	0.8101	0.8886	
5	1.0588	1.8144	
6	0.2539		

chestnuts. For this purpose, a multilayer perceptron neural network was trained by the backpropagation algorithm to yield the main parameters (Bekesiene et al., 2021). Based on the ANN approach construction, training, and testing stages, the examinations are presented below. These assessments support the suggestion that using an ANN approach in biological research helps select the cultivars that are optimal in situations when numerous and diverse physical traits have occurred (Mancuso et al., 1999; Feng et al., 2018; Farhadi et al., 2020; Öztekin et al., 2020; Singh et al., 2022). Accordingly, the use of ANNs for biological species identification has been found more appropriate (Dubey et al., 2006; Pandolfi et al., 2009). These results indicate that the selected physical properties to process and design equipment are critical traits that significantly impact the hybrid cultivars. This is in line with the research results on biological species. Indeed, previous studies (Taner et al., 2018; Öztekin et al., 2020) have suggested that the parameters, such as the length, width, thickness, GMD, sphericity, SNV, BD, TD, porosity, and SA, can be used for agricultural crop processing and equipment design.

The acquired R^2 and the RMSE values for both the test and training in the present study indicate that the ANN approach was a success. Based on the measured values and the calculated values by the ANN approach for each physical trait, and the error values (Table 5), the ε value of 0.079% obtained by the approach was lower than the acceptable limit of 10%. Therefore, these results indicate that the test data obtained by the ANN approach were compatible with the measured data. Further, these results supported the ideas that suggested the chestnut genotypes (Mancuso et al., 1999) and commercially important hybrid chestnut cultivars (Öztekin et al., 2020) could be classified with high accuracy by the ANN approach.

Bekesiene et al. (2021) emphasized that the number of hidden layers in an ANN approach is crucial to achieving a specified approximation order. In the present study, the structure of the ANN approach shows that the hidden layers influence approximation instruction for the randomly sufficient smooth function and the varying network parameters, as reported previously (Taner et al., 2018; Öztekin et al., 2020; Bekesiene et al., 2021).

4. Conclusion

In this study, an ANN-based classification method was proposed for chestnut varieties. The model can be used successfully to classify chestnut cultivars. In this way, innovative applications will be evaluated and the identification of chestnut varieties will be possible without the need for experts and expensive methods. Easily practicable identification and classification methods were chosen for classifying chestnut varieties in the present study. The determined identification and classification traits ranging from a simple device to a complex classification system can be used for designing many systems. Increasing the amount of data and varieties will further increase the reliability of this model made according to physical characteristics. Thus, it will be able to find a place in the application. It is predicted that the proposed model can be widely used in the identification of chestnut varieties due to its high speed, reliability, and accuracy. This method is thought to solve the problem of chestnut classification. In future studies, more chestnut varieties and data should be used and the classification ability of ANN and discriminant analysis models for chestnut varieties grown in different regions and climatic conditions in Türkiye can be developed. Furthermore, smartphone applications using the models proposed in this study can be developed. Moreover, it will be of interest

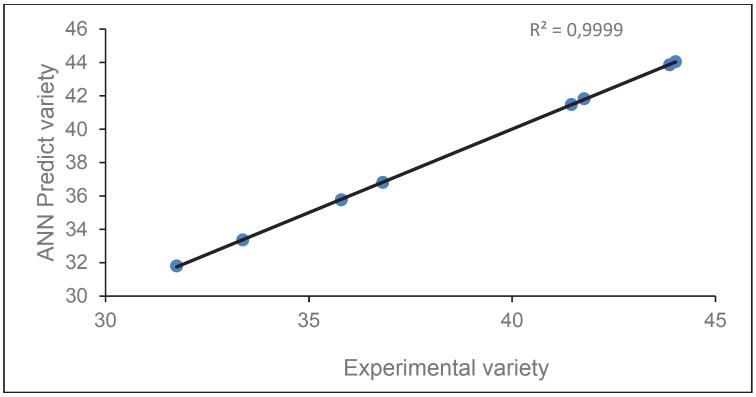

Figure 4. The relationship between the artificial neural network data and the measured data.

Table 5. The test and the relative error (ε) values for the artificial neural networks approach.

Cultivar	Measured data	Test data	ε (%)
Macit 55	31.75	31.80	0.160
Macit 55	33.38	33.36	0.060
Akyüz	43.88	43.85	0.051
Akyüz	41.46	41.48	0.044
Ali Nihat	36.82	36.81	0.042
Ali Nihat	35.80	35.76	0.120
Bouche de Betizac	44.02	44.04	0.039
Bouche de Betizac	41.77	41.82	0.117

to a wide audience, from meeting consumer needs to both producers and breeders.

Acknowledgments
I would like to thank A. Taner and E. Yeşiloğlu Cevher for their generous support of this study and N. Ocak for proofreading the manuscript.

Conflicts of interest
The author declares no conflict of interest.

References

Bechtler H, Browne MW, Bansal PK, Kecman V (2001). New approach to dynamic modelling of vapour-compression liquid chillers: ANN. Applied Thermal Engineering 2 (1): 941-953. https://doi.org/10.1016/S1359-4311(00)00093-4

Bekesiene S, Smaliukiene R, Vaicaitiene R (2021). Using artificial neural networks in predicting the level of stress among military conscripts. Mathematics 9 (6): 626. https://doi.org/10.3390/math9060626

Dubey BP, Bhagwa, SG, Shouche SP, Sainis JK (2006). Potential of artificial neural networks in varietal identification using morphometry of wheat grains. Biosystems Engineering 95 (1): 61-67. https://doi.org/10.1016/j.biosystemseng.2006.06.001

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. Applied Sciences 10 (9): 3315. https://doi.org/10.3390/app10093315

Feng L, Zhu S, Lin F, Su Z, Yuan K et al. (2018). Detection of oil chestnuts infected by blue mold using near-infrared hyperspectral imaging combined with artificial neural networks. Sensors 18 (6): 1944. https://doi.org/10.3390/s18061944

Jacobs RA (1988). Increased rate of convergence through learning rate adaptation. Neural Networks 1 (4): 295-307. https://doi.org/10.1016/0893-6080(88)90003-2

Kalogirou SA (2001). Artificial neural networks in the renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews 5 (4): 373-401. https://doi.org/10.1016/S1364-0321(01)00006-5

Keskenler MF, Keskenler EF (2017). From past to present artificial neural networks and history. Takvim-i Vekayi 5 (2): 8-18 (in Turkish with an abstract in English).

Mancuso S, Ferrini F, Nicese FP (1999). Chestnut (*Castanea sativa* Mill.) genotype identification: an artificial neural network approach. The Journal of Horticultural Science and Biotechnology 74 (6): 777-784. https://doi.org/10.1080/14620316.1999.11511188

Massantini R, Moscetti R, Frangipane MT (2021). Evaluating progress of chestnut quality: a review of recent developments. Trends in Food Science and Technology 113: 245-254. https://doi.org/10.1016/j.tifs.2021.04.036

Öztekin YB, Taner A, Duran H (2020). Chestnut (*Castanea sativa* Mill.) cultivar classification: an artificial neural network approach. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 48 (1): 366-377. http://dx.doi.org/10.15835/nbha48111752

Pandolfi C, Messina G, Mugnai S, Azzarello E, Masi E et al. (2009). Discrimination and identification of morphotypes of *Banksia integrifolia* (Proteaceae) by an Artificial Neural Network (ANN), based on morphological and fractal parameters of leaves and flowers. Taxon 58 (3): 925-933. https://doi.org/10.1002/tax.583020

Purushothaman S, Srinivasa YG (1994). A back-propagation algorithm applied to tool wear monitoring. International Journal of Machine Tools and Manufacture 34 (5): 625-631. https://doi.org/10.1016/0890-6955(94)90047-7

Serdar Ü, Akyüz B, Ceyhan V, Hazneci K, Mert C et al. Horticultural characteristics of chestnut growing in Turkey. Erwerbs-Obstbau 2018; 60: 239-245. https://doi.org/10.1007/s10341-017-0364-4

Singh D, Taspinar YS, Kursun R, Cinar I, Koklu M et al. Classification and analysis of pistachio species with pre-trained deep learning models. Electronics 2022; 11 (7): 981. https://doi.org/10.3390/electronics11070981

Taner A, Oztekin YB, Tekguler A, Sauk H, Duran H (2018). Classification of varieties of grain species by artificial neural networks. Agronomy 8 (7): 123. https://doi.org/10.3390/agronomy8070123

Visen NS, Paliwal J, Jayas DS, White ND (2002). Specialist neural networks for cereal grain classification. Biosystems Engineering 82 (2): 151-159. https://doi.org/10.1006/bioe.2002.0064

Yıldız T, Cevher EY (2022). Some mechanical properties of chestnut in relation to product processing and equipment design. Turkish Journal of Agriculture-Food Science and Technology 10 (8): 1565-1570. https://doi.org/10.24925/turjaf.v10i8.1565-1570.5332