https://doi.org/10.17221/21/2022-CJFS

Classification of hazelnuts according to their quality using deep learning algorithms

Nizamettin Erbaş[1], Gökalp Çınarer[2], Kazım Kılıç[1]*

[1]*Yozgat Vocational High School, Yozgat Bozok University, Yozgat, Turkey*
[2]*Department of Computer Engineering, Faculty of Engineering, Yozgat Bozok University, Yozgat, Turkey*

**Corresponding author: nizamettin.erbas@bozok.edu.tr*

Citation: Erbaş N., Çınarer G., Kılıç K. (2022): Classification of hazelnuts according to their quality using deep learning algorithms. Czech J. Food Sci., 40: 240–248.

Abstract: Hazelnut is a product with high nutritional and economic value. In maintaining the quality value of hazelnut, the classification process is of great importance. In the present day, the quality classification of hazelnuts and other crops is performed in general manually, and so it is difficult and costly. Performing this classification with modern agricultural techniques is much more important in terms of quality. This study was based on a model intended to detect hazelnut quality. The model is about the establishment of an artificial intelligence-based classification system that can detect the hidden defects of hazelnuts. In the developed model, the visuals used in the dataset are divided into training and test groups. In the model, hazelnuts are divided into 5 classes according to their quality using AlexNet architecture and modern deep learning (DL) techniques instead of traditional hazelnut classification methods. In this model developed based on artificial intelligence, a very good approach was presented with the accurate classification of 99%. Moreover, the values regarding precision and recall were also determined at 98.7% and 99.6%, respectively. This study is important in terms of becoming widespread information technology use and computer-assisted applications in the agricultural economics field such as product classification, quality, and control.

Keywords: artificial intelligence; image processing; food quality; computer science; marketing

Food and nutrition have always been an important issue since the history of humanity. Hazelnut, which is an important food and nutritional source, helps to carry natural vitamins to the human body thanks to the natural oils and fibres it contains. Hazelnut, which is also very rich in protein, supplies a large part of the daily protein requirement.

Hazelnut is also an advantageous product in terms of economic value. Although hazelnuts are produced in many countries in the world, the number of hazelnut-producing countries on an international trade scale is limited. The main producing and exporting countries are Turkey, Italy, the USA, and Spain (Kılıç and Alkan 2006; MAAF 2021). According to the data in 2019, a total of 1 125 178 t of hazelnuts were produced in the world. The average value of 1 t of hazelnuts is USD 7 021, and 1 kg is USD 7 (FAOSTAT 2019). Turkey earns a substantial income from the hazelnut trade.

Today, when the reason for the development of developed countries is examined, it is seen that agriculture has held great significance in the development process. Since quality and sustainable agriculture is the most important goal of the countries, it is necessary to apply modern agricultural techniques at every stage of agricultural production. In the last 10 years, food quality has become the focus of attention of societies to a larger extent (Gang 2018). Like in other sectors, in the agricultural sector also and in particular, the

https://doi.org/10.17221/21/2022-CJFS

technology is widely used for the quality classification of crops. With the development of computer technologies, innovative methods are also applied to the classification of products. Especially the increase in processor capacities and the use of video cards are important factors in the development of image processing techniques. With the widespread use of computer vision, the way has been cleared for the classification of products in different categories. While applying these classification processes, clustering methods, dimensional analysis of objects, determination of product type and classification in terms of quality are carried out.

The most important problem in the hazelnut industry, which has a very large market in the world, is that the quality and robustness of products cannot be determined correctly. In recent years, safe and high-quality production of foods has come into prominence (Öztürk and Kaşko Arıcı 2017) because image and robustness, namely quality, holds great significance in people's product preference (Kalkan et al. 2008). While a broken, crushed or rotten hazelnut is not preferred by the consumer, robust and distinct textured hazelnuts are preferred more. For this reason, the classification of quality levels of food products according to their images is an important stage during sales. However, today, farmers classify manually, but it takes a lot of time to distinguish products belonging to different classifications in manual classification. In addition, due to the loss of time and cost increase for the personnel in manual classification, such factors negatively affect the product quality classification process.

Thus, the classification of products by technological methods is of great importance. Thanks to the artificial intelligence technology from the technological methods, products are classified according to their quality (Kamilaris and Prenafeta 2018). The artificial intelligence, like in many fields, has been used in agricultural applications in recent years as well. Here, high-precision algorithms of artificial intelligence technologies, the use of which is rapidly increasing in every field, provide a chance for increasing quality and sustainability in the agricultural sector. As the digital world of the future rises with the artificial intelligence in order to be able to better assimilate this technology, using algorithms in agriculture will also be an important step (Di Caro et al. 2019). Thus, hazelnuts are classified according to their quality by using modern techniques instead of traditional hazelnut classification methods.

When these situations are considered, with an artificial intelligence-based hazelnut quality classification system:

– It will help to develop artificial intelligence-based automatic systems instead of manual classification applications in agriculture;
– Farmers will be facilitated;
– Products will be classified as economical;
– It will save time and space;
– Labour force costs and production costs will decrease and profitability will increase;
– Products will be marketed at a proper and correct price;
– Export will be easier;
– Better quality products will be offered to the market;
– It will be contributed to the regional and national economy;
– A positive contribution will be made to the promotion of products in the country.

In this study, a model that classifies hazelnuts into five categories in terms of cracks, holes, marks, cuts, and robustness was proposed by artificial intelligence algorithms. In the proposed model, hazelnuts were classified by artificial intelligence algorithms with respect to their quality properties.

In the study conducted by Solak and Altınışık (2018), hazelnuts were classified into three categories as small, medium and large by image processing techniques. The success rate in the classification was between 90% and 100%.

Cai and Liu (2019) proposed the betel nut classification method based on transfer learning. In the study performed with the AlexNet network, the betel nut classification accuracy was 89.4% in the dataset consisting of only 189 images.

Dheir et al. (2020) classified five different hazelnut species by a dataset containing 2 868 images. In the model in which artificial neural networks (ANN) and deep learning (DL) algorithms were used 98% accuracy was obtained.

In the study of Alnajjar (2021), hazelnut species were classified by ANN and image processing techniques. The dataset consisting of 1 595 images of four different hazelnut species and the model attained 100% accuracy in the test set.

Koç et al. (2020) classified 50 samples of pointed, black, and chubby hazelnut varieties by Gradient Boosting, Random Forest, and DL4J algorithms. In classification, maximum length, width, pressing durability, and weight of hazelnuts were used. Classification accuracy in Gradient Boosting, Random Forest and DL4J algorithms was determined to be 94, 100, and 94%, respectively.

In the next part of the research, hazelnut classification by the proposed model and methodology is exam-

https://doi.org/10.17221/21/2022-CJFS

ined, and the results of classification were evaluated and discussed in detail.

MATERIAL AND METHODS

Dataset. MVTec Anomaly Detection (MVTec 2021) is a dataset developed for industrial investigations and consisting of more than 5 000 images by categorising 15 different objects and textures. This covers a different number of test images for each category. There are 3 629 pictorial images, and 1 725 datasets reserved for testing (Bergmann et al. 2021). In this study, hazelnut images in the dataset were used to measure hazelnut quality. There are 514 hazelnut images with 1 024 × 1 024 resolution in 5 different categories in total in the data set (Bergmann et al. 2019). The original image numbers for each category are given in Table 1, and the dataset image examples are also given in Figure 1.

Data augmentation. The number of the data is important in the training performance of DL architectures. Images belonging to 4 categories in the dataset were made to quadruple by rotating 90 degrees to the right, 90 degrees to the left, and 180 degrees. Examples of original and augmented images are given in Figure 2. Images belonging to the good category were found adequately sufficient and no data augmentation was applied. In Table 1, the number of original and augmented images of each class is given.

Table 1. The number of original and augmented images of the categories in the dataset

Class	Original	Augmented
Cut	17	68
Crack	18	72
Good	391	391
Hole	18	72
Print	17	68

Convolutional Neural Network (CNN). Convolutional Neural Network (CNN) is the most popular DL model widely used for image analysis. CNN is a deep neural network architecture classifying the images and making automatically features related to filtering on the pixel matrices (Gu et al. 2018). CNN architectures consist of two parts: feature learning and classification. In the feature learning section, there are convolutional, activation, and pooling layers. Convolutional layer is the main layer used to determine the features of the images used in the study. It examines the features of the images used in my work and determines the impor-

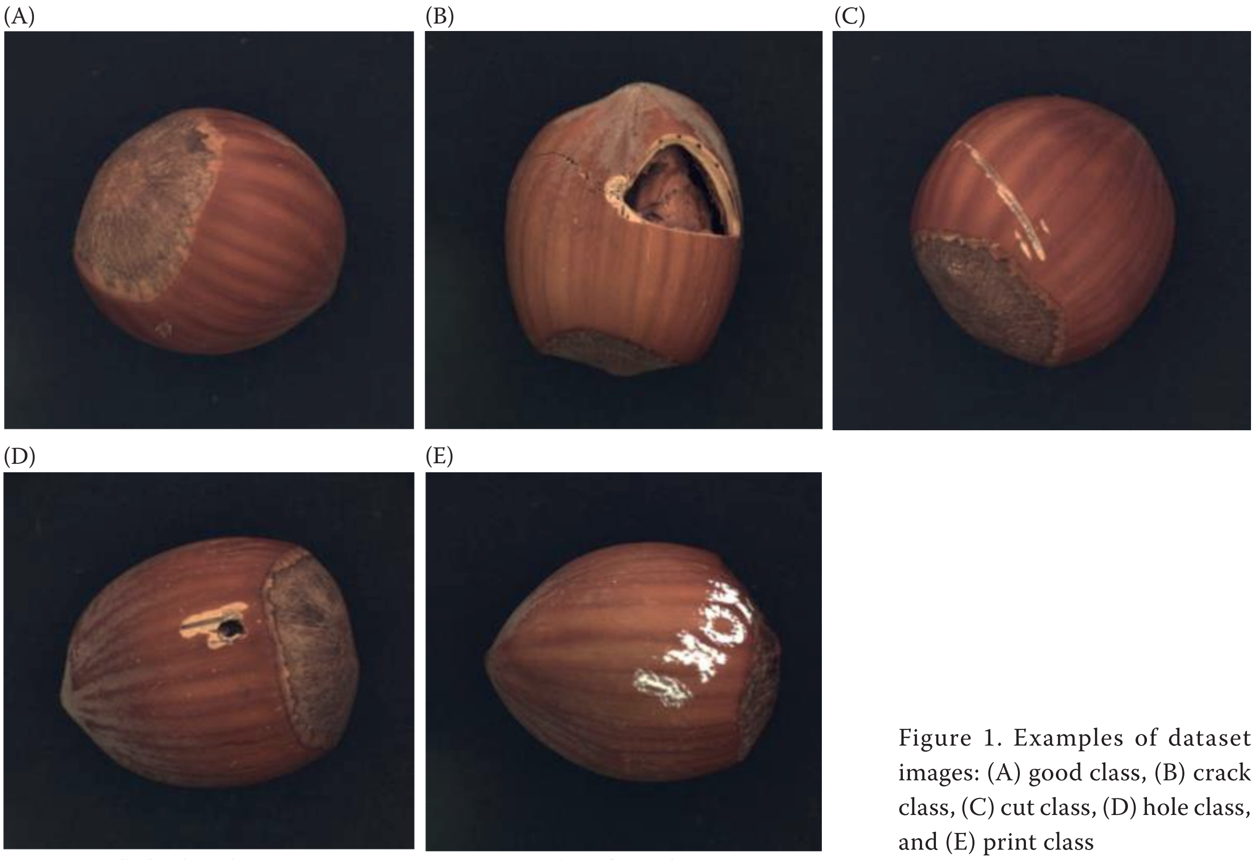

Figure 1. Examples of dataset images: (A) good class, (B) crack class, (C) cut class, (D) hole class, and (E) print class

https://doi.org/10.17221/21/2022-CJFS

(A) (B) (C) (D)

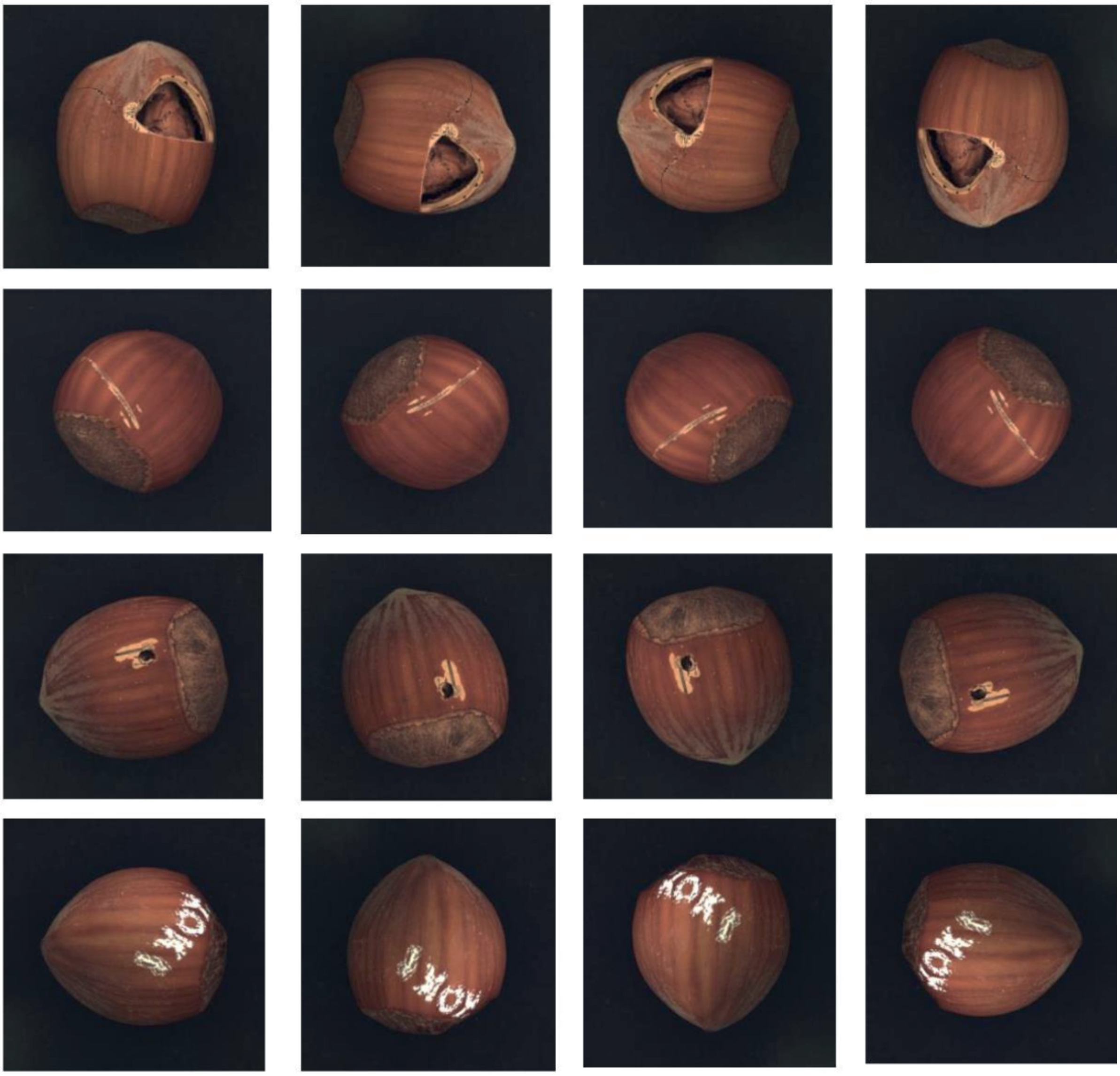

Figure 2. Enlarged versions of hazelnut images: (A) crack class, (B) cut class, (C) hole class, and (D) print class

tant features with the filters to be applied. In addition, the most appropriate feature is determined with the expanded matrix size used here. In the pooling layer, the weights of the architecture are controlled and the weights are reduced to make it suitable for the architecture. By applying the activation layer, the linear learning of the network is prevented and it is opened to evaluate with different parameters. Thus, better learning can be achieved with networks with nonlinear activation functions. The classification section also consists of flattened and fully tied layers. In the convolution layer, image pixels given as an entry are passed through convolution filters with certain weights. In the next stage, values obtained are applied to the activation function, and the linearity of the network is tried to be eliminated.

AlexNet. In the study, the AlexNet architecture, one of the popular CNN architectures, was used in the classification of hazelnut images (Krizhevsky et al. 2012). The input image size of this architecture is 227 × 227. The overall structure of the architecture consists of 5 convolutional layers and 3 fully tied layers. Filters of the size 11 × 11 are used in the first convolutional layer, 5 × 5 in the second convolutional layer, and 3 × 3 in all other convolutional layers. While maximum pooling is used in the first two convolutional layers, the next 3 convolutional layers are directly intertied. At the end of the fifth convolutional layer, maximum pooling is performed again. Rectified Linear Unit (ReLU) is used in the activation process of AlexNet architecture. In fully tied layers, dropout is used to prevent the network memorising. The AlexNet

https://doi.org/10.17221/21/2022-CJFS

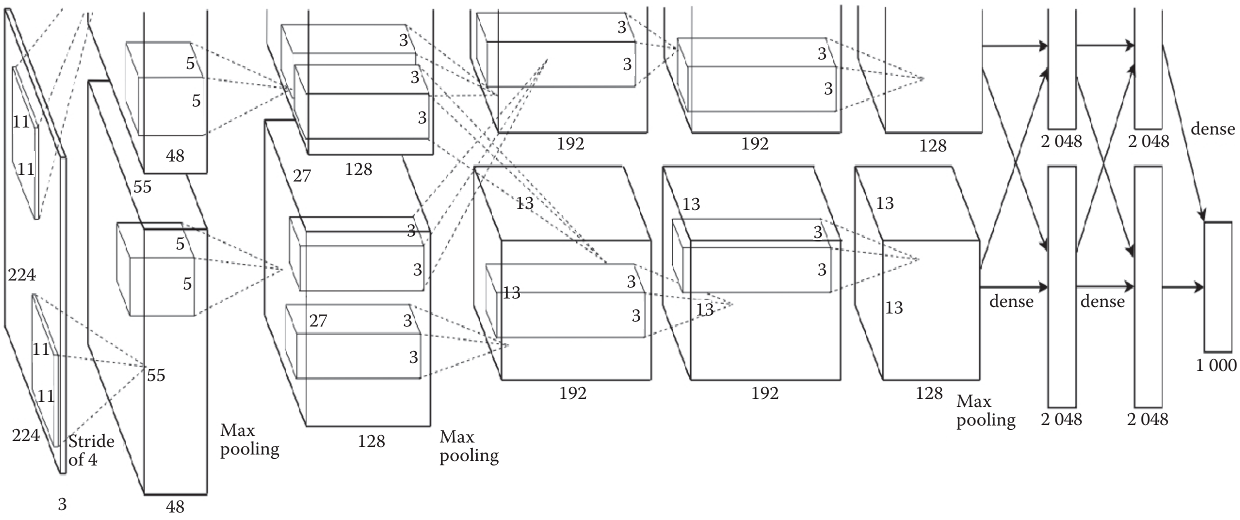

Figure 3. AlexNet architecture

architecture is shown in Figure 3 (Krizhevsky et al. 2012). Since AlexNet was the basic architecture that made DL popular first, this architecture was preferred in this model. ZFNet is used in classification in different architectures such as Google Net, but the AlexNet structure is the basis of this work, and it is focused on the results of the basic architecture instead of the architectures developed over this structure. It is a DL architecture that won the first place with its successful results in ImageNet competitions.

Transfer learning. The transfer learning approach is the use of the parameters of the network trained in a task in a similar new task. A lot of data is needed to train CNN architectures from scratch with images in the dataset. Using the transfer learning approach, those CNNs that cannot be trained from scratch with sufficient data are provided to achieve high accuracy performance in the new task (Fırıldık and Talu 2019). However, using the transfer learning approach in CNN architectures provides the opportunity for the network to progress faster according to training from scratch.

Experimental setup. In detecting hazelnut quality, images in the augmented dataset were divided into the training of 85% and the testing of 15%. ReLU was used for the training of the AlexNet architecture in the activation layers, and the Adam algorithm was also used for the optimisation of the network. The batch size was detected 64, and the number of training epochs 20.

Table 2. CNN architecture used in the experiment

Layer	Process	Feature map	Size	Kernel size	Stride	Activation	Train or not
Input	image	1	227 × 227 × 3	–	–	–	not
1	convolution	96	55 × 55 × 96	11 × 11	4	ReLU	not
–	Max pooling	96	27 × 27 × 96	3 × 3	2	ReLU	not
2	convolution	256	27 × 27 × 256	5 × 5	1	ReLU	not
–	Max pooling	256	13 × 13 × 256	3 × 3	2	ReLU	not
3	convolution	384	13 × 13 × 384	3 × 3	1	ReLU	not
4	convolution	384	13 × 13 × 384	3 × 3	1	ReLU	not
5	convolution	256	13 × 13 × 256	3 × 3	1	ReLU	not
–	Max pooling	256	6 × 6 × 256	3 × 3	2	ReLU	not
6	FC	–	9 216	–	–	ReLU	train
7	FC	–	4 096	–	–	ReLU	train
8	FC	–	4 096	–	–	ReLU	train
Output	FC	–	5	–	–	Softmax	train

CNN – Convolutional Neural Network; ReLU – Rectified Linear Unit; FC – fully-connected

https://doi.org/10.17221/21/2022-CJFS

The learning rate interval test (LRIT) method (Smith 2017) was used for the learning speed. The applications were implemented in the cloud environment using the Python programming language and the fast.ai (Howard and Gugger 2020) library.

In Table 2, pieces of information on the structure of the architecture used for feature extraction and classification from hazelnut images are given.

RESULTS AND DISCUSSION

As a result of the experiment, classification results obtained with the AlexNet architecture are given in Table 3. Hazelnut images were classified into 5 different ways. When Table 3 was examined, it was seen that 99% accuracy was achieved. In addition, precision and recall values were detected at 98.7% and 99.6%, respectively.

As a result of the experiment, precision, recall, *F*-score, and support values of each classification group are given in Table 4. When Table 4 is examined, it is seen that the AlexNet algorithm estimates the images of the crack group with a precision value of 93%. Apart from this, all groups reached a precision value of 100%. The AlexNet architecture correctly estimated all images of the good group and achieved the success of 100% in precision, recall, and *F*-score values.

Micro, macro, and weighted precision, recall, *F*-score results, and support values of classification results were given in Table 5. According to this, classification results of hazelnut quality by the artificial intelligence-based CNN model were determined quite high. When Table 5 is examined, it is seen that the suggested architecture was 100% in the macro recall value.

As a result of classification, the numbers of both correct and incorrect images obtained for each class are shown in Figure 4.

When the confusion matrix was examined, it was determined that it correctly estimated 53 images of the 54 images of the good category. The algorithm correctly estimated all the 13 images of the crack class, 14 images of the cut group, 9 images of the hole group, and 10 images of the print group. Only 1 image of a total of 100 images was misclassified.

The receiver operating characteristic (ROC) curve and area under curve (AUC) value showing the true-positive and false-positive ratio of the classification architecture are shown in Figure 5. ROC and AUC successfully represent image qualities. As can be seen from the ROC curve, the model proposed by the AlexNet architecture achieved high success in classifying hazelnut quality.

Classification of agricultural products by using high-level information technologies such as artificial intelligence and DL is the best example of a multidisciplinary study in the agricultural economics field. When literature studies are reviewed, studies on the classification of hazelnut quality by high-level information technology are rather limited. It is seen that previous studies were aimed at comparatively examining development levels of products rather than quality and economic contribution. Miraei Ashtiani et al. (2020) examined almond shells according to their thickness and width and performed a study to classify almond masses. They estimated almond quality with 96% accuracy by using different classification algorithms. In some studies, the classification process was applied according to product size and mass. In a study conducted by Vidyarthi et al. (2020), the mass and the original sizes of peanuts were estimated. In such studies, to determine the product size and calculate its mass is quite complex and difficult. In order to eliminate this situation, visual objects are used instead of mass calculations. By considering all kinds of conditions, in addition to the cracked and hole hazelnuts in the classification, overprinted hazelnuts were also included in the study. Dheir et al. (2020) classified hazelnuts into 5 different categories with 910 images and the CNN model belonging to the test group in a dataset consisting of 2 868 images. In the study, 98% accuracy was achieved. In this study, it is seen that the classification success of the AlexNet architecture model achieved a higher success of 99% with fewer images. Han et al. (2021) classified hazelnuts into 3 groups as good, medium, and poor. As a result of fivefold cross-validation with the CNN model, an average accuracy rate of 93.48% was reached. Narendra and Kini (2018) classified peanut quality with a similar machine learning model in agricultural classification. With the machine learning model they developed, the best prediction models in classifying peanut varieties were obtained 82.27% with Random Forest, 84.9% with Multilayer Perceptron, and 86.07% with

Table 3. Macro averaged performance summary of CNN architecture

Algorithm	Accuracy	AUC	Precision	Recall	*F*-score	Error rate
AlexNet	0.9900	1.0000	0.9857	0.9962	0.9909	0.2378

CNN – Convolutional Neural Network; AUC – area under curve

https://doi.org/10.17221/21/2022-CJFS

Table 4. Evaluation criteria obtained for each class

Classes	Precision	Recall	*F*-score	Support
Cut	1.00	0.98	0.99	54
Crack	0.93	1.00	0.96	13
Good	1.00	1.00	1.00	14
Hole	1.00	1.00	1.00	9
Print	1.00	1.00	1.00	10

Table 5. Micro, macro, and weighted average assessment criteria from all classes

Results	Precision	Recall	*F*-score	Support
Micro average	0.99	0.99	0.99	100
Macro average	0.99	1.00	0.99	100
Weighted average	0.99	0.99	0.99	100

libSVM algorithm. It was observed that machine learning algorithms achieve lower accuracy in classification. The DL model applied in this study eliminates this problem. In the study, the fact that the precision value was 98.5% and the recall value was 99.6%, an indicator of the success of the architectural model was applied. Moreover, the model estimated all products belonging to the 'good category' with 100% accuracy, precision, and recall results. It was seen that the AlexNet architecture was quite successful in classification. Schmidt-Hieber (2020) stated that the depth of neural network architectures, that is, the number of layers used, is important in the ReLU activation function. The high accuracy obtained in this research supports this. In a similar study that Han et al. (2021) conducted by applying different experiments, the good-, medium- and low-quality hazelnuts in the test set were classified. In the study, the accuracy values of 95.59, 90.00, 95.83, and 93.48% with different experimental groups were reached in the test data set. These results obtained in the CNN-model study conducted with different experimental clusters show that the quality classification has a great potential for an accurate, real-time, and non-destructive estimation.

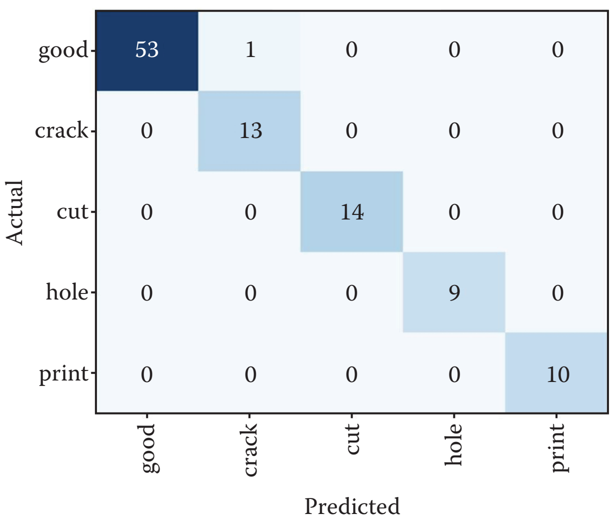

Figure 4. Confusion matrix

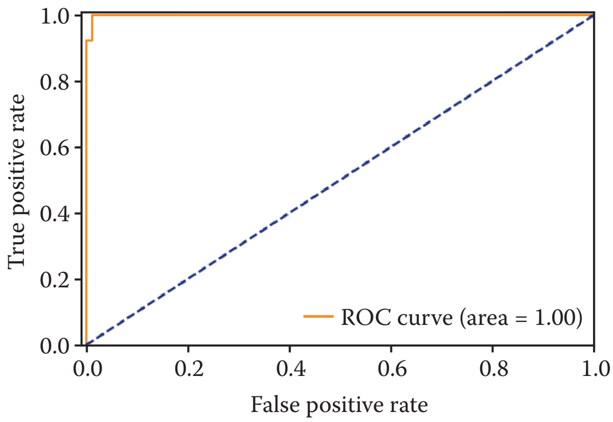

Figure 5. Receiver operating characteristic (ROC) curve

CONCLUSION

To predict quality levels from hazelnut images, DL classification model based on convolution was developed. In the model, hazelnut images were classified into five different categories and an average accuracy of 99% was achieved. In the study, three key factors were emphasised in detecting hazelnut quality: *i*) when confusion matrix results were analysed, the proposed AlexNet model performs better in classifying hazelnuts belonging to the good category; *ii*) the use of data augmentation methods can increase the performance of classification in this model; *iii*) when Max pooling and ReLU processes are applied, classification performs better. This study exhibits a great potential regarding combining imaging process techniques with DL to predict hazelnut quality with high accuracy. Agriculture has always been an important sector in the economic development of countries. For this reason, in order to ensure quality and sustainable development, modern techniques and technological applications should also be used in the marketing stage, like in other stages of agriculture.

In this study, the DL algorithms model which is an advanced information technology was used in hazelnut classification. With the proposed model using the CNN

https://doi.org/10.17221/21/2022-CJFS

application, hazelnuts were classified according to quality elements and great success was achieved. If the model is applied, classification will be able to be performed with high accuracy in the hazelnut industry. It is estimated that the model, which is expected to contribute to the use of information technology in the agricultural field, will also provide a great advantage with respect to the agricultural economics.

REFERENCES

Alnajjar M. (2021): Image-based detection using deep learning and Google Colab. International Journal of Academic Information Systems Research (IJAISR), 5: 30–35.

Bergmann P., Fauser M., Sattlegger D., Steger C. (2019): MVTec AD – A comprehensive real-world dataset for unsupervised anomaly detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, June 15–20, 2019: 9592–9600.

Bergmann P., Batzner K., Fauser M., Sattlegger D., Steger C. (2021): The MVTec anomaly detection dataset: A comprehensive real-world dataset for unsupervised anomaly detection. International Journal of Computer Vision, 129: 1038–1059.

Cai H., Liu S. (2019): Betel nut classification method based on transfer learning. In: 2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), Dali, China, May 24–27, 2019: 1039–1043.

Dheir I.M., Mettleq A.S.A., Elsharif A.A., Abu-Naser S.S. (2020): Classifying nuts types using Convolutional Neural Network. International Journal of Academic Information Systems Research (IJAISR), 3: 12–18.

Di Caro D., Liguori C., Pietrosanto A., Sommella P. (2019): Quality control of hazelnuts by means of NMR measurements. In: IOP Conference Series: Earth and Environmental Science, Vol. 275, 1st Workshop on Metrology for Agriculture and Forestry (METROAGRIFOR), Ancona, Italy, Oct 1–2, 2018: P1–P10.

FAOSTAT (2019): Faostat Statistical Database. [Dataset]. FAOSTAT. Available at https://www.fao.org/faostat/en/#data (accessed Jan, 2022).

Fırıldık K., Talu M.F. (2019): Investigation of transfer learning approaches used in Convolutional Neural Networks. (Evrişimsel sinir ağlarında kullanılan transfer öğrenme yaklaşımlarının incelenmesi). Computer Science, 4: 88–95. (in Turkish)

Gang L. (2018). The impact of supply chain relationship on food quality. Procedia Computer Science, 131: 860–865.

Gu J., Wang Z., Kuen J., Ma L., Shahroudy A., Shuai B., Chen T. (2018): Recent advances in convolutional neural networks. Pattern Recognition, 77: 354–377.

Han Y., Liu Z., Khoshelham K., Bai S.H. (2021): Quality estimation of nuts using deep learning classification of hyperspectral imagery. Computers and Electronics in Agriculture, 180: 105868.

Howard J., Gugger S. (2020): fastai: A layered API for deep learning. Information, 11: 108.

Kalkan H., Ince N.F., Tewfik A.H., Yardimci Y., Pearson T. (2008): Classification of hazelnut kernels by using impact acoustic time-frequency patterns. EURASIP Journal on Advances in Signal Processing, 2008: 247643.

Kamilaris A., Prenafeta S.X. (2018): A review of the use of convolutional neural networks in agriculture. The Journal of Agricultural Science, 156: 312–322.

Kılıç O., Alkan I. (2006): The developments in the world hazelnut production and export, the role of Turkey. Journal of Applied Sciences, 6: 1612–1616.

Koç C., Gerdan D., Eminoğlu M.B., Yegül U.I., Koç B., Vatan-daş M. (2020): Classification of hazelnut cultivars: Comparison of DL4J and ensemble learning algorithms. Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 48: 2316–2327.

Krizhevsky A., Sutskever I., Hinton G.E. (2012): Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25: 1097–1105.

MAAF (2021): Hazelnut in the World. Ankara, Turkey, Ministry of Agriculture and Forestry, General Directorate of Vegetation Report (MAAF): 1–10.

Miraei Ashtiani S.H., Rohani A., Aghkhani M.H. (2020): Soft computing-based method for estimation of almond kernel mass from its shell features. Scientia Horticulturae, 262: 109071.

MVTec (2021): MVTec AD – A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. [Dataset]. MVTec. Available at https://www.mvtec.com/company/research/datasets/mvtec-ad (accessed Jan, 2022).

Narendra V.G., Kini A.S. (2018): An intelligent classification model for peanut's varieties by color and texture features. International Journal of Engineering and Technology (UAE), 7: 250–254.

Öztürk D., Kaşko Arıcı Y. (2017): Analysis of production and marketing problems of hazelnut producers: A case of Samsun Province. Ordu University Journal of Social Science Research, 7: 21–34.

Schmidt-Hieber J. (2020): Nonparametric regression using deep neural networks with ReLU activation function. The Annals of Statistics, 48: 1875–1897.

Smith L.N. (2017): Cyclical learning rates for training neural networks. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, USA, March 24–31, 2017: 464–472.

https://doi.org/10.17221/21/2022-CJFS

Solak S., Altınışık U. (2018): Detection and classification of hazelnut fruit by using image processing techniques and clustering methods. Sakarya University Journal of Science, 22: 56–65.

Vidyarthi S.K., Singh S.K., Tiwari R., Xiao H.W., Rai R. (2020): Classification of first quality fancy cashew kernels using four deep convolutional neural network models. Journal of Food Process Engineering, 43: 1–13.

Received: January 31, 2022
Accepted: May 16, 2022
Published online: June 20, 2022