An improved pistachio detection approach using YOLO-v8 Deep Learning Models

Çınarer Gökalp[1*], *Çakır* Mübarek Mazhar[2]

[1]Yozgat Bozok University, Department of Computer Engineering, Turkey
[2]Yozgat Bozok University, Department of Mechatronics Engineering, Turkey

Abstract. Pistachios are an agricultural product widely used in the food industry. It is very important that pistachios are presented to the consumer in good quality on time. At the same time, whether the shells of pistachios are open or closed is an important criterion from a commercial industrial point of view. Pistachios with their shells open have a high unsaturated fat content, a high maturity level and an expensive market value. In this study, the open or closed status of pistachios was determined by using Artificial Intelligence-based deep learning models. For pistachio detection, 423 image data belonging to the Pesteh dataset were classified using models of the Yolov8 algorithm, which detects objects using convolutional neural networks. The data set is divided into 80% training, 10% validation and 10% testing. The performances of the models were evaluated with precision, recall, F1 and mAP score metrics. The highest test mAP value of the Yolov8 algorithm, which was run with image data consisting of pistachios, was obtained with the Yolov8-m model with 94.8%. The Yolov8-m model achieved a very successful result with 49.6 MB weight size, 11.0 ms inference time value and 0.33 hours training time value. In addition, the model's fast classification performance and small file size facilitate its applicability in the industrial field. The results show that the classification and detection of open and closed shell pistachios has been successfully carried out with Yolo models.
Keywords Yolov8, Pistachios Detection, Deep Learning, Artificial Intelligence

1 Introduction

The food and agriculture sectors are important business lines for the sustainability of human life. In addition, meeting the consumption needs of people around the world is becoming more difficult day by day. Meeting the increasing needs of such important sectors is only possible by applying technological developments in the sector.

Pistachio is a hard-shelled nut. When the structure of pistachios is examined, it consists of a lower sugar level of 10% compared to some other food products, a significant amount of protein of approximately 20%, 50% fat and 90% unsaturated fatty acids [1]. Contains potassium, vitamins C and E, as well as tocopherols, carotenoids and some phenolic compounds [2]. When its chemical components are examined, it can be seen that pistachios are a valuable, important and powerful food that can be consumed. Research shows that pistachios are at the top of the list of foods with antioxidant properties [3], that they have an effect on the control of the glycemic index, and that, when consumed in moderation, they are effective in weight control due to their satiety effect [4]. In addition to being a consumable food, it has an important place in the domestic and foreign markets as an important export product of the producing countries with various areas of use (hot confectionery, chocolate production, cake production, etc.) [5]. The effects of open or closed shells of pistachios commercially are as follows: The price of open-shelled pistachios is higher than closed-shelled pistachios, there is also the possibility of closed-shelled pistachios not ripening, and this is a factor affecting quality. In addition, customer (consumer) demand is fully met by open-shell pistachios [6]. A model developed using deep learning technology for the detection of closed-shell and open-shell pistachios is presented. The main objectives of the developed research include:

- Carrying out a basic study to automatically distinguish between the two types by using deep learning technology for the detection of closed-shell and open-shell pistachios in the food industry,
- Optimizing hyperparameters values to increase the accuracy of the Yolov8 algorithm,
- Finding the most appropriate algorithm based on deep convolutional neural networks with high accuracy to classify and detect pistachios and comparing the results with other algorithms.

2 Related Works

When the studies in this field in the literature are examined, the number of studies on pistachio classification is quite limited. In addition, there are studies classifying different types of nuts. Chen et al [7] addressed the issue of almond recognition and classification in their study. They used Yolov5 and Faster R-CNN algorithms to detect almonds. Both algorithms have successfully fulfilled their purposes.

* Corresponding author: gokalp.cinarer@bozok.edu.tr

They reached 0.79, 0.77 and 0.73 F1 values in the detection of green, brown and black almonds, respectively. Ng et al [8] detected nuts using the Mummy Nuts dataset. They also used Yolov3, Yolov5, CenterNet, Faster R-CNN, MobileNet SSD algorithms for detection and shared their comparative results. In the first stage of the study, which they carried out in two stages, they trained the algorithms with the original version of the dataset, and in the second stage, they trained the algorithms with more data by applying augmentation to the dataset data. To compare the two situations, the increase rates of the algorithms before and after the data augmentation process are shared. Sanlin et al [9] worked on walnut detection with deep learning. In their study, they created a dataset by collecting images from three different walnut species. They trained Yolov3, Yolov4, Yolov5 and Faster R-CNN algorithms with the dataset data they created. As a result of the study, the Yolov5 algorithm was found to be more successful and the Yolov5 algorithm was recommended for walnut detection. Divyanth et al [10] obtained images under different light and at different angles for coconut detection. By performing augmentation on the obtained data, 900 image data were obtained. They trained the Faster R-CNN algorithm with a dataset consisting of 900 image data. As a result of the training, the mAP value of the algorithm was stated as 0.886. When the literature was examined, it was seen that the applied models were limited to the latest Yolov5 algorithm.

3 Material and Methods

3.1 Dataset

Image data from the open source Pesteh dataset was used in the study [11]. The dataset, consisting of 423 images, is reserved for 80% training, 10% validation, and 10% testing. There are a total of 3096 label information in the image data in the dataset. Open shell pistachios are labeled '1' and closed shell pistachios are labeled '0'. Figure 1 shows the labeled images of the dataset.

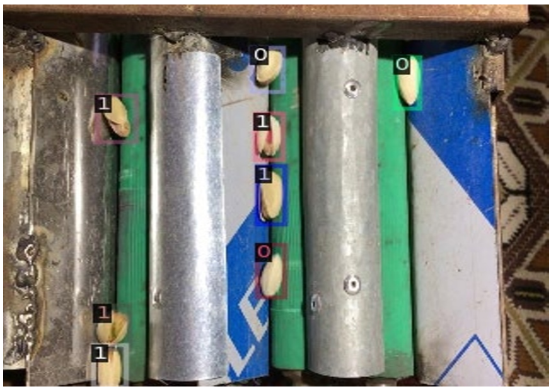

Fig. 1. Example of Labeled Image Data

3.2 Selection of Deep Learning Method and Preparation of the System

A computer with high system requirements is required to run deep learning algorithms. Another solution is Colab, which offers a virtual computer by Google. Colab offers a virtual environment where you can run artificial intelligence applications with powerful GPU and RAM. Deep learning algorithms were run via Python software from Colab. While writing the codes, data processing operations were carried out using Python-based libraries such as Pytorch, Tenserflow and Keras.

The most important features of deep learning algorithms are that they work quickly and have high accuracy rates. There are multiple algorithms developed for object detection. In their study, Viswanatha et al [12] stated that CNN models, SSD, Retinanet and other models were slower than the Yolo algorithm. Based on this, the models of Yolo's most recently developed Yolov8 [13] algorithm were used to detect pistachios. Yolov8 is a model with higher accuracy than other models on the coco dataset [14]. One of the innovations in the Yolov8 model is the anchor-free detection system. This innovation enables the acceleration of Non-Maximum Suppression operations. Additionally, the anchor free feature allows you to directly estimate the center of an object [14]. Instead of C3, which uses only the last output, the C2f structure, which combines all the outputs from Bottleneck, was used. The 6x6 kernel size in the bottleneck section has been changed to 3x3. The structure of the Yolov8 algorithm is shown in Figure 2.

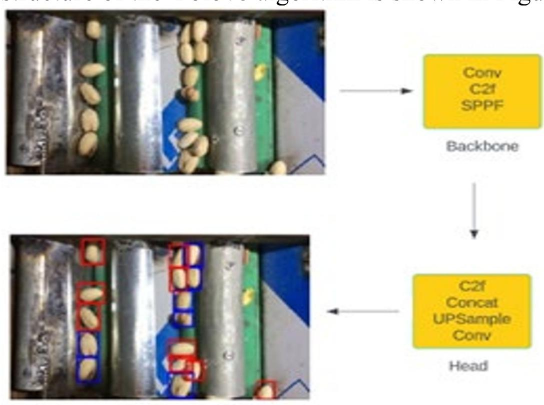

Fig. 2. Structure of the Yolov8 Model

In order to examine the results of the Yolov8 models, which will detect pistachios, under the right conditions, 5 models (n, s, m, l, x) must be run under equal conditions. For this reason, hyperparameter values of the models are the same in every model. Learning rate is a coefficient required to minimize error [15]. In our study, this value was determined as 0.01. SGD (Stochastic Gradient Descent) [16] one of the best methods in optimizing neural networks, was used as the optimization algorithm. Using momentum with SGD reduces oscillation by increasing convergence speed [17]. The momentum value is set to 0.937. Activation function is the hyperparameter value that controls how the outputs should change in order to bring them closer to the line in non-linear operations [18]. Activation function is set as SiLU. All models are set to run over 100 epochs. Yolov8 algorithm processes image data in groups while working on every 1 epoch. The batch size value, which determines the number of data in each image data group, is determined as 16. The dimensions of the images were reduced to 640x640, ensuring that all images were of equal size. Then, all models were run and the results were recorded.

3.3 Methods of Evaluation of Results

There are some values that allow us to get an idea about whether the algorithms are successful or not and to interpret the results. These values are based on whether the algorithm's prediction on an object is correct or incorrect. true positive (TP); if a pistachio is present in the image and the algorithm correctly guesses this object as a pistachio; false negative (FN); if there are pistachios on the image and the algorithm predicts that there are no pistachios; If there are no pistachios in the image, the algorithm predicts that there are pistachios, which is called false postive (FP) [19]. Depending on these values, precision is the ratio of TP value to the sum of TP and FP values; recall is calculated as the ratio of TP value to the sum of TP and FN values [19]. The mAP value is used to measure the accuracy of the detection algorithm on the entire object classes in a dataset [20].

4 Experimental Results

Yolov8 models were run with 423 image data and 3096 label information from the Pesteh dataset. The train results of the models run were recorded (see Table 1). When the precision values are examined, the highest precision value is the Yolov8-l model with 94.1%. The Yolov8-l model is followed by the Yolov8-m and Yolov8-x models with values of 93.4% and 92.7%, respectively. The highest recall value is 95.5%. This value belongs to the Yolov8-m model. Yolov8-x and Yolov8-s models with 94.4% and 93% recall values follow the Yolov8-m model. F1 score allows precision and recall values to be represented and evaluated through a single value. Our three highest models with F1 values of 94.4%, 93.5%, 93.4% are Yolov8-m, Yolov8-x, Yolov8-l, respectively. The Yolov8-l model has the highest mAP value, which brings all the values under a single expression. The mAP value of the Yolov8-l model is 98.1%. Yolov8-l model is followed by Yolov8-m model with 97.8% mAP value.

Table 1. Yolov8 Models Train Results.

Models	Precision	Recall	F1	mAP
Yolov8-n	0.864	0.920	0.891	0.952
Yolov8-s	0.926	0.930	0.927	0.970
Yolov8-m	0.934	**0.955**	**0.944**	0.978
Yolov8-l	**0.941**	0.929	0.934	**0.981**
Yolov8-x	0.927	0.944	0.935	0.968

Dataset data was divided into three groups: train, test and validation. Of the three parts, the train and validation parts were used during the learning of the models. The image data in the test section has not been seen by the models before. In other words, the test phase gives results close to the results under real conditions. For this reason, test results are very important to us. The results of the testing phase are seen in Table 2. When the precision values are examined, the highest precision value is the Yolov8-m model with 92.9%. The Yolov8-m model is followed by the Yolov8-l and Yolov8-s models with values of 89.8% and 89.4%, respectively. The highest recall value is 91.9%. This value belongs to the Yolov8-s model. Yolov8-m and Yolov8-n models with 91.3% and 89.6% recall values follow the Yolov8-s model. Our three highest models with F1 values of 92%, 90.6%, 89% are Yolov8-m, Yolov8-s, Yolov8-l, respectively. The Yolov8-m model has the highest mAP value, which brings all the values under a single expression. The mAP value of the Yolov8-m model is 94.8%. The Yolov8-m model is followed by Yolov8-s and Yolov8-n models with mAP values of 93.7% and 92.8%.

Table 2. Yolov8 Models Test Results.

Models	Precision	Recall	F1	mAP
Yolov8-n	0.853	0.896	0.873	0.928
Yolov8-s	0.894	**0.919**	0.906	0.937
Yolov8-m	**0.929**	0.913	**0.920**	**0.948**
Yolov8-l	0.898	0.884	0.890	0.916
Yolov8-x	0.867	0.841	0.853	0.875

Training time, inference time and weight size values of the models are given in Table 3. Considering the train times of the models as n, s, m, l and x, respectively, values of 0.1 hours, 0.17 hours, 0.33 hours, 0.52 hours, 0.81 hours were reached. In this context, the fastest model is the Yolov8-n model, while the slowest model is the Yolov8-x model. When the inference time values, defined as the inference speed per image, were examined, the Yolov8-n model gave the best result with 2.0 ms. This model is followed by s, m, l and x models with values of 4.2 ms, 11.0 ms, 17.1 ms and 26.5 ms, respectively. As with training time values, the model with the best inference time value is Yolov8-n and the model with the worst result compared to others is Yolov8-x. The evaluation order of the models in terms of weight size values is similar to the training time and inference time values. The best result is obtained in the Yolov8-n model with 5.96 MB, while the worst result belongs to the Yolov8-x model with 130 MB. Models with weight size values of 21.4 MB, 49.6 MB and 83.5 MB are Yolov8-s, Yolov8-m and Yolov8-l models, respectively. On the other hand, the values showing speed, weight size and accuracy rate are observed in a balanced manner in the Yolov8-m model.

Table 3. Duration Metrics of Yolov8 Models.

Models	Training Time (h)	Inference Time (ms)	Weight Size (MB)
Yolov8-n	0.105	2.0	5.96
Yolov8-s	0.172	4.2	21.4
Yolov8-m	0.337	11.0	49.6
Yolov8-l	0.520	17.1	83.5
Yolov8-x	0.818	26.5	130

5 Discussion

The aim of determining the shells of pistachios according to their open and closed status was achieved. The results showing the accuracy rate of the algorithm are of the desired quality. Loss functions decrease as the process progresses, as it should. In Yolov8-m, the train process is a very short time of 19.8 minutes, and in this short time, a high train mAP value of 97.8% and a test mAP value of 94.8% were reached. In addition, the small file size makes it easier to integrate into a computer system or an embedded system, and the disadvantage of a model with a large file size slowing down the system is avoided.

When the train mAP values of the Yolov8-m model were examined on a class basis, mAP values of 97.7% in open-shelled pistachios and 97.8% in closed pistachios were obtained. When the test mAP values were examined, mAP values of 94.7% in open-shelled pistachios and 95% in closed pistachios were obtained. This shows the stability of the model. The inference time value of the model per image is 11.0 ms, that is, approximately 1/100th of a second.

As a result of their study using the Rahimzadeh et al [11] Pesteh dataset, they reached the highest test accuracy value with 94.75% mAP value among the algorithms they used. The model that gives 91.87% accuracy is the ResNet50 model. Rahimzadeh et al. used images that were not included in the dataset during the testing phase and reached these results by averaging the accuracy values of the objects detected on the image. In our pistachio detection study, the Yolov8 algorithm, the most up-to-date Yolo model, was used. During the testing phase, the algorithm tested the data using the test image data in the dataset, and the test results show that the Yolov8-m model gives better results than the ResNet50 model. Additionally, Rahimzadeh et al in their study, stated the inference time per image and training time per epoch as 17.44 ms and 318 s, respectively. In our study, this value was recorded as 11.0 ms. In the Yolov8-m model, the training time per epoch is 11.88 s, which is approximately 1/27 of the training time per epoch of ResNet50.

Karadağ et al [21] determined the open and closed shell status of pistachios with the dataset they created. They used dual cameras in their setup and gave the results of the accuracy values separately for the dual cameras. They used the Faster R-CNN algorithm as a feature extractor in the detection of pistachios and ResNet152 for the classification process. They performed the testing independently of the dataset data. In this case, the accuracy values in the detection and classification processes of open-shelled pistachios and closed-shelled pistachios are 98% and 85%, respectively. The accuracy values they give for the right camera and left camera are 92.7% and 92.3%, respectively. In this study, 94.8% mAP value was reached in the test process of identifying and classifying pistachios according to their shells. This shows that the accuracy value of the two cameras used in their study by Karadağ et al. is higher than the accuracy value. When looked at on a class basis, values of 94.7% and 95% were obtained, respectively, in the detection of open-shelled and closed-shelled pistachios during the test phase. In their study by Karadağ et al., when the average of accuracy values for two classes is taken, it is 91.5%. As a result, it was seen that the Yolov8-m model was more successful in detection than Faster R-CNN and in classification than ResNet152.

6 Conclusion

This study deals with the detection and classification of pistachios according to whether their shells are open or closed. The data and methods used for the purpose are explained. Five different models of the Yolov8 model were used to detect pistachio shells. It has been observed that the results obtained are sufficient and according to the results, the most stable model is Yolov8-m.

Subsequent work will include optimization studies on the structure of the model to increase the accuracy of the algorithm. In addition, it is planned to increase data and add new image data on the dataset, which is directly related to the accuracy value of the algorithm. It is estimated that optimization efforts in the model structure and improvement of the dataset will increase the accuracy rate in object detection. Finally, as an example of an automation system, real-time data analysis integrated into an embedded system can be applied.

References

1. Barghchi, M. & Alderson, P. G. Pistachio (Pistacia vera L.). 68–98 (1989) doi:10.1007/978-3-642-61535-1_6.
2. Mateos, R., Salvador, M. D., Fregapane, G. & Goya, L. Why Should Pistachio Be a Regular Food in Our Diet? *Nutrients* **14**, 3207 (2022).

3. Bulló, M., Juanola-Falgarona, M., Hernández-Alonso, P. & Salas-Salvadó, J. Nutrition attributes and health effects of pistachio nuts. *British Journal of Nutrition* **113**, S79–S93 (2015).
4. Dreher, M. L. Pistachio nuts: composition and potential health benefits. *Nutr Rev* **70**, 234–240 (2012).
5. GEZGİNÇ, Y. & DUMAN, A. D. Antep Fıstığı İşleme Tekniği ve Muhafazasının Kalite Üzerine Etkisi. *Gıda* **29**, (2004).
6. N. F. Ince *et al.* Discrimination Between Closed and Open Shell (Turkish) Pistachio Nuts Using Undecimated Wavelet Packet Transform. *Biological Engineering* **1**, 159–172 (2008).
7. Chen, A., Nagar, Y., Shoshani, G. & Sharon, R. *Deep learning for almond fruit detection at different growth stages for orchards pest management using unmanned aerial vehicles*. (Wageningen Academic, 2023).
8. Ng, D. *et al.* Benchmarking Object Detection Models with Mummy Nuts Datasets. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* vol. 13852 LNCS 102–119 (Springer Science and Business Media Deutschland GmbH, 2023).
9. Sanlin, Z., Liping, Z., Weiqiang, Z., Zhuang, G. & Ziqiang., F. Identification and localization of walnut varieties based on YOLOv5. *Journal of Chinese Agricultural Mechanization* **43**, 167 (2022).
10. Divyanth, L. G. *et al.* Detection of Coconut Clusters Based on Occlusion Condition Using Attention-Guided Faster R-CNN for Robotic Harvesting. *Foods 2022, Vol. 11, Page 3903* **11**, 3903 (2022).
11. Rahimzadeh, M. & Attar, A. Detecting and counting pistachios based on deep learning. *Iran Journal of Computer Science* **5**, 69–81 (2022).
12. V, V., K, C. R. & C., R. A. Real Time Object Detection System with YOLO and CNN Models: A Review. (2022).
13. YOLOv8 . https://github.com/ultralytics/ultralytics.
14. Chaturvedi, P., Lavingia, K. & Raval, G. Detection of traffic rule violation in University campus using deep learning model. *International Journal of System Assurance Engineering and Management* **1**, 1–19 (2023).
15. Senior, A., Heigold, G., Ranzato, M. & Yang, K. An empirical study of learning rates in deep neural networks for speech recognition. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 6724–6728 (2013) doi:10.1109/ICASSP.2013.6638963.
16. Ruder, S. An overview of gradient descent optimization algorithms. (2016).
17. Qian, N. On the momentum term in gradient descent learning algorithms. *Neural Networks* **12**, 145–151 (1999).
18. Dubey, S. R., Singh, S. K. & Chaudhuri, B. B. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing* **503**, 92–108 (2022).
19. Li, M., Zhang, Z., Lei, L., Wang, X. & Guo, X. Agricultural Greenhouses Detection in High-Resolution Satellite Images Based on Convolutional Neural Networks: Comparison of Faster R-CNN, YOLO v3 and SSD. *Sensors 2020, Vol. 20, Page 4938* **20**, 4938 (2020).
20. Padilla, R., Netto, S. L. & Da Silva, E. A. B. A Survey on Performance Metrics for Object-Detection Algorithms. *International Conference on Systems, Signals, and Image Processing* **2020-July**, 237–242 (2020).
21. Karadağ, A. E. & Kılıç, A. Non-destructive robotic sorting of cracked pistachio using deep learning. *Postharvest Biol Technol* **198**, 112229 (2023).