Turkish Journal of Engineering
https://dergipark.org.tr/en/pub/tuje
e-ISSN 2587-1366

Application of Convolutional Neural Networks for Watermelon Detection in UAV Aerial Images: A Case Study

Iclal Cetin Tas*[1], Ali Musa Bozdogan[2], Sami Arica [3]

[1] *Ostim Technical University, Electrical-Electronics Engineering Department, Türkiye, iclal.cetintas@ostimteknik.edu.tr*
[2] *Çukurova University, Agriculture Engineering Department, , Türkiye, amb@cu.edu.tr*
[3] *Çukurova University, Electrical-Electronics Engineering Department, Türkiye, sarica@cu.edu.tr*

Cite this study: Cetin Tas, I., Bozdogan, A.M. & Arica, S.. (2025). Application of Convolutional Neural Networks for Watermelon Detection in UAV Aerial Images: A Case Study. Turkish Journal of Engineering, 9 (1), 1-11.

https://doi.org/10.31127/tuje.1481696

Keywords
Watermelon
Alexnet
VGG19
Classification
Agriculture

Research Article

Received:10.05.2024
Revised:11.09.2024
Accepted:23.09.2024
Published: 20.01.2025

Abstract

Although human dependence on agriculture decreases with developing technology, it continues. As many resources are increasingly restricted due to various climatic reasons, the importance of studies in this field increases. Applications using deep learning models are frequently encountered in the agricultural field. In particular, there are applications where deep learning models are used as a tool for optimum planting, land use, yield improvement, production/disease/pest control, and other activities.In this study, watermelons in an aerial view of a watermelon field were detected by utilizing the Alexnet deep learning architecture. To obtain yield, watermelons in watermelon fields should be specified and then counted. Aerial images are used for this application. The field image was divided into 50% overlapping sub-images, and each was classified as watermelon, leaf, and soil. Consequently, watermelon regions on the field image were specified. After training the Alexnet and Vgg19 network structure with the dataset, watermelons were to be identified by segmenting the images. It was observed that the Vgg19 network achieved 97.78% accuracy. The results of the experimental applications show that the Vgg19 can be applied for watermelon fruit and yield detection applications.

1. Introduction

In recent years, global warming has made its impact felt very intensely worldwide. Ensuring the continuity of the ecological cycle and the continuity of the food supply are issues on the agenda of the whole world. Many developed countries continue their necessary studies in this field through very different platforms. Countries attach more importance to agricultural policies than ever before. It is essential to increase the amount of product obtained from the unit area, which is among the basic elements of agricultural policies, and to use existing resources effectively, as in all areas. Production policies for many fruits and vegetables have been established according to previous years' criteria, such as cultivation area and yield. Watermelon is one of these fruits, and it is produced in many countries. According to 2021 data, China, Türkiye, and India are in the top three, respectively [1]. While designing sustainable agricultural practices, many technological approaches have helped. Precision agriculture aims to improve yield, lessen the human workforce, use resources economically, and protect the environment. It requires multi-disciplinary studies. Previous studies have shown that single or hybrid applications based on image processing, artificial intelligence, and deep learning approaches can achieve accepted success. It has been seen that image processing techniques in the agricultural field contribute significantly to precision agriculture by determining the product and its quality, allowing an increase in the efficiency of agricultural applications in different subjects. These application inputs include irrigation, leaf disease, fruit classifications, fertilization, spraying, etc. Digitalization and signal analysis play a crucial role in providing this contribution by enabling instant vegetation detection, soil moisture content measurement, identification of disease and weed locations, classification of agricultural materials, and sometimes mitigating individual-based negative effects through automation. Thus, it is estimated that the contribution of techniques based on image processing and artificial intelligence to sustainable agriculture will increase gradually [1]. The processes in which deep learning methods are most commonly used are classification and segmentation [2-4]. Hetal et al. proposed a method for detecting fruit by multi-feature extraction and succeeded in detecting it with 90% accuracy for different types of fruits. Dorj et al. developed an effective and suitable computer vision algorithm to detect and count citrus fruits on a tree and conducted a study to estimate yield, achieving a correlation coefficient R^2 of 0.93 using image processing techniques. Koç et al. proposed a method to classify fruits according to size and color characteristics. The proposed method was applied to orange, apple, and quince fruits. In the training process of the study, in which artificial classifiers were used, a success rate of 93.6% for nearest neighbor, 90.3% for decision trees, 88.3% for Naive Bayes, 92.6% for multilayer mesh structures, and 94.3% for random forest was achieved[5]. Automatic fruit recognition and object detection in large crop fields were performed by processing the image obtained by using an unmanned aerial vehicle [6-9]. Machine learning methods were used to identify and count pineapple and non-fruit crown images. A feature-defusing one-way analysis of variance (ANOVA) was used to optimize the performance of the machine learning classifier. The results showed that the pineapple and the crown part could be distinguished, and an accuracy performance of 94.4% was obtained with the counting process [6]. There are studies in many areas, such as seed [10], crop [11], leaf and fruit health [12–14], sweetness level [15], and yield of watermelon plant [16]. The image processing applications in the agricultural field are highly concentrated. Jiang et al. used image processing techniques to determine the inline line in maize [17]. Maharlooei et al. used image processing to detect and count different-sized soybean aphids on a soybean leaf. Experiments were carried out with soybean plants grown in a greenhouse. By collecting three datasets on different dates using replica plants from 4 soybean cultivars infested with a range of aphid densities, they captured images of infested soybean trifoliate leaves with different cameras in 2 different lighting conditions with different cameras used in different datasets. Images are taken with an inexpensive regular digital camera. The images showed satisfactory results in high-lighting conditions [18]. Sabanci et al. Another study aimed to classify potatoes in terms of size with the help of image processing techniques and artificial neural networks. Before the classification process, potatoes with an outer surface and deformity were identified by using the Herbaceous method and morphological processes, and they excluded them from classification and classified the potatoes without problems in size at the next stage. After the training period, they made multilayer artificial neural networks and examined their classification success [19]. In another study, Faster R-CNN was used, which allows better use of time for fruit detection using deep convolutional neural networks. First of all, sweet pepper was detected with this model, and then it was applied to six different fruits. F1 score values of rockmelon, 0.848 strawberry, 0.948 apple, 0.938 avocado, 0.932 mango, 0.942 orange, 0.915, sweetpepper 0.828 were obtained [20].

The watermelons in the images obtained by an unmanned aerial vehicle (UAV) from the watermelon field in Adana province in Turkey, one of the world's watermelon producers, were segmented using the gray level co-occurrence matrix (GLCM) and K-means clustering [21]. The dataset used in Ekiz's study is used in our study. Classification was made using the Haralick texture features, and it was seen that the classification could be made with an accuracy rate of 86.46%. Detection of related objects using different methods provides benefits in many areas. In this study, marking the watermelons on the image will allow us to determine the product amount and make a yield estimation. Detection of fruit defects before and after harvest is an important criterion in many respects. In this study, Radial Based on Probabilistic Neural Networks, fruit surface defects are classified as color and texture anomalies. A matrix was created using the texture and color characteristics of the defective area, and then the relevant areas were classified by the RBPNN method [22].

Although acceptable levels of success were achieved in studies conducted with traditional image processing techniques, it was observed that higher success rates were obtained in studies conducted with deep learning methods. In the study in which the diseases of peach fruit were detected by deep learning, a success rate of 99.4% was determined [23]. 99.7% success was achieved in the study in which plant leaves were made with deep learning methods [24]. Layered structures of convolutional neural networks and feature extraction in layers play an important role in increasing success.

This study detected the watermelons in the images obtained by an unmanned aerial vehicle with pre-trained convolution neural network architecture (CNN). In the following section, the information about the dataset used, the mathematical models of the applied methods, and their application areas are explained. In section 3, the experimental process and application results were evaluated. The outcome has been evaluated and discussed from different perspectives, and predictions have been made about future studies in section 4 and section 5, respectively.

The study's primary objective was to develop a method for accurately detecting and counting watermelons in aerial images, contributing to yield estimation and supporting better resource management. The dataset used in the study was composed of aerial images divided into overlapping sub-images, classified into three categories: watermelon, leaf, and soil. This classification allowed for precise segmentation of the watermelon regions in the field.

The study compared the performance of two CNN architectures, Alexnet and Vgg19. After training both networks on the dataset, the Vgg19 network demonstrated superior performance in identifying watermelon regions. This high level of accuracy suggests that the Vgg19 model is well-suited for watermelon fruit detection and yield estimation applications.

In addition to detailing the dataset and methods in section 2, the study also evaluated the experimental processes and results in section 3. The discussion in sections 4 and 5 provides insights into the broader implications of the findings and outlines potential future research directions, emphasizing the role of deep learning in advancing agricultural practices. Overall, the study's results highlight the effectiveness of deep learning models in agricultural applications and suggest promising.

2. Materials and Methods

2.1 Dataset

Images in the watermelon field were taken three meters above the ground using the DJI Mavic Pro Fly 4K's UAV brand. The image size was [1080 1920 3], and its bit depth was 24 bits (8 bits per color channel). All the algorithms were implemented by using MATLAB R2020b software. The field images were divided into 201x201 sub-blocks, and 150 were employed to train the network. These sub-blocks consisted of 50 watermelons, 50 leaves, and 50 soil. The Y component of YCbCr color space transformation was utilized to obtain a monochrome image from an RGB image (in Eq.(1) and see Figure 1).

$$Y = 0.299\,R + 0.587\,G + 0.114\,B \quad (1)$$

Figure 1. Sample Images

It displays sample images from the dataset (see Figure 1) and shows sub-blocks obtained from the aerial images and then converted to grayscale for each class (see Figure 2). Each block's image size is 201x201 pixels. 1500 images were obtained by applying the data augmentation process differently for each class.

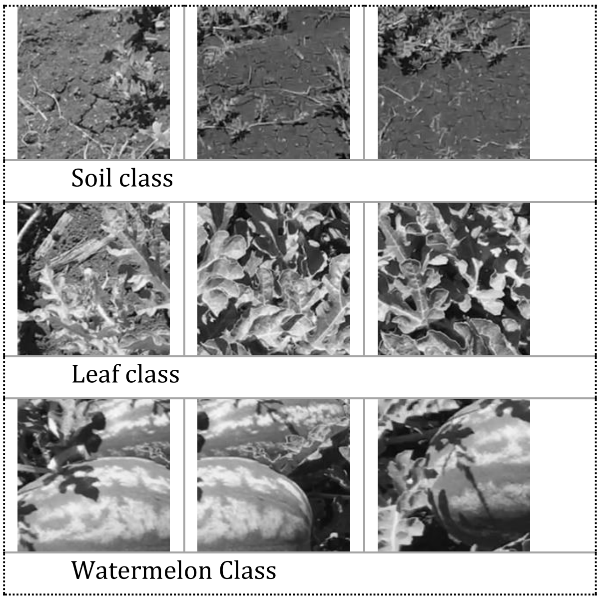

Figure 2. Training image samples for each class

The performance of deep learning models can be changed with slight differences. Performance is increased by using more and different layers. However, it is always possible to collect new data, so improvements are made by increasing the existing data. Since the size of the data set in our study was not very large, data augmentation was necessitated to increase the number of training samples. Using some data augmentation techniques, a relatively small dataset is transformed into a large database, and deep learning algorithms are trained with these datasets [25]. The basic benefit of data augmentation is to increase the accuracy of the test data after a healthier training process by creating additional training data by applying some deformation to the existing data. In the study, the data is augmented by using the following transformations: random rotation within [-20, 20] and [-5 5], degrees, translation along the x-axis or y-axis in the range of [-5 5] and [-50 50], scaling in x and y direction with varying scale in [1 1] and reflection along x or y axes.

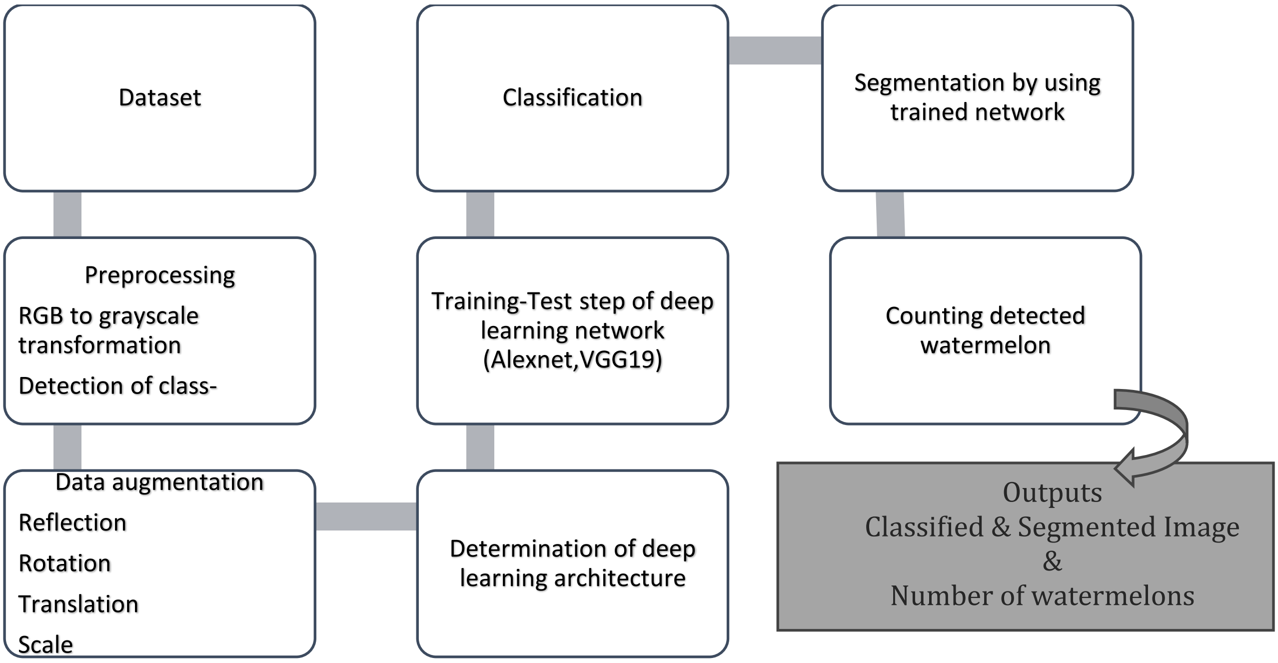

Figure 3. This study's flowchart

2.1. AlexNet

Deep learning is one of the most utilized methods in the agriculture area [26]. Convolution neural networks (CNNs) are widely employed due to their success in image classification. The main advantage of a classical machine learning classifier is that features are extracted automatically by adapting the filters used in each layer. Filters in the convolution layer in the network help extract spatial information from the image. A CNN model consists of several sections with three main building blocks. A section builds the convolutional layer to learn the features, the pooling layer to resample the image to reduce the size and, hence, the computational cost, and the non-linear (Relu) layer [27]. After the convolutional layers, the features extracted are flattened, and a feed-forward neural network follows. The final layer is the softmax layer, which determines the categories.

In this study, training in Alex.net architecture was done with the dataset. The trained network structure was used in the segmentation process. The Alexnet CNN model, designed with a deep learning architecture, won the ImageNet competition held in 2012. Alexnet was applied to identify leaf, fruit, and soil classes over watermelon images. Alexnet has a CNN-based architecture. It consists of 25 layers: an input layer, five convolution layers, three max pool layers, two dropout layers, fully connected layers, seven relu layers, two normalization, layers, and a softmax layer (Figure 4). The image at the input layer is size of 227x227x3. In the last layer, classification is made, and the value of the classification number in the input image is given. The filters are 11x11 in size, and the stride is 4 [28–31]. The Relu function in the Alexnet is in Eq.(2)

$$f(x) = \max(0, x) \tag{2}$$

The dropout layer is used to eliminate excessive.

Learning is memorization. Some connections that cause excessive learning are removed to prevent memorization of the network. In this way, the success of the network is increased. In the literature for automatic image recognition, labeling, and classification of fruits and vegetables, an automatic classification system was designed for Colombian fruits by training a convolutional neural network, and it was declared that 99.95% accuracy was achieved in the application results. The system consisting of 4980 images labeled in 22 classes, each corresponding to the images of the same type of fruit, was used for training and testing purposes. In addition, it has been tried to increase the reliability of the performance results of the model with the data duplication process. With the proposed method, transfer learning was applied by taking the parameters of a pre-trained model used for fruit classification as new initial parameters of the convolutional network, and it was observed that an increase in classification accuracy was obtained when trained with random initial weights compared to the same model [32]. Zhu et al. 's study [31] is one of the applications in which vegetable images are classified. ReLU is used instead of the output function of the Alexnet network, the traditional sigmoid function, which can accelerate the training of the deep learning network, and the tanh function. In the study, in which the direct effects of the size of the dataset on the accuracy rate were emphasized, it was seen that the accuracy rate of the deep learning method reached up to 92.1%, and this was the BP neural network (78%) and the SVM classifier (80.5%). Another paper has made an intelligent Mango glade classification based on Alexnet with the GrabCut algorithm. Preprocessing, object masking, and segmentation using GrabCut were performed on the images taken via the webcam. An Alexnet is adopted here to provide the mango grade output for the glade classification. As a result of the experiments, it was seen that a higher accuracy was obtained than the half-works in this field [30].

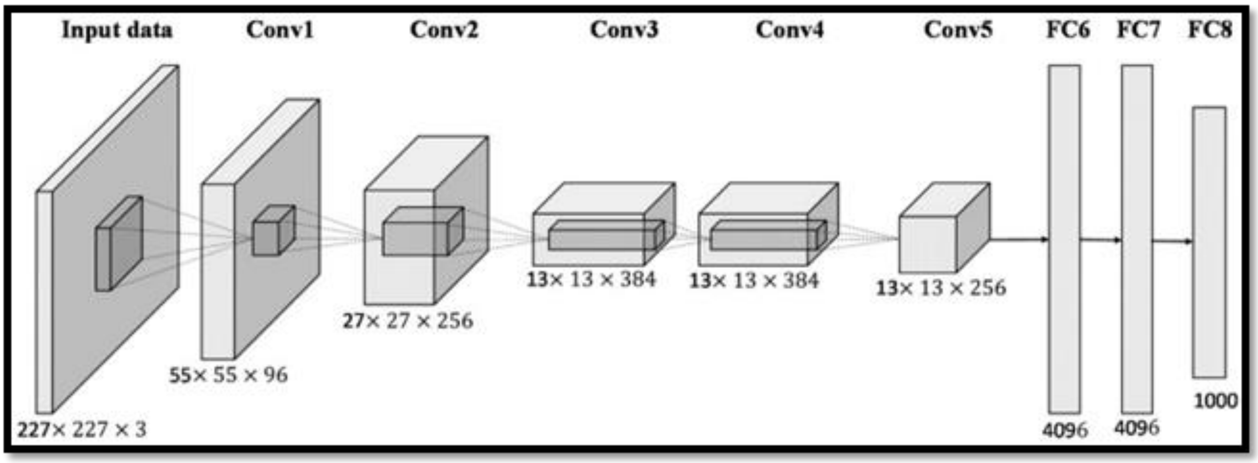

Figure 4. Alexnet architecture [31–34]

Tropical fruits obtained from the Supermarket Produce dataset with Alexnet were classified. This database includes shadows, exposure, occlusion, and reflection (fruits in a bag). It contains 2633 images of fruits divided into 15 categories with high variability and complexity. As a result, a 99.56% accuracy rate was obtained for all classes [35].

In this study, AlexNet deep learning architecture was used to detect watermelons in aerial images. AlexNet has a 25-layer structure and is a widely used model in deep neural network architectures. During the training process, the model was optimized for 200 epochs and 20 iterations per epoch. Stochastic gradient descent with momentum (SGDM) was used as the optimization algorithm. SGDM takes momentum into account to speed up the learning process and avoid local minima. During training, the mini-batch size was set to 10, and the model was shuffled after each epoch to provide better generalization performance. The maximum number of epochs was determined as 10 and the initial learning rate was selected as 1e-4. This small learning rate enabled the model to learn more precisely and helped to adjust the weights more accurately.

2.2. VGG19 Architecture

Vgg19 is a CNN developed that achieved high accuracy on the ImageNet dataset. The model consists of 19 layers, including convolutional and fully connected layers. Convolutional layers are used to extract features from the image, while fully connected layers are used to transform the features into the final classification decision. The Vgg19 model is loaded with pre-trained weights. These weights were trained on the ImageNet dataset, and feature extraction capabilities were learned for general object recognition tasks. During the training process, the output layer of the Vgg19 model is customized by aligning it with the classes in the dataset, and retraining the network is abused. The high kernel sizes used in the Alexnet architecture were reduced. Kernel sizes in Alexnet were not fixed, starting with 11 and continuing with 5 and 3. VGG16 fixed the kernel sizes. The idea was that 11x11 and 5x5 kernels could be replicated with more than one 3x3 kernel. The image to be included in the input layer is 224x224x3 in size. Vgg19 consists of 3 fully connected layers, 16 convolution layers, 1 SoftMax layer, and 5 MaxPoollayers (see Figure 5). The convolution layers' filters include 64, 128, and 256. [35–38]. The last layer is the classification layer. As in other deep architectures, in the VGG architecture, the height and width dimensions of the matrices decrease from the input layer to the output while the depth value increases. It has approximately 143 million parameters [39].

The VGG19 architecture and AlexNet differ significantly in terms of depth, complexity, and performance. VGG19 is a deeper network with 19 weight layers, while AlexNet has only 8 layers. This increased depth in VGG19 allows it to capture more complex and hierarchical features from the input data, which often leads to better performance in image classification tasks, especially for high-resolution images. This study used the VGG19 architecture, known for its deep neural network structure with 47 layers, to detect watermelons in aerial images. The training involved running the model for 200 epochs, with 20 iterations per epoch, ensuring sufficient learning time. The SGDM optimization algorithm was employed to enhance convergence speed by considering the momentum of past gradients, which helps the model avoid local minima. A mini-batch size 10 was used during training, and the dataset was shuffled after each epoch to improve generalization. The model was trained for 10 epochs, with an initial learning rate 1e-4. This lower learning rate enabled the model to fine-tune its weights with precision, allowing for more gradual and accurate adjustments during the training process.

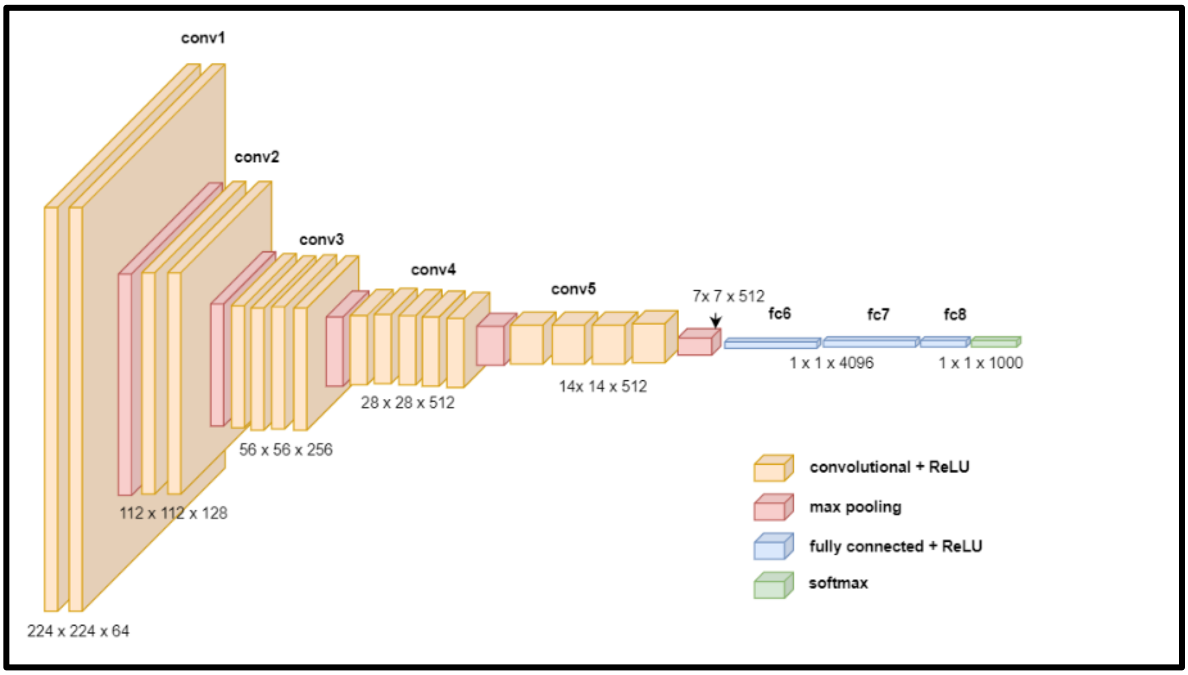

Figure 5. Vgg19 architecture [40]

3. Results

This study was performed on a computer with an Intel Core i7 processor, 16 GB RAM and 500 GB SSD storage capacity. It was performed in MATLAB 2020b version. The experimental procedures are explained in this section. The application of the proposed method aims to detect the location and number of watermelons in an image. Firstly, the data were grouped into three classes: watermelon, leaf, and soil. Data are balanced as all groups have an equal number of samples. The images were grouped manually for training. The class distinctions are clear in the sample images of the data set. The dominant class is included for some pictures containing elements belonging to two classes. Since each image is 2-dimensional grayscale, R, G, and B channels were taken to be the grayscale image to make it 3-dimensional. Data augmentation is applied. Some of the augmented images are shown in Figure 6.

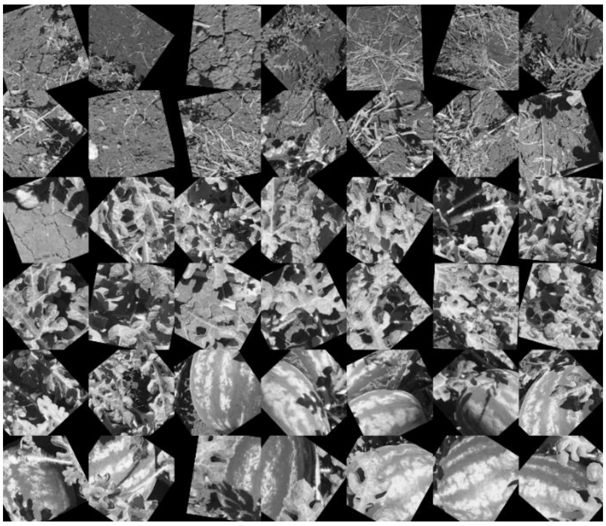

Figure 6. Augmented images

The training process was completed with augmented data for each class. The dataset is divided into %70 training and %30 for testing. It is aimed to increase the object detection capabilities of Alexnet and Vgg19 networks on the data set used in the study. The number of max epochs was 20 in the training phase. The batch size was fixed as 10. The maximum iteration number is equal to 200. It was used as 25 layers in the Alexnet architecture and 47 layers in the Vgg19 architecture. Softmax and classification layers are common in both network structures. A softmax layer is used, a softmax layer that applies a softmax function to the input. A classification layer is usually used after this layer. The task of the classification layer is to calculate the cross-entropy loss for classification and weighted classification tasks with mutually exclusive classes [41]. The softmax function can also be defined as a multiclass generalization of the logistic sigmoid function rather than the normalized exponentiation, as in Eq.(3).

$$y_r = \frac{\exp(a_r(x))}{\sum_{j=1}^{k}\exp\left(a_j(x)\right)} \tag{3}$$

$\boldsymbol{y_r}$ is the output unit function which is range0-1 (can be equal to or 1). $\sum_{j=1}^{k} \boldsymbol{y_j}$ is equal to 1. For multi-class classification problems can be shown as below (in Eq.(4)) :

$$P(c_r) = \frac{(\exp(a_r(x,0)))}{\sum_{j=1}^{k}\exp\left(a_j(x,0)\right)} \tag{4}$$

a_r represents conditional probability of sample class r, $P(c_r)$ is named class prior probability value. In a problem involving K classes, the new element is included in one of these classes by using the values taken from the previous softmax layer in the classification layer and the cross-etropy [39]. The loss function is desired in Eq.(5) :

$$Loss = \frac{-1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K} w_i\, t_{ni} Iny_{ni} \tag{5}$$

w_i represents the weight for class i, t_{ni} is the indicator that the nth sample belongs to the ith class and y_{ni} is the probability that the network associates the nth input with class i [42].

20 epochs applied per cycle. The iteration number is 10 per epoch. The learning rate is equal to 0.0001. The average processing time was 17 min 51 sec for Alexnet and 16 min 30 sec for Vgg19. Pre-trained networks trained to be customized on the dataset were applied to sample images. Consider the following aerial images are shown in Figure 1.

The aim is to determine the watermelon from an aerial image of the watermelon field. That process requires dividing the image into sub-images of 201x201 size by using a 50% sliding window. These sub-images are then applied to the trained network to obtain its category. All pixels in the pale of the sub-image are labeled as watermelon if it is as watermelon. Consequently, watermelons in the field image are segmented. That operation was demonstrated as in the following. The determined watermelons in these images are displayed in Figure 7 and Figure 8. It was observed that watermelons were detected in the images. The occluded areas or hidden parts by leaves of watermelons were not specified as watermelons because these areas were determined as leaf by networks. The segmentation may be improved by increasing the overlapping percentage and size of the sliding window used to extract sub-images. The results obtained from Alexnet can be seen in Figure 7.

Figure 7. Alexnet result images for the watermelon class

The same steps were applied to the VGG19 architecture; the results are shown in Fig. 8.

Figure 8. Vgg19 result images for watermelon class

The results were computed after 50 repetitions of the algorithm and averaging the outcomes. Since there was no significant improvement in accuracy when more application trials were performed, the analyses were performed on 50. The accuracy was computed between % 95.7. and % 100 for both networks.

A confusion matrix is often used in the literature to describe the performance of a classification model. The sensitivity, positive predictive value, negative predictive value, precision, sensitivity, and specificity of each class are given in Eq. (6-9):

$$Pr\,e\,cision\,(\%) = \frac{TP}{TP+FP}x100 \quad (6)$$

$$Recall(\%) = \frac{TN}{TN+FN}x100 \quad (7)$$

$$Accuracy(\%) = \frac{(TP+TN)}{TP+TN+FP+FN} \quad (8)$$

$$Fscore(\%) = \frac{2TP}{2TP+FP+FN} \quad (9)$$

Tp and fp represent true positives and true negatives and tn and fn are true negatives and false negatives. These are the values for which the correct result is obtained in 3 classes: watermelon, leaf, and soil. fp represents the error that occurs when the test result incorrectly confirms that data in a particular three classes does not actually exist. The average testing results are shown in Table 1.

Table 1. Test Results

Performance Criteria	Alexnet	Vgg19
Precision	0.9556	0.9778
Recall	0.9608	0.9792
Accuracy	%95.56	%97.78
F1-Score	0.9582	0.9785

At the last stage, a final processing was carried out to determine the number of watermelons whose locations were determined. In this process, the connected regions in the black-white images were not found and product counting was performed with the help of related functions to count them. The counting process is achieved by counting dependent regions first on a pixel and window basis and then on a group basis. As a result of counting in samples (see Figure 7 and Figure 8) results were obtained 3,2,3 respectively. These results have been determined to coincide with the watermelon numbers in the related images. The success calculated in the applications for the watermelon counting process is 75% for Alexnet and %80 for Vgg19. The main reason affecting the success rate here is the fruits hidden under the watermelon leaves.

Consequently, when the results obtained are examined, Vgg19 architecture has been observed to be more successful than Alexnet in classifying and segmenting watermelons for the dataset used in this study.

4. Discussion

In this study, we segmented watermelon in an aerial watermelon field image. The pre-trained Alexnet and Vgg19 networks were trained to categorize watermelon, leaf, and soil images of 201x201 size. The watermelons in a field image were detected by dividing the image into sub-images via 50% overlapping window and then estimating the category by using the trained network. These methods allowed for positioning the watermelons and then counting them. The counting provides the yield of the watermelon field. It may also be possible to estimate the size and hence the maturity of the watermelons.

Classification processes in previous studies were made only on plant components such as fruit or leaves. Additionally, the datasets in the studies include only the objects to be detected. In addition, it was observed that the objects detected were generally fruits[43] and vegetables such as oranges [44], tomatoes[45-46], carrots [47] etc., which were clearly distinguished from the leaf color. There are not many studies on watermelon fruit in the literature.

In this study, the color tones of the leaves and fruits are very close and in some cases cannot be seen with the naked eye when viewed from a certain distance. Due to the nature of the watermelon fruit, different color tones are not seen depending on its ripeness level. The fact that fruits are usually found more than once from the same region and in contact with each other can be considered as the limitations of this study. Because these situations directly affect the success rate. This situation makes applying different deep learning-based methods difficult and affects its performance. Semantic segmentation may be one of them. The direct counting of the detected watermelons may be incorrect since the nearby watermelons are segmented as a single connected object and the adjacent and touching watermelons should be divided. Although this, the outcomes of the proposed approach are promising and can be applied to watermelon segmentation.

The quality of the camera's optical sensor plays a key role in this process. Higher-quality sensors produce images with higher resolution, better contrast, and more accurate color representation. If image quality improves, the model can better distinguish fine details like watermelon edges and leaves, leading to better segmentation and classification results. The number of watermelons in an image can affect the model's performance. Along with the number, the location of the watermelons is also important. Classification of several watermelons in adjacent positions can be done with higher accuracy. However, if the watermelons are in independent positions and mixed with leaves, the detection and classification model may behave differently, even if there are many watermelons. More objects in the same image can lead to overlapping regions, increased complexity in feature extraction, and possibly confusion between classes such as watermelon and leaf.

As a result, in image classification-based studies where Alexnet and Vgg19 networks were used together, it was seen that Vgg19 was more competent and gave better results than Alexnet, similar to this study.

5. Conclusion

In this study, we examined the effects of CNN structure in the classification of watermelon plants. It was found that pre-trained Alexnet and Vgg19 networks were used to classify watermelon fruit. The results obtained showed that the Vgg19 network was more successful. In previous studies, deep learning networks were generally used only for classification. In this study, networks trained on the dataset and whose capabilities were increased. It was segmented by testing its classification capabilities on different images. It has been observed that deep architectures can also provide good results for this classification problem. According to the results obtained, it is thought that the method used can be used for automated agricultural machines. In future studies, if the training is implemented using devices with special hardware, such as a high-performance server with GPU support, the inference (test) can be faster and it is envisaged that it can be implemented in real-time. It can be used to facilitate processes such as farmers monitoring the growth of fruit in the field and yield estimation. For future research, existing and different data sets can be used after object detection is implemented and developed with internet of things structures.

Funding

This research received no external funding

Author contributions

Iclal Cetin Tas: Theoretical formalism, Performed The Analytic, Software, Validation, Writing. **Ali Musa Bozdogan:** Conceptualization, Methodology ;**Sami Arica**: Conceptualization, Methodology

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Ağin, O., Malasli, M. Z., & Tarihi, G. (2016). Görüntü İşleme Tekniklerinin Sürdürülebilir Tarımdaki Yeri ve Önemi: Literatür Çalışması. Journal of Agricultural Machinery Science, 12(3), 199–206.
2. Aydın, V. A. (2024). Comparison of CNN-based methods for yoga pose classification. Turkish Journal of Engineering, 8(1), 65–75. https://doi.org/10.31127/TUJE.1275826
3. Pajaziti, A., Basholli, F., & Zhaveli, Y. (2023). Identification and classification of fruits through robotic systems by using artificial intelligence. Engineering Applications, 2(2), 154–163.
4. Hyder, U., & Talpur, M. R. H. (2024). Detection of cotton leaf disease with machine learning model. Turkish Journal of Engineering, 8(2), 380–393.
5. Koç, D. G., & Vatandaş, M. (2021). Classification of Some Fruits using Image Processing and Machine Learning. Turkish Journal of Agriculture - Food Science and Technology, 9(12), 2189–2196. https://doi.org/10.24925/TURJAF.V9I12.2189-2196.4445
6. Wan Nurazwin Syazwani, R., Muhammad Asraf, H., Megat Syahirul Amin, M. A., & Nur Dalila, K. A. (2022). Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning. Alexandria Engineering Journal, 61(2), 1265–1276. https://doi.org/10.1016/J.AEJ.2021.06.053
7. Akar, Ö, Saralioğlu, E., Güngör, O.,& Bayata, H. F.(2024). Semantic segmentation of very-high spatial resolution satellite images: A comparative analysis of 3D-CNN and traditional machine learning algorithms for automatic vineyard detection. International Journal of Engineering and Geosciences, 9(1), 12-24 https://doi.org/10.26833/ijeg.1252298
8. Gholami, A. (2024). Exploring drone classifications and applications: a review. International Journal of Engineering and Geosciences, 9 (3), 418-442 https://doi.org/10.26833/ijeg.1428724
9. Villi, O., & Yakar, M. (2024). Sensortechnologies in unmanned aerial vehicles: types and applications.Advanced UAV, 4(1), 1-18
10. Mukasa, P., et al. (2022). Nondestructive discrimination of seedless from seeded watermelon seeds by using multivariate and deep learning image analysis. Computers and Electronics in Agriculture, 194, 106799. https://doi.org/10.1016/J.COMPAG.2022.106799
11. Zheng, Y. Y., Kong, J. L., Jin, X. B., Wang, X. Y., Su, T. L., & Zuo, M. (2019). CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in Precision Agriculture. Sensors, 19(5), 1058. https://doi.org/10.3390/S19051058
12. Jayakumar, D., Elakkiya, A., Rajmohan, R., & Ramkumar, M. O. (2020). Automatic Prediction and Classification of Diseases in Melons using Stacked RNN based Deep Learning Model. 2020 International Conference on System, Computation, Automation and Networking, ICSCAN 2020. https://doi.org/10.1109/ICSCAN49426.2020.9262414
13. Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanehkaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. Computers and Electronics in Agriculture, 173, 105393. https://doi.org/10.1016/J.COMPAG.2020.105393
14. Abu-Nasser, S. S. A. N., & Bassem, S. (2018). Rule-Based System for Watermelon Diseases and Treatment. International Journal of Academic Information Systems Research (IJAISR), 2(7), 1-7. https://www.researchgate.net/publication/326927417_Rule-Based_System_for_Watermelon_Diseases_and_Treatment
15. Nazulan, W. N. S. W., Asnawi, A. L., Ramli, H. A. M., Jusoh, A. Z., Ibrahim, S. N., & Azmin, N. F. M. (2020). Detection of Sweetness Level for Fruits (Watermelon) with Machine Learning. 2020 IEEE Conference on Big Data and Analytics, ICBDA 2020, 79–83. https://doi.org/10.1109/ICBDA50157.2020.9289712
16. Koc, A. B. (2007). Determination of watermelon volume using ellipsoid approximation and image processing. Postharvest Biology and Technology, 45(3), 366–371. https://doi.org/10.1016/J.POSTHARVBIO.2007.03.010
17. Jiang, G., Wang, Z., & Liu, H. (2015). Automatic detection of crop rows based on multi-ROIs. Expert Systems with Applications, 42(5), 2429–2441. https://doi.org/10.1016/J.ESWA.2014.10.033
18. Maharlooei, M., Sivarajan, S., Bajwa, S. G., Harmon, J. P., & Nowatzki, J. (2017). Detection of soybean

aphids in a greenhouse using an image processing technique. Computers and Electronics in Agriculture, 132, 63–70. https://doi.org/10.1016/J.COMPAG.2016.11.019
19. Fen, İ. Ü., Enst, B., & Der, K. (2022). Kadir SABANCI ve ark. Retrieved June 30, 2022, from http://www.tuik.gov.tr/bitkiselapp/
20. Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., & McCool, C. (2016). DeepFruits: A Fruit Detection System Using Deep Neural Networks. Sensors, 16(8), 1222. https://doi.org/10.3390/S16081222
21. Ekiz, A., Arica, S., & Bozdogan, A. M. (2019). Classification and Segmentation of Watermelon in Images Obtained by Unmanned Aerial Vehicle. ELECO 2019 - 11th International Conference on Electrical and Electronics Engineering, 619–622. https://doi.org/10.23919/ELECO47770.2019.8990605
22. Capizzi, G., Lo Sciuto, G., Napoli, C., Tramontana, E., & Wozniak, M. (2015). Automatic classification of fruit defects based on co-occurrence matrix and neural networks. undefined, 861–867. https://doi.org/10.15439/2015F258
23. Makalesi, A., & Aslan, M. (2021). Derin Öğrenme ile Şeftali Hastalıklarının Tespiti. European Journal of Science and Technology, 23, 540–546. https://doi.org/10.31590/ejosat.883787
24. Doğan, İ. T. F. (2018). The Comparison Of Leaf Classification Performance Of Deep Learning Algorithms. SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES, 1.
25. Salamon, J., & Bello, J. P. (2017). Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. IEEE Signal Processing Letters, 24(3), 279–283. https://doi.org/10.1109/LSP.2017.2657381
26. Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., & Valenzuela, A. (2020). A Review of Convolutional Neural Network Applied to Fruit Image Processing. Applied Sciences, 10(10), 3443. https://doi.org/10.3390/APP10103443
27. Aydın, V. A.(2024). Comparison of CNN-based methods for yoga pose classification. Turkish Journal of Engineering, 8 (1), 65-75.
28. İnik, Ö., et al. (2022). GAZİOSMANPAŞA BİLİMSEL ARAŞTIRMA DERGİSİ (GBAD). Gaziosmanpasa Journal of Scientific Research Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. Retrieved July 5, 2022, from http://dergipark.gov.tr/gbad
29. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A Review on Deep Learning Techniques Applied to Semantic Segmentation. Retrieved April 2017, from https://doi.org/10.48550/arxiv.1704.06857
30. Hemmer, M., Van Khang, H., Robbersmyr, K. G., Waag, T. I., & Meyer, T. J. J. (2018). Fault classification of axial and radial roller bearings using transfer learning through a pretrained convolutional neural network. Designs, 2(4), 1–17. https://doi.org/10.3390/DESIGNS2040056
31. Khvostikov, A., Aderghal, K., Benois-Pineau, J., Krylov, A., & Catheline, G. (2018). 3D CNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies. Retrieved September 15, 2022, from http://arxiv.org/abs/1801.05968
32. Alvarez-Canchila, O. I., Arroyo-Pérez, D. E., Patiño-Saucedo, A., Rostro González, H., & Patĩo-Vanegas, A. (2020). Colombian fruit and vegetables recognition using convolutional neural networks and transfer learning. Journal of Physics: Conference Series, 1547(1), 012020. https://doi.org/10.1088/1742-6596/1547/1/012020
33. Yeh, J. F., Lin, K. M., Lin, C. Y., & Kang, J. C. (2021). Intelligent Mango Fruit Glade Classification Using AlexNet with GrabCut Algorithm. 2021 IEEE International Conference on Consumer Electronics-Taiwan, ICCE-TW 2021. https://doi.org/10.1109/ICCE-TW52618.2021.9603074
34. Han, X., Zhong, Y., Cao, L., & Zhang, L. (2017). Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. Remote Sensing, 9(8), 848. https://doi.org/10.3390/RS9080848
35. Patino-Saucedo, A., Rostro-Gonzalez, H., & Conradt, J. (2018). Tropical fruits classification using an alexnet-type convolutional neural network and image augmentation. Lecture Notes in Computer Science, 11304, 371–379. https://doi.org/10.1007/978-3-030-04212-7_32/COVER
36. Mascarenhas, S., & Agarwal, M. (2021). A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. https://doi.org/10.1109/CENTCON52345.2021.9687944
37. Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2023). Transfer learning for image classification using VGG19: Caltech-101 image data set. Journal of Ambient Intelligence and Humanized Computing, 14(4). https://doi.org/10.1007/s12652-023-08495-5
38. Doğuş Gülgün, O., & Erol, H. (2020). Classification performance comparisons of deep learning models in pneumonia diagnosis using chest X-ray images. Turkish Journal of Engineering (TUJE), 4, 129–141. https://doi.org/10.31127/TUJE.652358
39. Şahin, N., Alpaslan, N., İlçin, M., & Hanbay, D. (2023). Evrişimsel sinir ağı mimarileri ve öğrenim aktarma ile bitki zararlısı çekirge türlerinin sınıflandırması. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 35(1), 321–331. https://doi.org/10.35234/FUMBD.1228883
40. Alibabaei, K., et al. (2022). A review of the challenges of using deep learning algorithms to support decision-making in agricultural activities. Remote Sensing, 14(3). https://doi.org/10.3390/RS14030638
41. Iqbal, A., Usman, M., & Ahmed, Z. (2022). An efficient deep learning-based framework for tuberculosis detection using chest X-ray images. Tuberculosis, 136, 102234. https://doi.org/10.1016/J.TUBE.2022.102234
42. Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and

Statistics, p. 738). Springer. Available online: https://www.springer.com/gp/book/9780387310732
43. Gulzar, Y. (2023). Fruit image classification model based on MobileNetV2 with deep transfer learning technique. Sustainability (Switzerland), 15(3), 1906. https://doi.org/10.3390/SU15031906
44. Asriny, D. M., & Jayadi, R. (2023). Transfer learning VGG16 for classification of orange fruit images. Journal of System and Management Sciences, 13(1). https://doi.org/10.33168/JSMS.2023.0112
45. Chen, H. C., et al. (2022). AlexNet convolutional neural network for disease detection and classification of tomato leaf. Electronics, 11(6), 951. https://doi.org/10.3390/ELECTRONICS11060951
46. Nguyen, T. H., Nguyen, T. N., & Ngo, B. V. (2022). A VGG-19 model with transfer learning and image segmentation for classification of tomato leaf disease. AgriEngineering, 4(4). https://doi.org/10.3390/agriengineering4040056
47. Zhu, H., Yang, L., Fei, J., Zhao, L., & Han, Z. (2021). Recognition of carrot appearance quality based on deep feature and support vector machine. Computers and Electronics in Agriculture, 186, 106185. https://doi.org/10.1016/J.COMPAG.2021.106185