Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 12 NO: 2 PAGE: 531-543 YEAR: 2023
DOI:10.17798/bitlisfen.1263130

Identification of Walnut Variety from The Leaves Using Deep Learning Algorithms

Alper Talha KARADENİZ[1*], Erdal BAŞARAN[2], Yüksel ÇELİK[3]

[1] *Digital transformation and software department, Trabzon University, Trabzon, Turkey*
[2] *Department of Computer Engineering, Karabuk University, Karabuk, Turkey*
[3] *Department of Computer Technology, Agri Ibrahim Cecen University, Agri, Turkey*
(ORCID ID 0000-0003-4165-3932) (ORCID ID:0000-0001-8569-2998) (ORCID ID:0000-0002-7117-9736)

Keywords: Walnut Dataset, CNN, Machine learning, Deep Learning

Abstract
In order to determine the variety from walnut leaves, each leaf must be examined in detail. Species that are very similar in color and shape to each other are very difficult to distinguish with the human eye. Examining and classifying plant leaves belonging to many classes one by one is not appropriate in terms of time and cost. Studies on walnut varieties in the literature are generally classified as a result of experimental studies in the laboratory environment. There are two or three different classes in studies using walnut leaf images. In this study, firstly, a unique walnut dataset obtained from 1751 walnut leaf images obtained from 18 different walnut varieties was created. Classification was made using deep learning methods on the original walnut dataset. It has been tested with CNN models, which are widely used in the literature, and some performance metrics are recorded and the results are compared. The images were first preprocessed for cropping, denoising and resizing. Classification was made using CNN models on the original dataset and augmented dataset with data augmentation method. It was seen that the VGG16 CNN model gave the best results both in the original dataset and the augmented dataset. In this model, the accucarcy result found with the original data set was 0.8552, while the accuracy result in the enhanced data set was 0.9055. When the accuracy values are examined, it is seen that walnut varieties are classified successfully.

1. Introduction

Plants are one of the essential resources for our world, and these resources need to be transported to the future healthily [1]. Demand for food crops is increasing due to the increasing global population and the challenges posed by climate change. However, as the need for agricultural nutrients increases, the costs must be minimal. Plants with the appropriate genotype should be selected to use the resources effectively. This will help increase productivity and efficiency. The automatic and correct recognition of walnut varieties is important for agricultural engineers and walnut growers.

The diagnosis of plant leaves is utilized by a detailed examination of each leaf. It isn't easy to distinguish species similar in color and shape from each other with the human eye. Examining and classifying one by one plant leaves belonging to many classes is not appropriate in terms of time and cost. Therefore, taking leaf images and automatically diagnosing and classifying them in the computer environment provides a lot of convenience in terms of time and cost [2]. By using artificial intelligence techniques, software studies have been carried out to provide faster and more accurate results than the functions performed by the human eye [3].

Feature extraction is a difficult process for machine learning. However, it is necessary for the classification process and affects the classification performance [4]. With the developing technology, the speeds and capacities of Central Processing Units (CPUs) and Graphics Processing Units (GPUs) have increased. With these developments, serious performances have been achieved in the data processing stage, leading to the emergence of

* Corresponding author: alperkaradeniz@trabzon.edu.tr
Received: 10.03.2023, Accepted: 09.05.2023

deep learning architectures [5]. On the other hand, deep learning-based studies for the detection and classification of plant leaf diseases have been studied to evaluate their deep learning potential [6].

Considering the mentioned reasons and examining the classification studies carried out in recent years, it is seen that deep neural networks, especially convolutional neural networks (CNN), delivered better results compared to traditional machine learning [7]. One of the most important reasons for the widespread use of CNN algorithms is automatic feature extraction [8]. Furthermore, a matrix (raw image) is used as input to the model, not a vector (feature vector) [9].

Vasif Nabiyev et al. developed a method for plant identification using CNN and Transfer Learning. In the Oxford Flowers Dataset study, a fine-tuning approach was used to transfer learning from the ImageNet domain. MobileNetV2 trained on the ImageNet database was used as a pre-trained network, and an accuracy of 0.9897 was achieved. In addition, positive results were obtained by writing a mobile application [10]. Ibtesam M. Dheir et al. The dataset consisting of 2868 images and five different nut species was classified. In the model, there are 4 convolution layers and these layers use the ReLU activation function. After the convolution layers, there is the Max Pooling layer and then the smoothing layer. The first of the last two layers consists of 512 hidden layers and a total of 2,603,205 parameters that can be trained by the network. The last layer is the output layer and Softmax is used as the activation function. A success of 0.98 was achieved in the study [11]. Yixue Liu et al. conducted a study classifying 21 types of grape leaves. After working with preprocessing and CNN algorithms, they developed the Grad-CAM algorithm to analyze the effect of image complementary preprocessing on the classification results and obtained very successful outcomes. As a result of the tests performed using the Googlenet model, the success rate was 0.974 [8]. Daniel Nkemelu et al. worked on classifying 12 different types of plant seedlings. Tests were conducted with K-Nearest Neighbor (KNN), Support Vector Machine (SWM), and a CNN model they created. As a result of tests with several preprocessing and CNN algorithms, the highest accuracy rate of 0.926 was reached [12]. Yu Sun et al designed a 26-layer deep learning model for classification with large-scale data obtained from the natural environment. The proposed model was tested on the BJFU100 dataset and achieved a success rate of 91.78%. Considering the results, it seems that the model is promising for smart forestry [13]. In the first study with the original walnut data set, a new ResNet-based model was proposed. In the proposed model, ResNet architecture was used for feature extraction, Atom Search Optimization algorithm was used for feature selection and SVM was used for classification. As a result of the experimental tests, a success rate of 81.77% was achieved [14].

In addition to plant identification in the literature, disease diagnosis studies from leaf images have become very popular in recent years. Umit Atila et al. proposed the EfficientNet deep learning architecture by using the Plant Village dataset to classify plant leaf diseases and compared the performance of this model with other state-of-the-art deep learning models. The EfficientNet architecture and other deep learning models are trained using the transfer learning approach, and all layers of the models are set to be trainable in transfer learning. As a result of the tests performed, an accuracy of 0.999 was obtained [4]. Rakesh Chandra Joshi et al. suggested an automatic recognition system for the viral infection of Vignamungo, a legume variety usually grown in the Indian subcontinent. This proposed automated system is based on deep learning and is named VirLeafNet. Data used for system training were obtained from images of healthy, mildly infected, and severely infected leaves obtained over multiple periods. Test results of the proposed models; VirLeafNet-1 was found to be 0.912, VirLeafNet-2 0.964 and VirLeafNet-3 0.974 [15]. Lucas M. Tassis et al. proposed an automatic CNN-based model for the detection of lesioned images from coffee trees. In the first stage of the proposed model, Mask R-CNN network was used for segmentation. In the second stage, UNet and PSPNet networks were used for segmentation. In the final stage, ResNet was applied for the classification process. As a result of the experimental tests, the success rate was found to be 942% [16]. A. Anagnostis et al. has created an accurate and fast object detection system that can identify anthracnose leaves in walnut trees for use in real agricultural environments and has achieved a 0.87 verification. It has been concluded that this system is a viable solution for real-time discovery, monitoring, and decision-making [17] [18].

The producers must buy the right walnut saplings to grow walnuts. It is challenging even for experts to distinguish the walnut variety from the leaves. It takes 3-5 years to see fruits in walnut. Therefore, during this period, the producer spends on an undesired walnut variety or cannot be produced in the region. In this case, the producer tries to change the variety by top-working or has to reestablish the walnut orchard. Saplings that are not namely true have been brought to the courts in many places, causing disagreements between the grower and the nursery. Every year, many legal cases are filed between the nursery and the producers who purchased saplings that do not belong to the desired walnut variety.

In this study, classification with CNN models was utilized to identify walnut varieties from walnut leaves. Before giving the data set input to the CNN models, preprocessing methods were applied, and experimental test results were compared. To improve the model, the data augmentation process was applied to the data set, experimental tests were performed again, and the comparison process was utilized.

With this study, many undesirable issues can be prevented, such as the purchase and planting of wrong saplings, time loss until fruiting, loss of seedlings not planted in a suitable climate, and court processes. Thus, more successful establishment and finalization of walnut orchards will be possible.

2. Generating Original Dataset

2.1. Dataset

Our data set was created by sampling from the walnut orchard in the Application Garden of Yalova Atatürk Horticultural Central Research Institute. A total of 1751 leaves from 18 different cultivars were photographed one by one. The nomenclature in the data set is registered by the institute.

Walnut varieties were determined in advance, their leaves were cut from their branches and displayed on a white background. Images were created using the Canon EOS 100D camera, in daylight, close-up and automatically. All leaf images were taken from trees of a predetermined variety on the same day and within a few hours. Imaging is applied the same for each leaf, but there may be differences in the person's posture, sun, shade and angle when taking the photo.

Examples of 18 types of walnuts are shown in Figure 1.

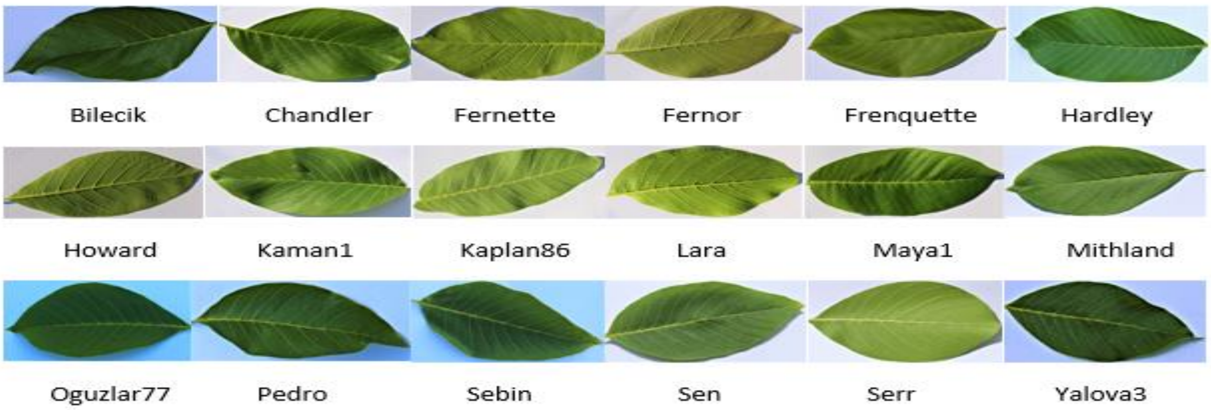

Figure 1. Leaf image examples of walnut species.

The number of leaf images of walnut cultivars in the original walnut dataset is shown in Table 1.

Table 1. Leaf image numbers of walnut species.

Number	Name	Data Count
1	Bilecik	96
2	Chandler	82
3	Fernette	89
4	Fernor	104
5	Frenquette	126

6	Hardley	95
7	Howard	85
8	Kaman1	98
9	Kaplan86	84
10	Lara	63
11	Maya1	74
12	Mitland	147
13	Oguzlar77	59
14	Pedro	77
15	Sebin	88
16	Sen	157
17	Ser	119
18	Yalova3	108
		Total : 1751

2.2. Preprocessing

Before experimental tests with CNN models, preprocessing was applied to the images in the dataset to improve the images. There are errors and noises in the images due to shooting. In order to make a better classification from the raw images taken for the data set, only the area with the leaf was determined. The leaves were cut from the edge lines and preprocessed with the help of image processing methods.

In these preprocesses, in order to convert the first image to black and white, the Localrange of image filter 7.7 neighborhood was applied to the images, and the local ranges of the images were obtained. The most appropriate black and white images were obtained based on these local range values and 20 threshold values.

As a second step, morphological methods were used to eliminate noise in the image. The morphological operations used are erosion, dilation, and closing, respectively. The morphological structuring element is created as a parameter in the erosion and dilation processes. This element is a neighborhood matrix with two-dimensional or multidimensional binary values in which actual pixels are included in the morphological calculation, and false pixels are not. In the Erosion process, a square configuration element with a width of 10 pixels is created. At the same time, the bright regions surrounded by dark-toned regions in the image are narrowed, while the dark-toned regions covered by bright regions are enlarged. After this process, dilation is applied with a square element width of 75 pixels. In the dilation process, while the bright regions surrounded by dark-toned regions in the image expand, the dark-toned regions surrounded by bright regions weaken and even disappear depending on the size of the building element and the dark-toned region. The disk method is applied for the closing process instead of the configuration element. With this method, morphological operations are provided to work faster. The noise on the image is minimized by applying the closing, dilation, and erosion operations to the binary image sequentially.

Objects on the image are labeled and determined on the 2-d binary image (Label connected components in 2-D binary image). Finally, only the leaf region dataset images were obtained by taking the outer frame of the leaf, which is the largest of the objects. Data preprocessing steps are shown in Figure 3.

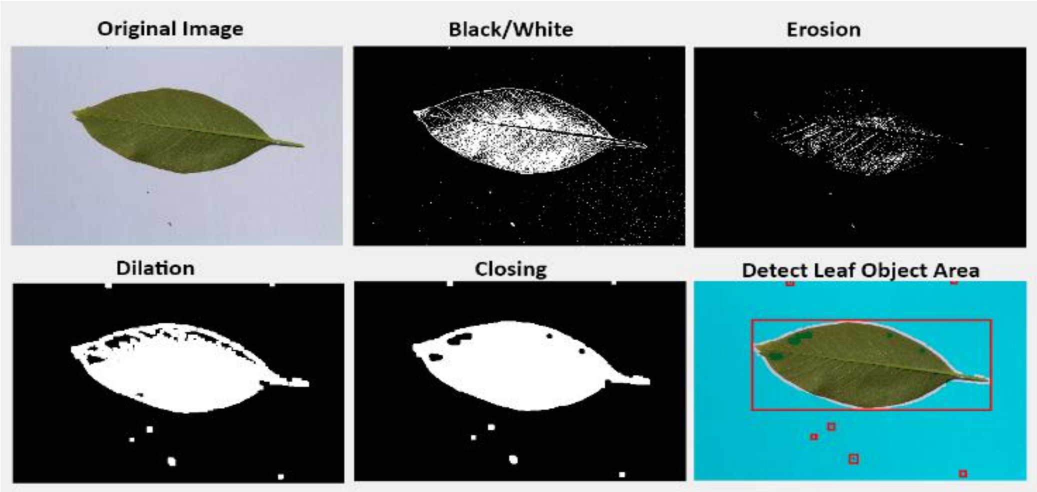

Figure 3. Data preprocessing steps.

After the pre-processing processes were completed, the images were scaled to 600 x 600 dimensions for the experimental tests to run more efficiently.

2.3. Data Augmentation

Data augmentation techniques are a widely used method in deep learning to increase the generalization ability of the model [19]. Data augmentation techniques were applied to our data set consisting of walnut leaf images. These techniques are applied between 0-30 degrees brightness, shift, zoom and flip operations and 1-1.5 degrees rotation. As a result of the data augmentation processes, the number of images in the augmented data set has been increased by approximately 4 times and consists of 6606 images.

3. Material and Method

3.1. CNN Models

CNN models consist of multiple deep layers that do different tasks. CNN models basically consist of 3 layers: convolution, pooling and fully connected layer [18]. In other words, CNNs consist of trainable sections placed one after the other. After receiving the input data in CNN, the training process is carried out by making layer-by-layer operations. Finally, it gives an output to compare the expected value with the generated value. Error occurs as much as the difference between the output and the expected result. This error is transferred to the weights in the network with the back propagation algorithm. The weights are updated at each iteration to reduce the error [5]. The general CNN architecture is shown in Figure 2.

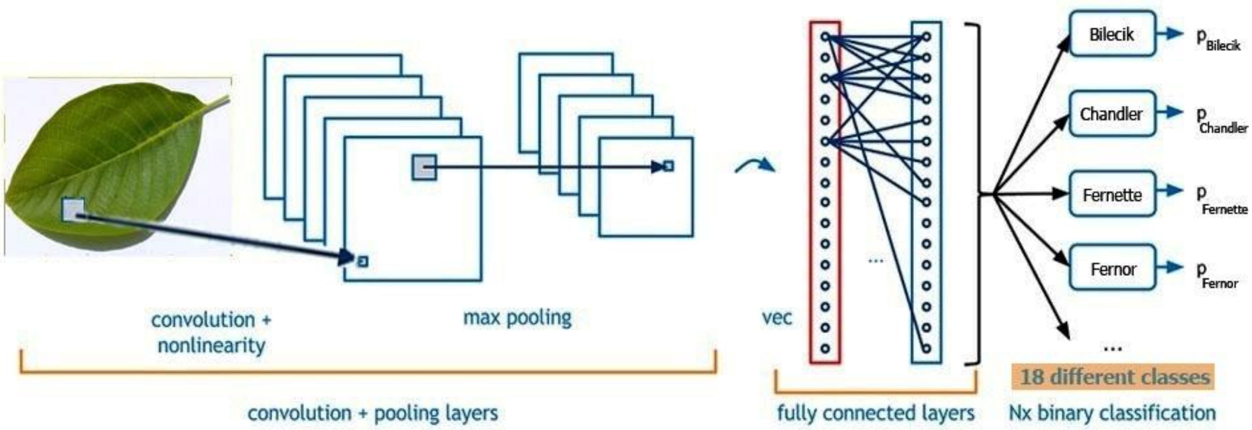

Figure 2. CNN's general architecture.

In the study, Alexnet, InceptionNetV3, VGG16, VGG19, ResNet50, ResNet101, EfficientNet, Darknet19 and GoogleNet CNN models were used for experimental tests.

3.1.1. AlexNet

The AlexNet model is a CNN architecture consisting of 10 layers and approximately 61 million parameters. This CNN architecture was first introduced in the ImageNet competition held in 2012. The first layer of AlexNet is the input layer with an image size of 227 x 227. Then there are 5 convolution layers. After the convolution layer comes 3 fully connected layers. As the last layer, there is the output layer, the Softmax layer. In AlexNet architecture, ReLu activation function is used to increase efficiency between convolution layers and fully connected layer. The Softmax layer is used as the output layer and each output value represents a class value [20] [21].

3.1.2. InceptionV3

InceptionV3 was developed by Google. In addition to previous versions, batch normalization and factorization have been added. The first aim is to reduce the number of parameters and connections to reduce costs, while not reducing the efficiency of the network. Softmax is used in the last layer, the output layer. The Inception V3 architecture consists of 42 layers, including the input layer, which takes a 299×299 pixel image [22].

3.1.3. VGG16

VGG-16 is a CNN architecture with approximately 138 million parameters proposed in 2014. Instead of using many hyper-parameters, this architecture applies 3 x 3 filters and 2 x 2 pooling at each step. There are three layers in the full connection layer, the first two of which are ReLU and the last one is Softmax. VGG-16 contains 16 layers, and the input layer works by taking images of 224 × 224 pixels [23].

3.1.4. VGG19

VGG network architecture was introduced by Simonyan et al. [22]. The VGG19 architecture starts with five block convolutional layers and is configured with three fully connected layers. Convolutional layers are 3 x 3, and ReLU activation is performed after each convolution layer, followed by 2 x 2 pooling. One thousand fully connected layers are used, and the Softmax activation function is used for the output [24].

3.1.5. ResNet50

ResNet50 consists of a network architecture based on a large number of stacked residual volumes. These residual units are used as building blocks to form the ResNet50 network. Each residual unit consists of convolution and pooling layers. ResNet50 CNN architecture, consisting of 224 x 224 pixel input images, was defined in 2015. It is a CNN architecture that is recommended to prevent distortions in inputs with a large number of dimensions [25].

3.1.6. ResNet101

The difference between Resnet101 and Resnet50 is that it has 17 more redundant blocks in the Conv4 layer [26].

3.1.7. EfficientNet

CNN architectures are generally developed with a fixed resource and more resources are used to improve accuracy as needed. This model can identify factors that carefully balance network depth, width, and resolution and perform better by systematically examining scaling. This observation proposes a new scaling method that equally scales all depth/width/resolution dimensions using a simple but highly effective composite coefficient. The model was 8.4 times smaller than the best CNN and 6.1 times faster at inference, while in ImageNet, it found 97.1% accuracy. EfficientNet also achieved an accuracy of 91.7% in the CIFAR-100 dataset and 98.8% in the Flowers dataset [27].

3.1.8. DarkNet19

DarkNet19 is a CNN algorithm capable of clustering up to 1000 clusters. There are 64 layers on the DarkNet19 CNN architecture. These layers are the input layer, Convolution Layer, Batch Normalization (BN), LeakyReLU, maximum pooling, overall average pooling, Softmax, and output layers. The LeakyRelu function, an improved version of the traditional ReLU type, is used for activation [28].

3.1.9. GoogleNet

The size of the receiving area in the GoogleNet network is 224×224, taking RGB color channels with average subtraction. The total number of layers used for the construction of the network is about 100. In GoogleNet architecture, a pooling layer with a filter size of 5 x 5, a 1 x 1 convolution layer with 128 filters, a fully connected layer with 1024 units and ReLU, and Softmax for the classifier are used [29].

4. Experiment Test

4.1. Performance Metrics

There are 18 different types in our data set. Therefore, multiple classifications were made. The confusion matrix is used to measure the performance of this type of classification. Values such as True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) are obtained from the confusion matrix.

Here, TP represents the number of correctly classified images for each category, and TN represents the sum of correctly classified images outside of the category they should be. FN is the number of misclassified images from the appropriate category. FP is the number of images misclassified outside of the intended category.

Performance criteria used in our study; Accuracy (Acc), Precision (Sen), Specificity (Spe), Precision (Pre) and F-Score.

Sensitivity is the ratio of the number of correctly predicted positive images to the total number of positive images. Specificity is the ratio of the number of correctly predicted negative images to the total number of negative images. Accuracy is the ratio of the number of correctly classified samples to the total number of samples. Precision is the ratio of correctly predicted positive outcomes to all positive outcomes. In addition, Recall (Rec) is a metric that shows how many of the transactions we should predict positively are positively predicted. It has the same formula as Recall Sensitivity. The F1-Score value shows us the harmonic mean of the Preve Rec values.

The necessary calculations for the mentioned performance metrics are shown in equations 1-6 [30] [31].

$$Spe = \frac{TN}{TN+FP} \tag{1}$$

$$Sen = \frac{TP}{Tp+FN} \tag{2}$$

$$Acc = \frac{TP+TN}{TP+FN+TN+FP} \tag{3}$$

$$Pre = \frac{TP}{TP+FP} \tag{4}$$

$$Rec = \frac{TP}{TP+FN} \tag{5}$$

$$F1 - Score = 2 * \frac{Pre*Rec}{Pre+Rec} \tag{6}$$

4.2. Training

Our study's original and expanded walnut data sets were randomly divided into 70% training and 30% test sets. The learning rate for all optimization methods was decided as 0.001.

Before starting the training, the images were set with CNN input sizes as 227 × 227 pixels for AlexNet, 224 × 224 pixels for ResNet50, ResNet101, GoogleNet, EfficientNetB0, VGG16 and VGG19, 299 × 299 pixels for Inception V3, and 256 x 256 pixels for DarkNet19.

In experimental studies, the mini-batch size is set to 16. Table 2 shows the hyper-parameters used in all experiments in our research.

Table 2. Hyperparameters used in experimental CNN tests.

CNN Name	Image Size	Learning Rate	Epoch
AlexNet	227 x 227		
VGG16	224 x 224		
VGG19	224 x 224	0.0001	32
DarkNet19	256 x 256		
Inception V3	229 x 229		
EfficientNetB0	224 x 224		
Googlenet	224 x 224		
Resnet101	224 x 224		
Resnet50	224 x 224		

5. Results and Discussions

Our study aims to examine the classification success of the walnut data set we created by using the most popular CNN models in the literature and comparing the CNN models' success rates.

Experimental studies were carried out with all the CNN models mentioned, on both the original and augmented datasets. As a result of these studies, Accuracy, Sensitivity, Specificity, Precision and F1-Score results were found. The values obtained from tests with the original dataset are shown in Table 3. The accuracy rates of the CNN models according to the test results are given in Figure 4.

Table 3. Experimental test results of CNNs

CNN Model	Accuracy	Sensitivity	Specificity	Precision	F1-Score
AlexNet	0.7771	0.7514	0.9869	0.7757	0.7584
DarkNet19	0.7695	0.7362	0.9865	0.7448	0.7377
GoogleNet	0.7524	0.7251	0.9855	0.7330	0.7264
InceptionV3	0.6439	0.6112	0.9790	0.6309	0.6131
ResNet50	0.7695	0.7362	0.9865	0.7575	0.7425
ResNet101	0.7524	0.7197	0.9855	0.7411	0.7272
VGG16	**0.8552**	***0.8315***	***0.9915***	***0.8300***	***0.8363***
VGG19	0.8400	0.8137	0.9906	0.8220	0.8157
EfficientNet	0.7467	0.7230	0.9851	0.7334	0.7245

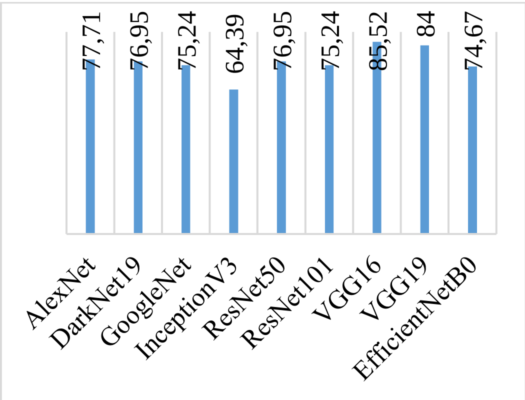

Figure 4. Accuracy rates of CNN models.

As seen in Table 3, VGG16 achieved the highest success with an accuracy rate of 85.52%.

Figure 5 shows the confusion matrix of the highest performing model, VGG16.

DaTa Type	Frenquett	Bilecik	Chandler	Fernette	Fernor	Hardley	Howard	Kaman1	Kaplan86	Lara	Maya1	Mithland	Oguzlar77	Pedro	Sebin	Sen	Serr	Yalova3
Frenquette	33	4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4	0
Bilecik	0	21	0	0	0	1	0	0	0	0	0	0	5	0	0	2	0	0
Chandler	0	0	19	0	0	1	1	0	1	4	5	0	0	1	0	0	0	0
Fernette	3	0	1	27	4	0	0	1	4	1	0	1	1	1	0	1	6	0
Fernor	0	0	1	0	12	1	1	0	0	1	0	0	1	0	2	0	0	0
Hardley	0	2	1	0	5	23	2	0	0	0	0	0	2	3	2	5	5	1
Howard	0	0	0	0	2	0	13	0	0	0	0	0	0	0	0	0	0	0
Kaman1	0	0	2	0	7	1	5	28	2	4	3	0	0	1	0	0	1	0
Kaplan86	0	0	0	0	0	0	0	0	4	0	8	0	0	0	0	0	0	0
Lara	0	0	0	0	0	0	4	0	1	9	1	0	0	0	0	0	0	0
Maya1	0	0	0	0	0	0	0	0	13	0	5	1	0	0	0	0	0	0
Mithland	2	0	1	0	0	0	0	0	0	0	0	42	2	1	0	1	3	0
Oguzlar77	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
Pedro	0	0	0	0	0	1	0	0	0	0	0	0	0	16	0	0	0	0
Sebin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0
Sen	0	2	0	0	0	0	0	0	0	0	0	0	2	0	0	38	2	6
Serr	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	15	0
Yalova3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25

Figure 5. Confusion Matrix obtained with the VGG16.

Table 4. Experimental test results of CNNs with Augmented Dataset

CNN Model	Accuracy	Sensitivity	Specificity	Precision	F1-Score
AlexNet	0.8631	0.8504	0.9920	0.8543	0.8518
DarkNet19	0.8992	0.8872	0.9941	0.8878	0.8870
GoogleNet	0.8675	0.8512	0.9922	0.8605	0.8532
InceptionNetV3	0.8061	0.7919	0.9886	0.7968	0.7928
ResNet50	0.8704	0.8547	0.9924	0.8634	0.8577
ResNet101	0.8583	0.8424	0.9917	0.8520	0.8460
VGG16	***0.9055***	0.8900	***0.9945***	***0.8894***	0.8891
VGG19	0.9045	***0.8918***	0.9944	0.8941	***0.8924***
EfficientNetB0	0.8505	0.8341	0.9912	0.8408	0.8365

The graphs of the accuracy rates of the test results of the augmented dataset and CNN models are given in Figure 6. As seen in Table 4, as a result of the experimental tests performed with the augmented data set, VGG16 granted the highest success with an accuracy rate of 90.55%, as in the original data set. However, VGG19 achieved better results on some performance metrics having the best results in Sensitivity and F1-Score. In addition, VGG16 and VGG19 produced very close results in almost all metrics. DarkNet19 also performs quite well relative to VGG models.

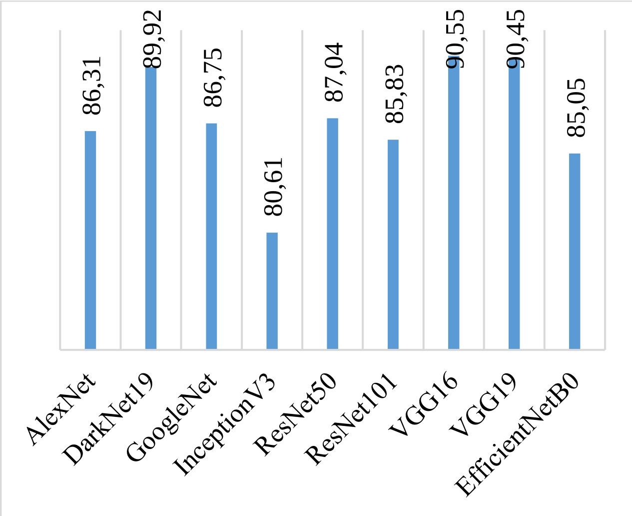

Figure 6. Accuracy rates of CNN models with the augmented dataset.

Figure 7 shows the confusion matrix of VGG16, which delivered the highest performance in the experimental test results in the augmented data set.

DaTa Type	Frenquette	Bilecik	Chandler	Fermette	Fermor	Hardley	Howard	Kaman1	Kaplan86	Lara	Maya1	Mithland	Oguzlar77	Pedro	Sebin	Sen	Serr	Yalova3
Frenquette	144	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0
Bilecik	0	106	0	0	1	0	0	0	0	0	0	0	0	0	0	2	1	3
Chandler	0	1	92	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0
Fernette	0	0	0	104	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Fernor	0	1	0	4	109	0	2	2	2	3	0	0	0	0	0	0	0	0
Hardley	0	0	3	0	0	103	0	2	0	0	0	0	2	1	0	1	0	0
Howard	0	0	0	1	3	0	94	1	0	1	0	0	0	0	0	0	0	0
Kaman1	0	0	0	0	0	0	1	113	0	0	0	0	0	1	0	0	0	0
Kaplan86	0	0	2	0	0	1	0	0	49	3	43	0	0	1	0	0	0	0
Lara	0	0	1	2	1	0	4	1	0	62	1	0	0	0	1	0	0	2
Maya1	0	0	2	0	0	0	1	0	44	1	40	0	0	0	0	0	0	0
Mithland	1	0	0	0	0	0	0	0	0	0	0	169	0	0	0	0	0	0
Oguzlar77	0	0	0	0	0	0	0	0	0	0	0	0	67	0	0	1	0	2
Pedro	0	1	4	1	0	3	1	0	1	0	0	0	0	79	1	0	0	0
Sebin	0	0	0	0	0	0	0	1	0	0	0	0	6	0	95	0	2	0
Sen	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	177	1	2
Serr	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	136	1
Yalova3	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	120

Figure 7. Confusion matrix of VGG16 according to augment dataset.

As can be seen in Tables 3 and 4, the performance of the augmented dataset is better in all CNN models than in the original dataset. In addition, VGG19 Sensitivity and F1-Score performance metrics in the augmented data set offered better results than VGG16 indicating the importance of the number of images in the data set for CNN models.

6. Conclusions

Identification of the walnut variety from the leaves can be utilized as a result of a detailed examination of each leaf. Since these leaves are very similar in color, shape, and texture, it is difficult to distinguish them by traditional methods. Within the scope of this study, a unique walnut dataset containing 18 different classes and 1751 walnut leaf images has been brought to the literature. The original and augmented version of the data set were classified separately using nine different CNN models in the literature. The performance results were compared. Looking at all performance calculation metrics in both datasets, VGG16 was the best performing CNN model. While the accuracy result of VGG16 found with the original data set was 0.8552, the accuracy in the augmented data set was 0.9055. Considering the success rates, walnut varieties were classified successfully with the deep learning methods.

In the future, in addition to the existing CNN models, developing a different and more successful deep learning model is aimed to use as a mobile application for nurseries and walnut growers.

Dataset availability Link to access the data set: https://github.com/TechResearchLab/Walnut-Leaves-Dataset

Acknowledgment

I would like to thank my father Prof. Dr. Turan KARADENİZ, my sister, I know Dr. Instructor Tuba BAK, Director of the Institute, who allowed us to take the walnut images from the application garden of Yalova Atatürk Horticultural Central Research Institute. We would like to thank Yılmaz BOZ and the staff of the Institute.

Contributions of the authors

This article was produced from a PhD thesis.

Conflict of Interest Statement

This article was produced from the thesis work of the first author and second and third authors are thesis advisors.
The authors declare that there is no conflict to interest related to this paper.
Ethics committee approval is not required for the prepared article.

References

[1] K. Kulkarni, Z. Zhang, L. Chang, J. Yang, P. C. A. da Fonseca, and D. Barford, "Building a pseudo-atomic model of the anaphase-promoting complex," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 69, no. 11, pp. 2236–2243, 2013.
[2] F. Doğan and İ. Türkoğlu, "Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması," *Sak. Univ. J. Comput. Inf. Sci.*, vol. 1, no. 1, pp. 10–21, 2018.
[3] S. Solak and U. Altınışık, "Görüntü işleme teknikleri ve kümeleme yöntemleri kullanılarak fındık meyvesinin tespit ve sınıflandırılması," *Sak. Univ. J. Sci.*, vol. 22, no. 1, pp. 56–65, 2018.
[4] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecol. Inform.*, vol. 61, p. 101182, 2021.
[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[6] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease detection and classification by deep learning," *Plants*, vol. 8, no. 11, p. 468, 2019.
[7] B. Yanıkoğlu and E. Aptoula, "İçeriğe Dayalı İmge Erişim Yöntemleri ile Bitki Tanıma," 2016.
[8] Y. Liu, J. Su, G. Xu, Y. Fang, F. Liu, and B. Su, "Identification of grapevine (vitis vinifera l.) cultivars by vine leaf image via deep learning and mobile devices," 2020.
[9] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, and N. Carvalhais, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.
[10] T. Karahan and V. Nabiyev, "Plant identification with convolutional neural networks and transfer learning," *Pamukkale Üniversitesi Mühendislik Bilim. Derg.*, vol. 27, no. 5, pp. 638–645, 2021.
[11] I. M. Dheir, A. Soliman, A. Mettleq, and A. A. Elsharif, "Nuts Types Classification Using Deep learning," *Int. J. Acad. Inf. Syst. Res.*, vol. 3, no. 12, pp. 12–17, 2019.
[12] D. K. Nkemelu, D. Omeiza, and N. Lubalo, "Deep convolutional neural network for plant seedlings classification," *arXiv Prepr. arXiv1811.08404*, 2018.
[13] Y. Sun, Y. Liu, G. Wang, and H. Zhang, "Deep learning for plant identification in natural environment," *Comput. Intell. Neurosci.*, vol. 2017, 2017.
[14] A. T. Karadeniz, Y. Çelik, and E. Başaran, "Classification of walnut varieties obtained from walnut leaf images by the recommended residual block based CNN model," *Eur. Food Res. Technol.*, vol. 249, no. 3, pp. 727–738, 2023.
[15] R. C. Joshi, M. Kaushik, M. K. Dutta, A. Srivastava, and N. Choudhary, "VirLeafNet: Automatic analysis and viral disease diagnosis using deep-learning in Vigna mungo plant," *Ecol. Inform.*, vol. 61, p. 101197, 2021.
[16] L. M. Tassis, J. E. T. de Souza, and R. A. Krohling, "A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images," *Comput. Electron. Agric.*, vol. 186, p. 106191, 2021.
[17] A. Anagnostis *et al.*, "A deep learning approach for anthracnose infected trees classification in walnut

orchards," *Comput. Electron. Agric.*, vol. 182, p. 105998, 2021.

[18] A. Anagnostis, G. Asiminari, E. Papageorgiou, and D. Bochtis, "A convolutional neural networks based method for anthracnose infected walnut tree leaves identification," *Appl. Sci.*, vol. 10, no. 2, p. 469, 2020.

[19] A. Dobrescu, M. V. Giuffrida, and S. A. Tsaftaris, "Doing more with less: a multitask deep learning approach in plant phenotyping," *Front. Plant Sci.*, vol. 11, p. 141, 2020.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Prepr. arXiv1409.1556*, 2014.

[24] T. Carvalho, E. R. S. De Rezende, M. T. P. Alves, F. K. C. Balieiro, and R. B. Sovat, "Exposing computer generated images by eye's region classification via transfer learning of VGG19 CNN," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 866–870, 2017.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[26] Y. Dong, H. Zhang, C. Wang, and Y. Wang, "Fine-grained ship classification based on deep residual learning for high-resolution SAR images," *Remote Sens. Lett.*, vol. 10, no. 11, pp. 1095–1104, 2019.

[27] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, pp. 6105–6114, 2019.

[28] W. Setiawan and A. Purnama, "Tobacco Leaf Images Clustering using DarkNet19 and K-Means," in *2020 6th Information Technology International Seminar (ITIS)*, pp. 269–273, 2020.

[29] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[30] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.

[31] E. Başaran, "Classification of white blood cells with SVM by selecting SqueezeNet and LIME properties by mRMR method," *Signal, Image Video Process.*, pp. 1–9, 2022.