Research Article/Araştırma Makalesi

gazi JOURNAL OF ENGINEERING SCIENCES

Gazi AKADEMİK YAYINCILIK

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Measuring the Effect of Data Augmentation Methods for Improving the Success of Convolutional Neural Network

Kürşad Uçar*[a], Hasan Erdinç Koçer[b]

Submitted: 07.06.2022 Revised: 17.08.2022 Accepted: 01.09.2022 doi:10.30855/gmbd.0705031

Keywords: Classification rate, Convolutional Neural Network, Data augmentation, Deep learning

[a] Selcuk University, Technology Faculty, Dept. of Electric Electronic Engineering 42030 - Konya, Türkiye
Orcid: 0000-0001-5521-2447
e mail: kucar@selcuk.edu.tr

[b] Selcuk University, Technology Faculty, Dept. of Electric Electronic Engineering 42030 - Konya, Türkiye
Orcid: 0000-0002-0799-2140

*Corresponding author: kucar@selcuk.edu.tr

ABSTRACT

With the intensive work done, deep learning finds many useful areas. However, obtaining a sufficient amount of data required by deep learning is not always an easy task. To overcome this difficulty, deep network trainers prefer to develop their datasets by using a set of algorithms. With the increased amount of data, deep networks can be trained more successfully. Data augmentation (DA) is one of the most widely used methods of increasing the amount of data. With DA, the number of sounds and images that a convolutional neural network (CNN) can classify can be increased. In this study, the number of images belonging to 6 classes that do not have enough images to train the CNN successfully enough was increased by DA methods. First, the amount of data was increased by applying three different DA methods separately and all three together. The original dataset and created datasets in which DA methods were used are used to train 15 CNNs with different parameters. Then, their effects on CNN have been investigated. As a result, a success increase of over 5% was observed by increasing the data.

Evrişimsel Sinir Ağlarının Başarısının Artırılmasında Veri Arttırma Yöntemlerinin Etkisinin Ölçülmesi

ÖZ

Yapılan yoğun çalışmalarla derin öğrenme birçok kullanım alanı bulmaktadır. Ancak derin öğrenmenin gerektirdiği yeterli miktarda veriyi elde etmek her zaman kolay bir iş değildir. Bu zorluğun üstesinden gelmek için derin ağ eğiticileri, bir dizi algoritma kullanarak veri kümelerini geliştirmeyi tercih ederler. Artan veri miktarı ile derin ağlar daha başarılı bir şekilde eğitilebilir. Veri artırma (DA), veri miktarını artırmanın en yaygın kullanılan yöntemlerinden biridir. DA ile bir evrişimsel sinir ağının (CNN) sınıflandırabileceği ses ve görüntü sayısı artırılabilir. Bu çalışmada, CNN'yi yeterince başarılı bir şekilde eğitmek için yeterli görüntüye sahip olmayan 6 sınıfa ait görüntü sayısı DA yöntemleri ile artırılmıştır. İlk olarak, üç farklı DA yöntemi ayrı ayrı ve üçü birlikte uygulanarak veri miktarı artırılmıştır. DA yöntemlerinin kullanıldığı orijinal veri seti ve oluşturulan veri setleri, 15 CNN'yi farklı parametrelerle eğitmek için kullanılmıştır. Daha sonra CNN üzerindeki etkileri araştırılmıştır. Sonuç olarak veriler artırılarak %5'in üzerinde bir başarı artışı gözlemlenmiştir.

Anahtar Kelimeler: Sınıflandırma oranı, Evrişimli Sinir Ağı, Veri arttırma, Derin öğrenme

To cite this article: K. Ucar and H. E. Kocer, “Measuring the Effect of Data Augmentation Methods for Improving the Success of Convolutional Neural Network,” *Gazi Journal of Engineering Sciences,* vol. 8, no. 3, pp. 430-438, 2022. doi:10.30855/gmbd.0705031

1. Introduction

Today, machines can easily and quickly perform quite different tasks, such as disease detection [1], object separation [2], face recognition [3], and ripe fruit detection [4]. Machine learning and deep learning algorithms have a large share for the machines reaching this stage. Compared to machine learning, deep learning has made a breakthrough in the solution of classification problems with its emergence [5]. Especially in image classification and pattern recognition areas, the superior success of convolutional neural networks (CNNs) compared to classical methods is seen [6]. Deep learning makes it possible to work with a large amount of data compared to machine learning and classify it with higher accuracy. For this reason, it is a known fact that deep learning algorithms give better results if they have more data that can be classified [7], [8]. To increase deep learning performance, many different data are needed [9]–[12]. However, as it is not easy to overcome this difficulty, sufficient data cannot be obtained to train the classifiers [13]. Especially in fields such as military and medical imaging, it is not possible to obtain sufficient data due to reasons such as protocol and confidentiality [14]. In addition, even if it is desired to increase the amount of data through data collection, it may not always be possible to meet this need because the collected data will start to look similar to each other. At the same time, collecting large amounts of data and labeling them are both tedious and time-consuming. For this reason, data replication or data augmentation (DA) [15], [16] is preferred in classification applications with deep networks where data deficiency is encountered. DA is the process of obtaining new data with some transformations while preserving the labels of existing data [15]. In [17], since the images were destroyed in the pre-processing before CNN processing, the decrease in the dataset was eliminated with the images produced with the Generative adversarial network (GAN). Similarly, synthetic data was produced by using DA methods to alleviate the class imbalance in [18]. Since DA has many options and variables, it is important to choose the right method or approach. For this purpose, which DA method is more suitable for the dataset was investigated in [19]. Otherwise, it will cause the classifier to misclassify or delay the training process due to the data that does not affect it. Even data with little or no effect on the classifier increase the training period and are generally harmful to the user. Another effect of DA on CNN classification is that it improves generalization performance [7], [20], [21]. Thus, overfitting, which is one of the main problems of artificial intelligence applications, is prevented. With generalization, the classification success of the deep network is valid not only for training and test data but also for data that it has never seen.

In this paper, the effect of different DA approaches on deep learning is examined. First, three datasets were created by applying the traditional DA methods: shifting, rotating, and flipping separately. Then, by applying these three methods together, another dataset was obtained. The fifth dataset was created by cleaning the erroneous images that occurred while applying DA. The 15 created deep networks with different parameters were trained and tested with the original dataset and the datasets produced with DA. Thus, by comparing DA methods, we showed that DA increases the success of CNN. In addition, the erroneous images that occurred while applying DA were removed from the dataset, and the effect of this was demonstrated.

The organization of the rest of the paper is as follows. In Section 2, information about the CNNs and DA methods used is given. The created datasets are explained in Section 3. Experiments and results are presented in Section 4. Finally, in the last section, the work is concluded.

2. Material Method

2.1. Structures of convolutional neural networks

Since CNNs are similar to animal systems in terms of topology, the studies of speech and image classifiers have intensified. One of the reasons why CNN is preferred so much is that it does not need a feature extraction step on the raw data to achieve high accuracy. Thus, it is possible to work on raw images. However, since each neuron is linked with all the pixels of the image, it requires high memory and processing power. Spatial stability in the image is obtained by shared weights [22].

Subsampling, which is a stage of CNNs, is an important method in object recognition, as it helps to achieve invariance against image disorders. Local connections on neurons in sequential layers take advantage of spatial dependencies on the image [23]. Connections between layers, convolutional layers, and pooling are the basic building blocks of the CNN architecture. Properties at each position in the input matrix are determined by filters in the convolution layer [24]. By pooling, the dimensions of the features obtained with the convolution layer are reduced. Thus, the calculation speed is increased, and the possibility of overfitting is reduced [25]. Maximum pooling divides the image into rectangles that do not overlap the others and gives the maximum value within the rectangles. Thus, it helps to reduce the amount of data that needs to be worked by reducing the size of the image. Feature maps are created by applying the convolution process to the image. The number of filters used determines the number of feature maps created.

CNNs with different structures were trained to see the effect of the different DA approaches. CNNs with 2, 3, and 4 convolution layers are used. In addition, the performances were examined by changing the filter numbers that affect the performance of the CNN. 3x3 kernel size filters are used in the convolution layers. The CNNs used are shown in Fig 1. In each structure, 2x2 max pooling was applied after the first and second convolution layers. After the third and fourth convolution layers, no pooling was applied. At the end of the convolution layers, there is a fully connected layer with 128 nodes. The output layer consists of 6 nodes.

The numbers of filters on the CNN with two convolution layers are 32-32, 64-64, 32-64, 64-96, and 32-96. The numbers of filters in the CNN with three convolution layers are 32-32-32, 64- 64-64, 16-32-64, 32-64-96, and 64-96-96. The numbers of filters in the CNN with four convolution layers are 32-32-32-32, 64-64-64-64, 32-32-64-64, 16-32-64-96, and 32-48-64-96.

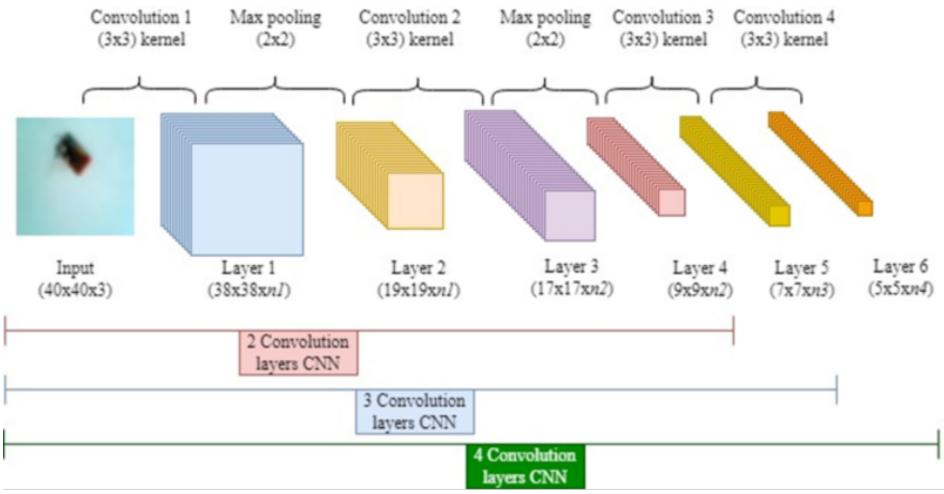

Figure 1. 2, 3, and 4 convolution layers CNN architectures (*n1*, *n2*, *n3*, *n4* are filter numbers)

2.2. Data augmentation

DA means creating new datasets by modifying the original datasets. Synthetic copies are created after processing with various techniques. For images, the amount of data is increased by methods such as flipping, rotation, cropping, shifting, and color tone changes. The other DA methods are used: frequency masking and scaling for sounds, replacement with synonyms for texts, random word deletion, etc. In this study, new datasets were created by applying shifting, flipping, and rotating these among the DA methods.

3. Dataset

The created dataset consists of 6 classes. Classes consist of hand tools, tape, electronic cards, lego, battery, and plant leaves.
Hand Tool: Clamping hand tools of the gripper type constitute this class. They are images of tools such as pliers and side cutters of different sizes and colors.
Tape: Tape images were created using different tape types, such as isolated, double-sided, and duct tape. The classification was made on the circular shapes of the tapes.
Electronic Card: Electronic cards belonging to different electrical circuits were used. As the circuits on them were different, electronic cards with different elements and sizes were used. Some cards also had connection cables. Images of the cards were taken from the surface where only the circuit elements are

seen.
Lego: Random shapes were created using legos as a child's toy. Four different colors were used: yellow, red, blue, and green. The lego class was created with the images taken by creating different combinations in both shape and color.

Table 1. Dataset labels and training-testing image numbers

DA methods	Dataset label	Number of images		
		Train	Val.	Test
-	Dataset1	3360	720	720
Only flipping	Dataset2	11760	2520	2520
Only rotating	Dataset3	11760	2520	2520
Only shifting	Dataset4	11760	2520	2520
Flipping, rotating, and shifting	Dataset5	11760	2520	2520
Flipping, rotating, and shifting (cleared)	Dataset6	11760	2520	2520

Battery: This class consists of images of pen and square batteries belonging to different brands.
Leaf: The leaf was created by viewing the dried leaves of the plane tree. The sizes of the leaves can be different as well as their openings. In addition, some torn leaves were also used.

The elements of the classes were displayed randomly, not in a fixed position and angle, as seen in the sample images given in Fig 2. A total of 4800 images were taken, with 800 images originating from each class. The images taken have 40x40x3dimensions. No special lighting was used for the images, but they were taken under normal room conditions with sunlight. Images were taken at different hours with different light intensities during the day. Objects placed on a white surface were imaged with a fixed distance camera.

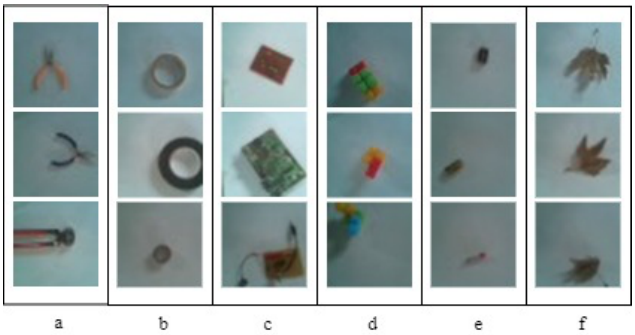

Figure 2. Original dataset, a) hand tools, b) tape, c) electronic card, d) lego, e) battery, and f) leaf.

DA methods were applied to each image randomly. Since randomness was the foremost factor in obtaining the original images, rotation, horizontal shifting, vertical shifting, and flipping processes were also applied to each image randomly to increase the amount of data. To create the augmented data, we employed the following combinations of DA methods to the images:

- only rotating,
- only horizontal and - or vertical shifting,
- only flipping horizontally and - or vertical flipping,
- rotating, shifting, and flipping.

Thus, four separate datasets were prepared, including 2800 images from each class.

Two problems arise with the process of filling the empty regions when creating new images from the images of objects that are beyond or close to the borders of the image. Due to these problems, after the DA methods were applied, some distorted images that do not resemble the real image as in Fig 3 were also formed. The first problem arises if part of the object is beyond the boundaries of the image and the image is shifted or rotated from the side where the object is located to the other side. The filling process is in the color of the object since the regions where no value is assigned by shifting and rotating will be filled with the boundaries of the object. Thus, an image different from the background is created in the border region where the object is located. This problem causes the object to appear as it is not actually. Fig 3(a) and (b) show the distorted samples caused by the filling in the image after the rotation and shifting process, respectively. The second problem is the disappearance of objects completely or partly. This problem also arises when shifting or rotating is done toward the edge where the object is located (Fig 3(c)). Since the images created with such problems will not be healthy for the datasets, they have been removed from the datasets. Another dataset was created with the remaining images after the distorted images were eliminated. New images were produced to replace the deleted images.

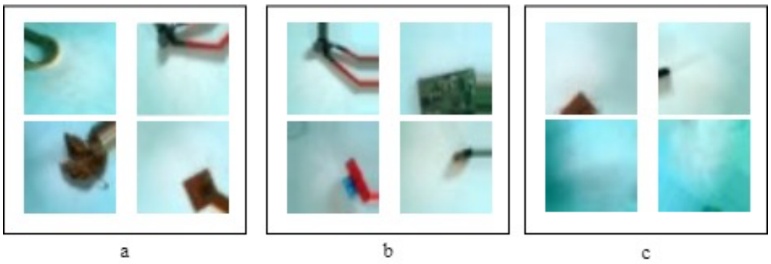

Figure 3. a) Distorted images that occur during the rotation process, b) incorrect images that occur during the shifting process, c) images in which all or some object parts are missing

4. Experiments and Result

Since the study was conducted to show the effect of DA on classification, CNNs with the same parameters were trained and tested with original and increased data. Using the original images, 560 of the 800 images in each class were used for training, while the test and validation data consist of 120 images separately. In the created dataset, 1960 of 2800 images from each class were used for training, and 420 images were used for testing and validation.

Table 1 shows the datasets created and the total number of images in the training, validation, and testing processes. In addition, for the rest of the article, instead of naming the datasets with the related DA methods, they were used with given labels.

CNNs created for the study have three different numbers of convolution layers: 2, 3, and 4. The best result is obtained by trying 5 different combinations of filter numbers in each CNN layer. In all CNNs, the number of nodes in the fully connected layer is 128. As a result, a total of 15 different CNNs were trained and tested with 6 datasets. Each CNN's input is 40x40x3, which is the original size of the images. In Table 2, Table 3, and Table 4, the training and testing accuracy rates of CNNs with two, three, and four convolution layers for each dataset are presented as percentages, respectively. The results are given by choosing the highest test accuracy out of 30 epochs.

Table 2. Train-test accuracies of CNNs with two convolution layers.

Number of filters	**32-32**		**32-64**		**64-64**		**64-96**		**32-96**	
Dataset	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**
Dataset1	84.32	86.39	84.26	84.31	86.67	87.22	85.69	87.92	84.44	88.06
Dataset2	92.58	91.15	95.44	95.32	93.29	93.10	92.98	93.02	93.02	93.21
Dataset3	94.60	93.85	94.84	94.92	94.56	94.21	95.28	95.08	95.36	95.71
Dataset4	90.00	88.14	93.85	93.85	92.50	92.82	93.02	92.30	92.50	93.81
Dataset5	87.54	89.13	88.10	89.13	89.01	90.08	89.05	91.11	88.89	91.03
Dataset6	92.90	92.86	92.38	94.84	92.09	94.09	93.02	94.29	93.17	93.97

Table 3. Train -test accuracies of CNNs with three convolution layers.

Number of filters	**32-32-32**		**64-64-64**		**16-32-64**		**32-64-96**		**64-96-96**	
Dataset	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**
Dataset1	85.42	87.36	89.17	90.28	84.89	85.42	87.36	89.17	90.28	84.89
Dataset2	93.65	94.56	95.99	96.47	94.92	93.65	94.56	95.99	96.47	94.92
Dataset3	96.98	96.51	97.98	97.26	95.08	96.98	96.51	97.98	97.26	95.08
Dataset4	95.24	96.58	96.35	97.06	95.75	95.24	96.58	96.35	97.06	95.75
Dataset5	91.59	93.37	93.65	95.60	89.33	91.59	93.37	93.65	95.60	89.33
Dataset6	95.48	95.99	96.19	97.18	95.00	95.48	95.99	96.19	97.18	95.00

The best result was obtained in Dataset3 in 2 convolution layer CNNs. In general, the highest success was achieved with CNN, whose filter numbers were 32 and 96, respectively. In Table 3, it is seen that the best high accuracy rates are achieved in CNNs with the highest number of filters. It is seen that the accuracy rates are higher for each dataset according to two convolution layer CNN. In four convolution layer CNNs, the highest accuracy rate for all datasets was achieved with two combinations. The train's success has exceeded 99%.

In Table 5, in order to see the effects of the number of convolutional layers and filters on CNN on classification, the average of the accuracy rates of all datasets for each CNN is presented.

Table 4. Train -test accuracies of CNNs with four convolution layers.

Number of filters	**32-32-32-32**		**64-64-64-64**		**32-32-64-64**		**16-32-64-96**		**32-48-64-96**	
Dataset	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**	**Val.**	**Test**
Dataset1	86.25	88.89	90.56	92.50	90.28	86.25	88.89	90.56	92.50	90.28
Dataset2	94.92	95.68	97.86	97.46	96.94	94.92	95.68	97.86	97.46	96.94
Dataset3	96.75	97.22	98.02	97.26	97.86	96.75	97.22	98.02	97.26	97.86
Dataset4	96.31	96.59	97.54	97.98	96.59	96.31	96.59	97.54	97.98	96.59
Dataset5	92.22	92.42	95.71	95.71	94.05	92.22	92.42	95.71	95.71	94.05
Dataset6	95.83	96.83	96.51	96.51	95.91	95.83	96.83	96.51	96.51	95.91

In order to summarize the results of 15 CNNs that were trained and tested according to their datasets, the average of accuracy rates of all CNNs for each dataset is given in Table 6.

As it can be clearly understood from the tables above, higher training and test accuracy rates can be achieved by reproducing data with data enhancement methods. It was observed that datasets created using only one type of DA method (rotation, shifting, or flipping) gave better results in CNN training and testing compared to the combination of three types of DA (rotation + flipping + shifting) method. Considering the experimental results of the study, the highest test accuracy rate was reached in the increased dataset by using the rotation alone. When only one DA method is compared, it is seen that shifting has the lowest accuracy rate. This is because the shifting process distorts the image.

In the last two lines of Table 2-4, the effect of distorted images in data enhancement methods can be seen. Although there are fewer images in Dataset6, it gives results close to Dataset5. In two convolution layer CNNs, the cleaned dataset has a lower accuracy than the uncleaned dataset. This situation reverses as the number of convolution layers increases. While CNNs with 3 convolution layers have reached similar accuracy rates for the mentioned datasets, the cleaned dataset gives better results in 4 convolution layer CNNs. However, even the dataset with distorted images reached higher accuracy than the original dataset.

When the maximum and average test accuracy rates presented in Table 6 are examined, it is seen that the accuracy rate of the original dataset with a lower number of images varies more depending on the CNN structure. This leads to the necessity of choosing the appropriate CNN structure for datasets with low visibility. Although there are distorted images due to rotation in Dataset3, it generally has better classification success than Dataset2. The reason for this is that it is possible to create images different from the original with the rotation process. Thus, the distinctiveness of classes increases.

Table 5. Average accuracy rates of all datasets for each CNN

2 convolution Layers		**3 convolution layers**		**4 convolution layers**	
Number of filters	**Test**	**Number of filters**	**Test**	**Number of filters**	**Test**
32-32	90.25	32-32-32	94.06	32-32-32-32	94.60
32-64	92.06	64-64-64	95.64	64-64-64-64	96.61
64-64	91.92	16-32-64	93.23	32-32-64-64	95.77
64-96	92.29	32-64-96	96.24	16-32-64-96	94.75
32-64	92.63	64-96-96	96.35	32-48-64-96	96.16

Table 6. Averages and maximum accuracy rate for each dataset

	Average Accuracy for 15 CNNs	**Max. Accuracy for 15 CNNs**		
Dataset	**Test**	**Test**	**Number of convolution layer**	**Number of filter**
Dataset1	88.55	92.50	4	64-64-64-64
Dataset2	95.29	97.46	4	64-64-64-64
Dataset3	96.32	98.21	3	64-96-96
Dataset4	95.58	98.25	4	32-48-64-96
Dataset5	93.49	97.30	4	64-64-64-64
Dataset6	95.79	97.42	4	32-48-64-96

It is understood from Table 5 that as the number of convolution layers and filters in the CNN structure increases, CNN performance increases.

Table 7. The confusion matrices

Dataset	Classes	Hand tools	Tape	Electronic card	Lego	Battery	Leaf
Dataset1	Hand tools	0.9547	0.0044	0.0100	0.0056	0.0228	0.0006
	Tape	0.0095	0.8983	0.0206	0.0006	0.0644	0.0167
	Electronic card	0.0044	0.0150	0.8389	0.0028	0.0550	0.0839
	Lego	0.0394	0.0261	0.0167	0.8500	0.0678	0.0000
	Battery	0.0033	0.0806	0.0244	0.0033	0.8817	0.0067
	Leaf	0.0017	0.0156	0.0956	0.0000	0.0000	0.8872
Dataset2	Hand tools	0.9857	0.0052	0.0017	0.0013	0.0060	0.0000
	Tape	0.0027	0.9541	0.0027	0.0000	0.0365	0.0040
	Electronic card	0.0017	0.0268	0.8863	0.0000	0.0348	0.0503
	Lego	0.0040	0.0044	0.0003	0.9817	0.0095	0.0000
	Battery	0.0014	0.0400	0.0122	0.0002	0.9449	0.0013
	Leaf	0.0002	0.0135	0.0186	0.0000	0.0033	0.9644
Dataset3	Hand tools	0.9779	0.0044	0.0014	0.017	0.0144	0.0000
	Tape	0.0019	0.9424	0.0040	0.0006	0.0487	0.0024
	Electronic card	0.0008	0.0097	0.9195	0.0000	0.0379	0.0321
	Lego	0.0011	0.0027	0.0006	0.9881	0.0075	0.0000
	Battery	0.0005	0.0217	0.0035	0.0002	0.9738	0.0003
	Leaf	0.0002	0.0167	0.0154	0.0000	0.0051	0.9627
Dataset4	Hand tools	0.9859	0.0027	0.0014	0.0000	0.0098	0.0002
	Tape	0.0016	0.9283	0.0033	0.0000	0.0625	0.0043
	Electronic card	0.0008	0.0125	0.8967	0.0000	0.0360	0.0540
	Lego	0.0046	0.0005	0.0002	0.9868	0.0079	0.0000
	Battery	0.0032	0.0219	0.0040	0.0003	0.9690	0.0016
	Leaf	0.0003	0.0121	0.0156	0.0000	0.0040	0.9681
Dataset5	Hand tools	0.9797	0.0065	0.0013	0.0019	0.0100	0.0006
	Tape	0.0041	0.8922	0.0067	0.0013	0.0895	0.0062
	Electronic card	0.0016	0.0244	0.8602	0.0000	0.0563	0.0575
	Lego	0.0019	0.0014	0.0003	0.9889	0.0075	0.0000
	Battery	0.0040	0.0425	0.0187	0.0022	0.9316	0.0010
	Leaf	0.0003	0.0171	0.0227	0.0000	0.0032	0.9567
Dataset6	Hand tools	0.9894	0.0016	0.0013	0.0041	0.0037	0.0000
	Tape	0.0073	0.9505	0.0056	0.0011	0.0256	0.0100
	Electronic card	0.0021	0.0100	0.9051	0.0003	0.0484	0.0341
	Lego	0.0010	0.0000	0.0000	0.9963	0.0027	0.0000
	Battery	0.0043	0.0316	0.0183	0.0038	0.9417	0.0003
	Leaf	0.0003	0.0087	0.0248	0.0000	0.0016	0.9646

Confusion matrices were also used to see the classification success of each class. In the study, instead of showing 90 different confusion matrices, it was preferred to show the confusion matrices belonging to each dataset. While calculating these matrices, the results of 15 CNNs tested with a dataset were collected and normalized. Confusion matrices of datasets are shown in Table 7. The most successfully classified class is the lego class. The battery class stands out as the class with the lowest classification success. The most successful classification in Dataset1 belongs to legos. Outside of legos, the highest rate belongs to leaves with 0.8819. As can be understood from here, Dataset1 is not considered very successful in terms of classification. It is seen that the battery class has the lowest classification success. With the data increase in Dataset2, the success of the classification also emerges. The success difference between the battery class with other classes has been closed with the lowest classification success in Dataset1. In general, the highest classification belongs to Dataset3. In fact, the lego class was predicted at 99.93% correctly in Dataset3. Although there are also distorted images in Dataset3 and Dataset4, higher accuracy was achieved with only the rotation process. With the application of more DA in

Dataset5, the classification success is lower than Dataset2, Dataset3, and Dataset4. In addition, Dataset5 has the lowest lego class classification success at 98.78% among all datasets. When confusion matrixes of Dataset5 and Dataset6 are compared, it is seen that cleaning contributes positively to the classification rate of hand tool, lego, and battery classes, where the average accuracy of Dataset6 is lower than Dataset5's.

5. Conclusion

In this study, the effect of DA methods on the success of deep learning was investigated. In this context, 3 different DA methods were applied to images belonging to 6 different classes. Raw data and datasets created by DA methods were separately trained on CNN. In order to see the effect of DA independently from the CNN structure, training and testing procedures were carried out with 15 different CNN structures. When all test results are examined, it is clearly seen that; a higher classification success was achieved with the datasets created with DA compared to the raw datasets. Even with unreal images created due to DA methods, higher success has been achieved than raw data. Thus, the data which does not sufficient for CNN to reach a high accuracy rate was expanded with the DA method, allowing CNN to reach a higher classification rate.

Acknowledgment

This research is supported in part by Selçuk University Coordinatorship of Faculty Member Traning Program via grant number 2019 - ÖYP - 008.

Conflict of Interest Statement

The authors declare that there is no conflict of interest

References

[1] M. Ouhami, A. Hafiane, Y. Es-Saady, M. El Hajji, and R. Canals, "Computer vision, IoT and data fusion for crop disease detection using machine learning: A survey and ongoing research," *Remote Sensing*, vol. 13, no. 13, p. 2486, 2021. doi:10.3390/rs13132486

[2] Ö. Şengül, S. Öztürk, and M. Kuncan, "Color based object separation in conveyor belt using PLC," *Avrupa Bilim ve Teknoloji Dergisi*, no. 18, pp. 401–412, 2020. doi:10.31590/ejosat.685936

[3] F. Tabassum, M. I. Islam, R. T. Khan, and M. R. Amin, "Human face recognition with combination of DWT and machine learning," *Journal of King Saud University-Computer and Information Sciences*, 2020. doi:10.1016/j.jksuci.2020.02.002

[4] A. D. M. Africa, A. R. V Tabalan, and M. A. A. Tan, "Ripe fruit detection and classification using machine learning," *Int J*, vol. 8, no. 5, 2020. doi:10.30534/ijeter/2020/60852020

[5] Y. Q. Lv, K. Liu, F. Cheng, and W. Li, "Visual tracking with tree-structured appearance model for online learning," *Iet Image Processing*, vol. 13, no. 12, pp. 2106–2115, 2019, doi:10.1049/iet-ipr.2018.6517

[6] F. Özyurt, T. Tuncer, E. Avci, M. Koç, and İ. Serhatlioğlu, "A novel liver image classification method using perceptual hash-based convolutional neural network," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3173–3182, 2019. doi:10.1007/s13369-018-3454-1

[7] L. Nanni, M. Paci, S. Brahnam, and A. Lumini, "Comparison of different image data augmentation approaches," *Journal of Imaging*, vol. 7, no. 12, p. 254, 2021. doi:10.3390/jimaging7120254

[8] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017. doi:10.48550/arXiv.1712.04621

[9] Q. Lin *et al.*, "Classifying functional nuclear images with convolutional neural networks: a survey," *IET Image Processing*, vol. 14, no. 14, pp. 3300–3313, 2020. doi:10.1049/iet-ipr.2019.1690

[10] H. Alhichri, Y. Bazi, and N. Alajlan, "Assisting the Visually Impaired in Multi-object Scene Description Using OWA-Based Fusion of CNN Models," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 10511–10527, 2020. doi:10.1007/s13369-020-04799-7

[11] A. Khémiri, A. K. Echi, and M. Elloumi, "Bayesian versus convolutional networks for Arabic handwriting recognition," *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9301–9319, 2019. doi:10.1007/s13369-019-03939-y

[12] F. Zhou, Y. Hu, and X. Shen, "MSANet: multimodal self-augmentation and adversarial network for RGB-D object recognition," *The Visual Computer*, vol. 35, no. 11, pp. 1583–1594, 2019. doi:10.1007/s00371-018-1559-x

[13] A. Fawzi, H. Samulowitz, D. Turaga, and P. Frossard, "Adaptive data augmentation for image classification," in *2016 IEEE international conference on image processing (ICIP)*, 2016, pp. 3688–3692.

[14] R. Hao, K. Namdar, L. Liu, M. A. Haider, and F. Khalvati, "A comprehensive study of data augmentation strategies for prostate cancer detection in diffusion-weighted MRI using convolutional neural networks," *Journal of Digital Imaging*, vol. 34, no. 4, pp. 862–876, 2021. doi:10.1007/s10278-021-00478-7

[15] D. A. Van Dyk and X.-L. Meng, "The art of data augmentation," *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1–50, 2001. doi:10.1198/10618600152418584

[16] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019. doi:10.1186/s40537-019-0197-0

[17] V. Kukreja, D. Kumar, and A. Kaur, "GAN-based synthetic data augmentation for increased CNN performance in Vehicle Number Plate Recognition," in *2020 4th international conference on electronics, communication and aerospace technology (ICECA)*, 2020, pp. 1190–1195.

[18] H. Lee, H. Lee, H. Hong, H. Bae, J. S. Lim, and J. Kim, "Classification of focal liver lesions in CT images using convolutional neural networks with lesion information augmented patches and synthetic data augmentation," *Medical physics*, vol. 48, no. 9, pp. 5029–5046, 2021. doi:10.1002/mp.15118

[19] L. Taylor and G. Nitschke, "Improving deep learning with generic data augmentation," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2018, pp. 1542–1547. doi:10.1109/SSCI.2018.8628742

[20] P. M. Blok, F. K. van Evert, A. P. M. Tielen, E. J. van Henten, and G. Kootstra, "The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R-CNN," *Journal of Field Robotics*, vol. 38, no. 1, pp. 85–104, 2021. doi:10.1002/rob.21975

[21] M. M. A. Monshi, J. Poon, V. Chung, and F. M. Monshi, "CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR," *Computers in biology and medicine*, vol. 133, p. 104375, 2021. doi:10.1016/j.compbiomed.2021.104375

[22] Y. LeCun *et al.*, "Handwritten digit recognition with a back-propagation network," *Advances in neural information processing systems*, vol. 2, 1989.

[23] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. de Leon Ferreira, "Deep learning for biological image classification," *Expert Systems with Applications*, vol. 85, pp. 114–122, 2017. doi:10.1016/j.eswa.2017.05.039

[24] A. Arora, P. Chakraborty, and M. P. S. Bhatia, "Analysis of Data from Wearable Sensors for Sleep Quality Estimation and Prediction Using Deep Learning," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 10793–10812, 2020. doi: 10.1007/s13369-020-04877-w

[25] X. Jiang, S. C. Satapathy, L. Yang, S.-H. Wang, and Y.-D. Zhang, "A Survey on Artificial Intelligence in Chinese Sign Language Recognition," *Arabian Journal for Science and Engineering*, pp. 1–36, 2020. doi:10.1007/s13369-020-04758-2