All Sciences Proceedings
All Sciences Proceedings
http://as-proceeding.com/

1st International Conference on Scientific and Innovative Studies

April 18-20, 2023 : Konya, Turkey

© 2023 Published by All Sciences Proceedings

ICSIS
https://as-proceeding.com/index.php/icsis

Measuring the Effect of Data Augmentation in a CNN-Based Deep Neural Network Model

Halit Çetiner [*], Sedat Metlek [2]

[1]*Vocational School of Technical Sciences, Isparta University of Applied Sciences, Türkiye*
[2]*Vocational School of Technical Sciences, Burdur Mehmet Akif Ersoy University, Türkiye*

[*]*(halitcetiner@isparta.edu.tr)*

Abstract **–** Traditional classification methods have difficulty in meeting the changing needs according to the ever-increasing data piles. With the development of processors with high performance as memory and processing capabilities, deep learning-based methods have been widely used. A large amount of data is needed to train a deep learning-based model, which is a computational science field. CIFAR-10, which contains images of 10 different objects in the world, is a benchmark dataset used effectively in image identification and classification. The proposed deep learning-based models should be tested in a computer environment in order to be used in real life. The proposed model performs the testing process with images that it has never encountered during the training phase. In this article, a deep learning model is proposed that performs classification on the CIFAR-10 dataset, which contains images of objects in the world. An effective classification method has been developed by removing the overfitting effect, if any, on the proposed model. Proposed model, classification process was carried out both with and without data augmentation. The data set used was expanded with random crop, scale transformation, vertical and horizontal flipping data augmentation techniques. In the experimental studies, there was a big difference between the performance of the process using the data augmentation technique and the process without any augmentation. Using different augmentation techniques together or individually did not improve model performance. Proposed model achieved success rates of 91.93%, 93.63% and 90.49%, respectively, including train accuracy, precision, recall. According to the results obtained, it can be said that the study has achieved results that can compete with the literature.

Keywords – Data Augmentation, CNN, CIFAR-10, Deep Learning, Classification

I. Introduction

Classical classification methods have become unable to meet the changing needs of the ever-increasing data stack. With the creation of many open source data sets and the development of high-performance processors in capacity, deep learning-based methods have been widely used. Deep learning-based models, on the other hand, contain very different architectures. In general, CNN-based methods that automatically extract distinctive features are used in image classification. In recent years, Convolutional Neural Network (CNN)-based solutions have been increasing for many different computer tasks, from object segmentation [1] to object detection [2]. The success of deep learning models that provide these solutions depends on many issues, from the number of filters in the convolution layers to the depth of the network. In order to get a better result from similar studies in the literature, careful selection of the parameters of the deep learning model affects the performance. In this study, a well-adjusted deep learning model was developed for the classification of the CIFAR 10 dataset.

Deep neural networks can give better performance results than machine learning techniques using classical feature extraction methods [3]. When the studies in the literature are examined, it is seen that the structure and parameters of the layers used vary depending on the depth of the CNN architectures developed. AlexNet consists of 5 convolution layers and 3 dense layers [4]. The architecture called GoogleNet consists of a single dense layer and a total of 22 layers [5].

With the emergence of AlexNet architecture, which is one of the deep neural networks, CNN-based models are rapidly applied for various tasks in the field of computer vision, especially image recognition, classification and segmentation problems [6]–[8]. Deep learning models can show superior performance in solving problems where big data is increasing day by day. Besides, it deals with the overfitting problem and the gradient vanishing problem. The existence of the stated problem is understood from the failure to achieve the desired success in the test performance results, while good results are obtained in the training performance [3]. In order to prevent the overfitting problem in deep CNN-based networks, the data set is expanded with techniques such as translation, horizontal and vertical flipping [9]–[11]. This expansion process is called data augmentation in the literature. CNN models can be applied on the original data set as well as on the applied data augmentation.

In this article, a deep learning model is proposed that performs classification on the CIFAR-10 dataset, which contains images of objects in the world. An effective classification model has been developed by removing the overfitting effect, if any, on the proposed model. Proposed model, classification process was carried out both with and without data augmentation. Obtained performance results are given comparatively. At the same time, controversial analyzes were carried out by making comparisons with recent studies on the benchmark dataset CIFAR-10. In comparisons, our proposed model achieves state-of-the-art (SOTA) results.

The remainder of the article is organized as follows. In Section 2, similar studies in the literature are examined technically. In Section 3, the data set used, the applied data augmentation techniques, and the proposed model are explained in detail. In Section 4, the performance outputs obtained from two different uses of the proposed model are presented. The performance results obtained with the use of accuracy, precision, recall performance metrics, which are widely used in the literature, are shown with graphics. In Section 5, the study is concluded with the results obtained.

II. Related Works

Data augmentation techniques such as random crop, scale transformation, vertical and horizontal flipping have a significant effect on overfitting the overfitting problem that is frequently encountered in CNN models [12]. Deep learning models can give good performance results when the amount of data in the data set is increased in a balanced way. There are different data augmentation techniques, especially sample pairing, neural augmentation, smart augmentation methods, Generative Adversial Networks (GANs). In the sample pairing method, one of the training images is randomly selected and superimposed on the original image [13]. In real applications, it experiences hindrance due to training tricks. Neural augmentation and smart augmentation methods produce new images by reducing the errors of the weight values of the neural network [14], [15]. Generally, there is a need to increase datasets in the biomedical field. In this sense, studies are carried out to increase CNN performance by increasing medical images with synthetic images. Adar et al. uses GAN architectures to synthesize Region of Interest (ROI) regions containing liver lesion [16]. CNN architecture was used to classify the data set formed with the synthesized images. Although there are many data augmentation techniques, the most appropriate data augmentation techniques have been determined for the given problem. At the same time, the effect of data augmentation techniques on classification was measured and compared with the results obtained without applying any data augmentation technique.

Zhu and Chen performed the performance test of the model they developed based on the densities, on the CIFAR-10 dataset [17]. Krishna and Kalluri performed image classification using pre-trained architectures [18]. The importance of determining suitable lot sizes is emphasized for stable convergence and good test performances [19]. While the small batch size provides a more stable and flexible training, increasing the mini batch size has a positive effect on the test performance [19].

III. MATERIALS AND METHOD

A. *Material*

CIFAR 10 is an image dataset created for use in machine learning and computer vision applications [20]. In the dataset consisting of 10 classes, each image is 32x32 in size. There are 80,000 images in total in the data set, which consists of 6,000 images in each class. A total of 60,000 images, 50,000 of which are training images and 10,000 of which are test images, were used in the data set. Fig. 1 shows some images in the classrooms. These images are generally given the class names in the table specified in Table 1.

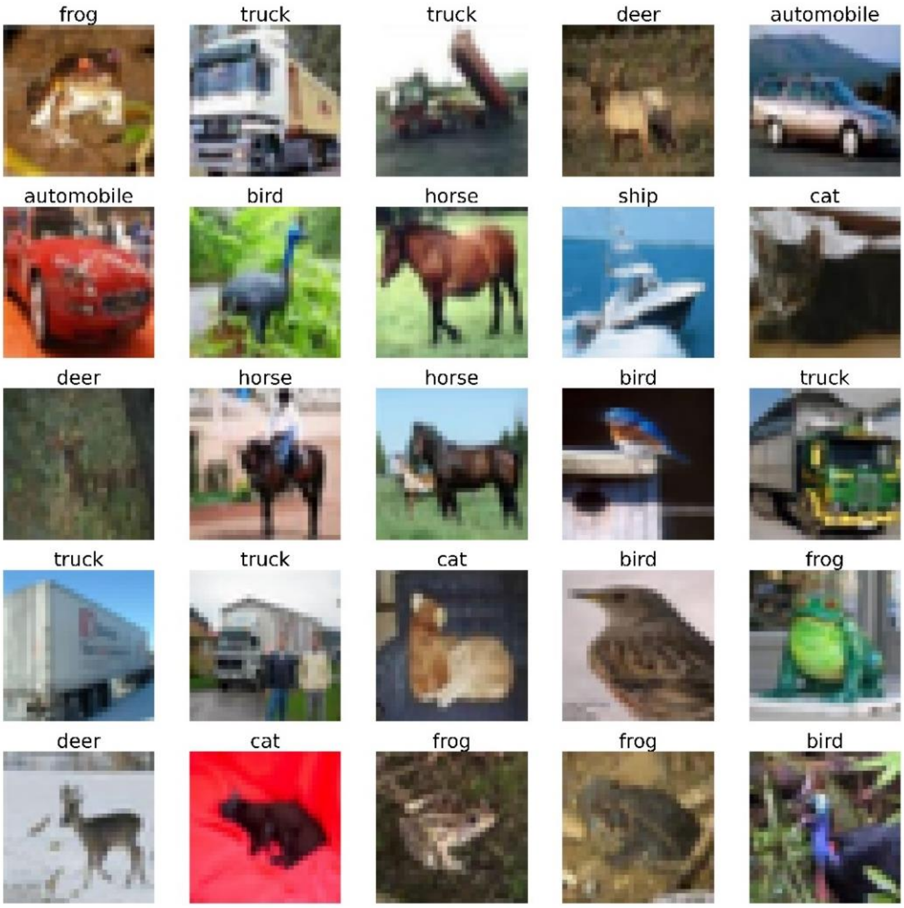

Fig. 1. Some images in the dataset

Table 1. Classes in the dataset

Label	Class Name
0	Air plan
1	Automobile
2	Bird
3	Cat
4	Deer
5	Dog
6	Frog
7	Horse
8	Truck

B. *Proposed Model*

The widely used CIFAR-10 dataset was used to benchmark the models proposed in this article. The performance results obtained with and without data augmentation on the used data set were obtained. The analysis of the model was carried out according to the performance results obtained. 3x3 window size filters and ReLU activation function are used in all of the convolution layers used in the proposed model. 2x2 window sizes are used in the max pooling layers used. In the dropout layers, a neuron drop rate of 0.6 was used. In general, a CNN model consisting of 6 convolution layers, 6 Batch Normalization layers, 3 Max Pooling layers, 5 Dropout layers, 1 Flatten, 1 fully connected layer, a total of 23 layers with 1 classification layer has been proposed. The first layer of the proposed model consists of 32x32x3 images. A convolution layer with 8 filters is defined in the second and fourth layers. In the third, fifth, ninth, eleventh, fifteenth, and seventeenth layers, the Batch Normalization layer is defined. In the sixth, twelfth and eighteenth layers, the max pooling layer has been added. A dropout layer has been applied in the seventh, thirteenth, nineteenth and twenty-second layers. In the twentieth layer, the Flatten layer is defined, which transforms the feature matrices into one-dimensional vectors. In the twenty-first layer, the dense layer with 512 neurons with ReLU activation function is defined. Classification layer with softmax activation function is applied in the twenty-third layer.

.

IV. Results and Discussion

The data set was expanded with random crop, scale transformation, vertical and horizontal flipping data augmentation techniques used in data augmentation. However, in order to measure the effect on the results, two separate experimental studies were carried out on the original image and on the data augmented images. The performance metric results obtained as a result of the experimental studies are shown with graphics in Fig. 2-9 under this section. At the same time, SOTA comparisons are presented in Table 2. In experimental studies, while Adam optimization method was applied, 32 value was preferred as the batch size value. In the training models performed, the number of epochs was determined as 100. The Aug expression in Fig. 2-9 is the abbreviation of Augmentation.

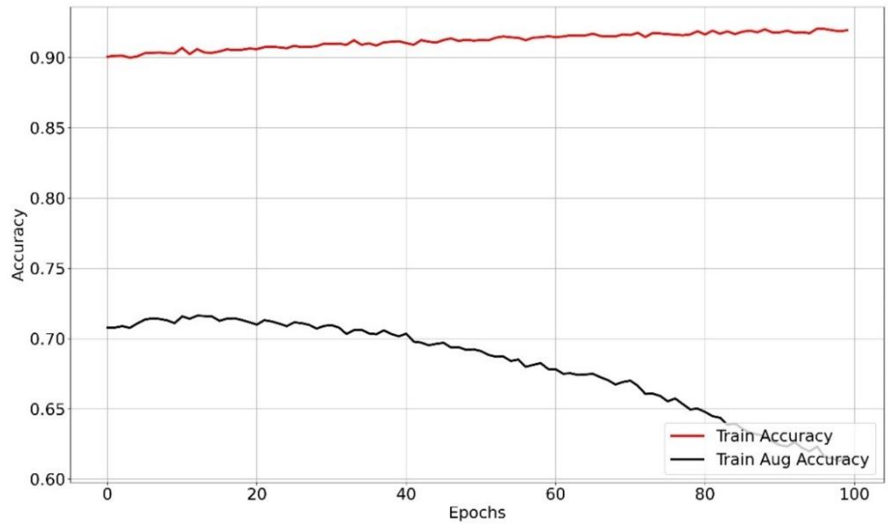

Fig. 2. Train accuracy results of proposed model

In Fig. 2, with the proposed model, a success rate of 91.9% train accuracy was achieved over the original data. However, a success rate of 61.6% was achieved on images with data augmentation.

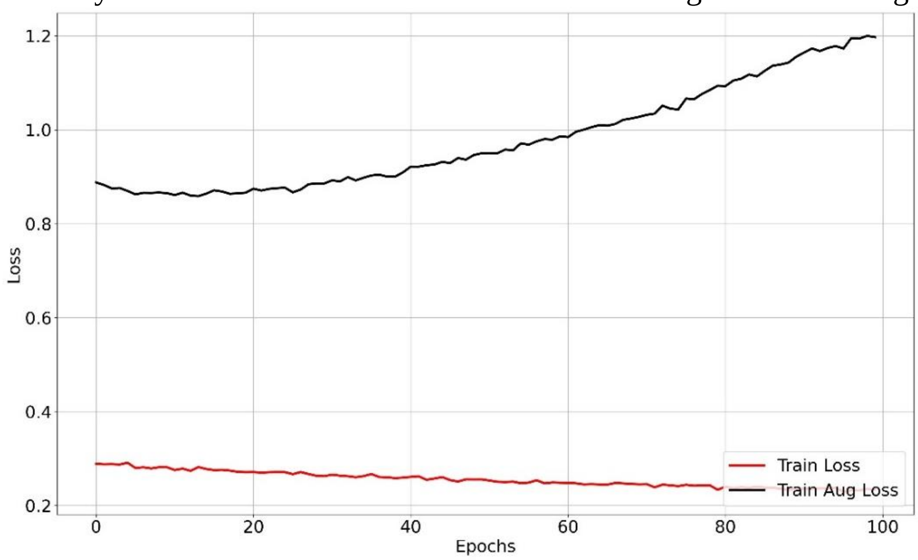

Fig. 3. Train loss results of proposed model

In Fig. 3, a train loss value of 0.23% has been reached over the original data with the proposed model. However, the loss value increased up to 1.19% over the data augmented images.

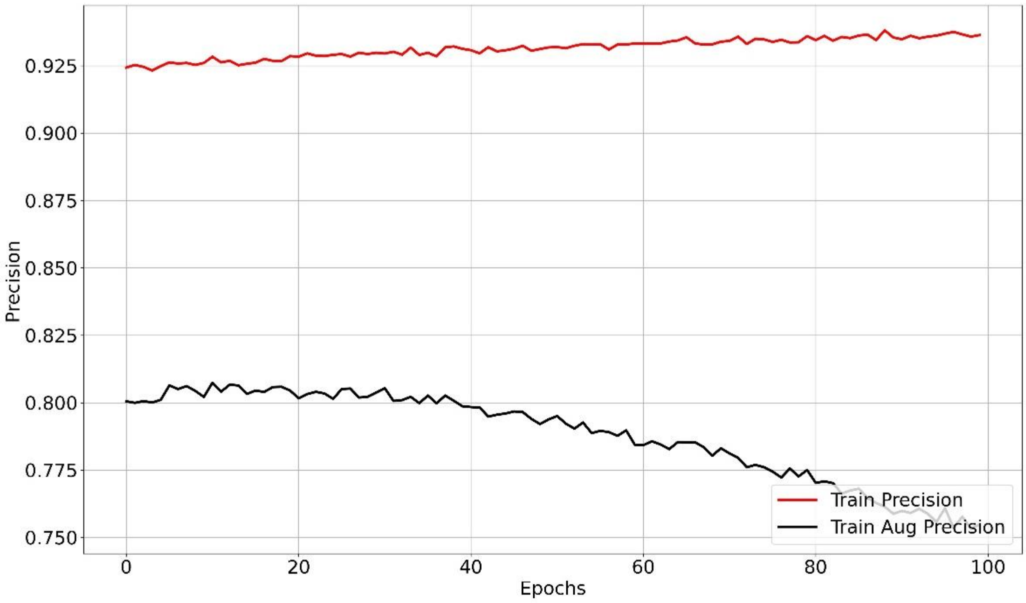

Fig. 4. Train precision results of proposed model

In Fig. 4, 93.63% train precision success rate was achieved over the original data with the proposed model. However, 75.48% train precision rate was achieved on the images with data augmentation.

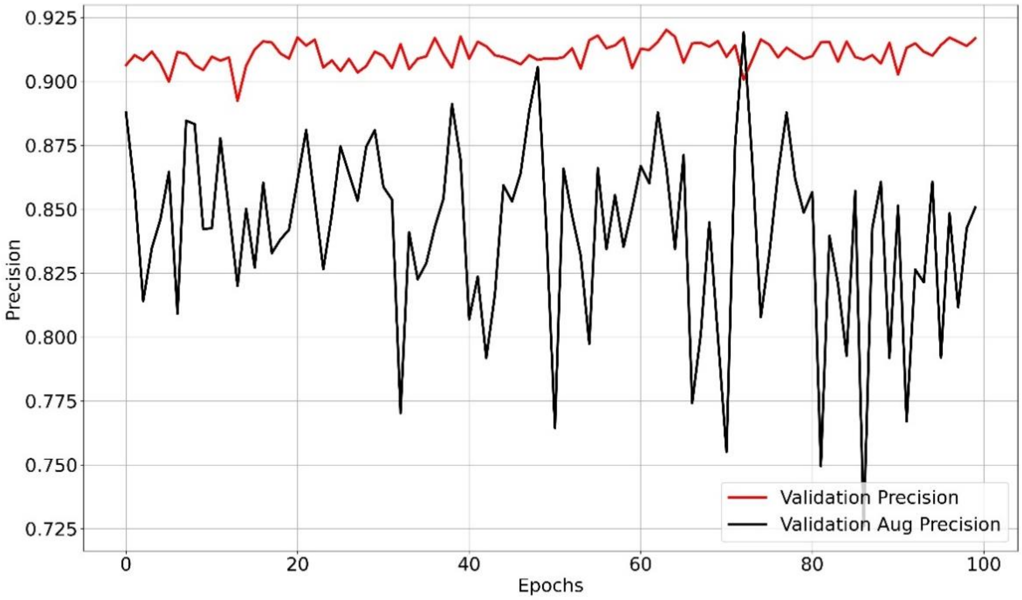

Fig. 5. Validation precision results of proposed model

In Fig. 5, 91.68% validation precision success rate was achieved over the original data with the proposed model. However, 85.07% validation precision was achieved on the images with data augmentation.

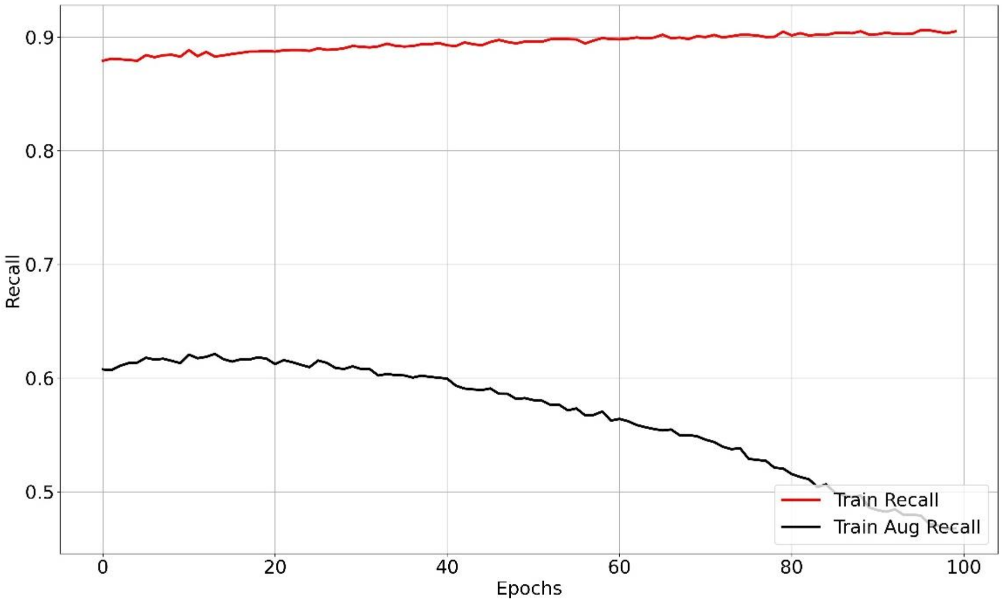

Fig. 6. Train recall results of proposed model

Proposed model achieved a train recall success rate of 90.49% over the original data. However, a train recall rate of 46.78% was achieved on data augmented images. Both results are presented in Fig. 6.

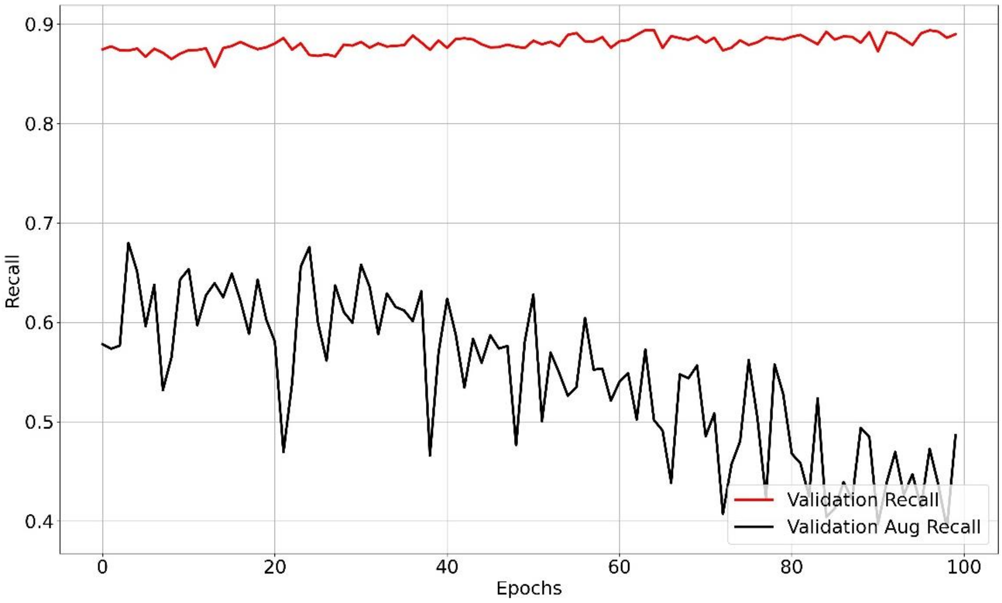

Fig. 7. Validation recall results of proposed model

Proposed model achieved a validation recall success rate of 88.98% over the original data. However, a validation recall rate of 48.62% was achieved on images with data augmentation. Both performance results are presented in Fig. 7.

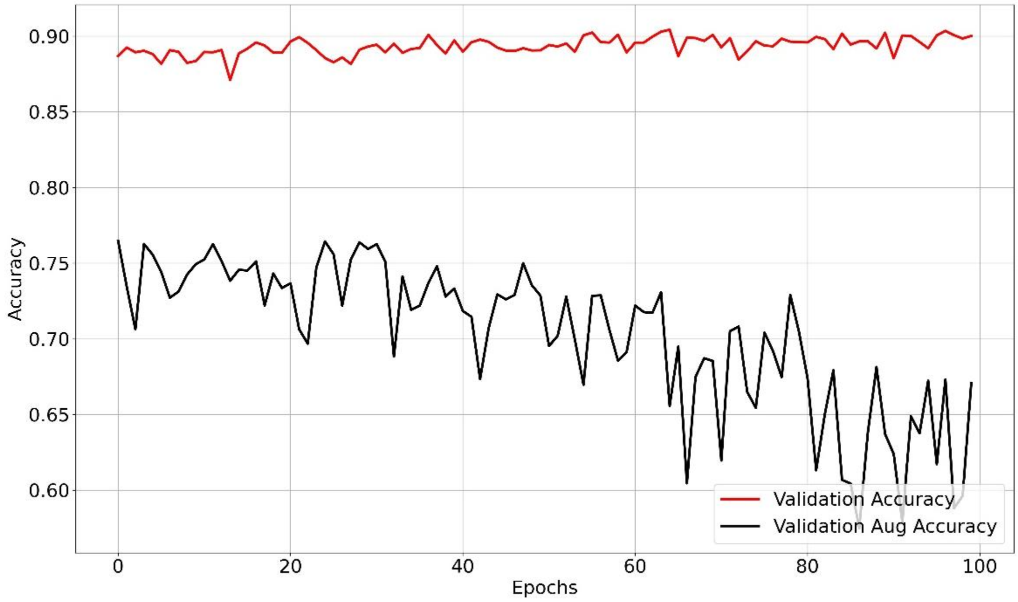

Fig. 8. Validation accuracy results of proposed model

The proposed model achieved 90.02% validation accuracy over the original data. However, a validation accuracy rate of 67.08% was achieved with data augmented images. Both performance results obtained are presented in Fig. 8.

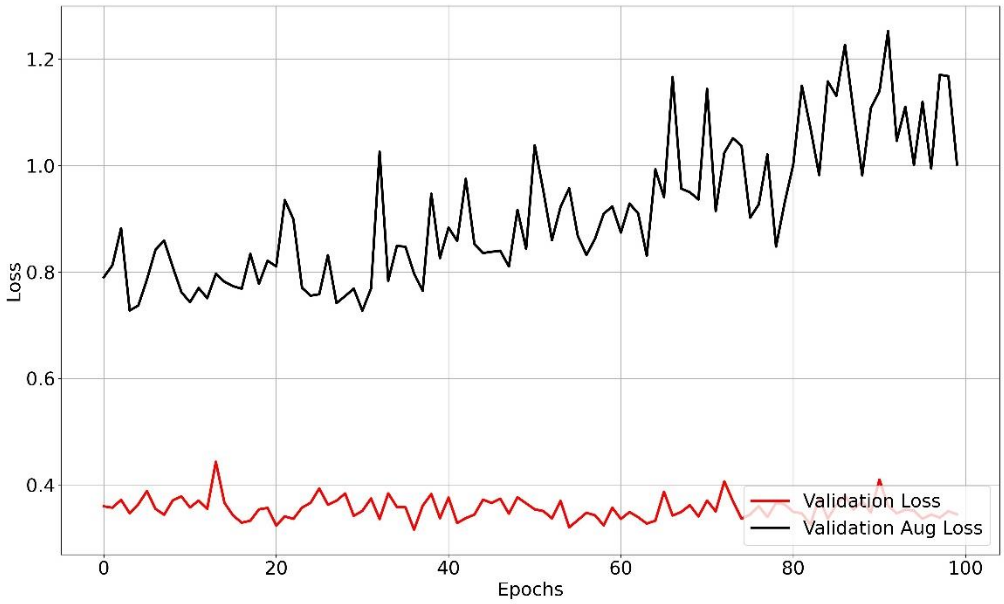

Fig. 9. Validation loss results of proposed model

Proposed model reached a validation loss of 0.34% over the original data. However, the loss value on the data augmented images increased up to 1.00% validation loss value. Both performance results obtained are presented in Fig. 9.

Table 2. Classification performances on the CIFAR-10 dataset

Reference Study	Model	Accuracy (%)	Precision (%)	Recall (%)
[21]	HOG	53		
[21]	PI	59		
[21]	VGG16	60		
[21]	Inception ResNet V2	90.74		
[22]	CNN+AdaBoost	88.4		
[23]	Inception V3	70.1		
[24]	GoogleNet	68.95		
[24]	ResNet50	52.55		
CNN without Augmentation	CNN	91.9	93.63	90.49

Giuste and Vizcarra achieved 53% and 59% accuracy rates with Histogram of oriented gradients (HOG) and Pixel Intensities (PI) based methods [21]. Here, the authors aimed to improve performance by combining the advantages of multiple methods. There are studies using many pre-trained models, including GoogleNet and ResNet50 [24]. Similar to these studies, image classification was performed with the InceptionV3 model. Lee et al. performs image classification with a deep convolution-based CNN model [22]. The model results of the softmax activation function were combined with the Adaboost algorithm. All of the studies with references [21], [22], [23] and [24] compared use the CIFAR-10 dataset.

Table 3. Train performance results of proposed model

Model	Train				Validation			
	Accuracy (%)	Loss (%)	Precision (%)	Recall (%)	Accuracy (%)	Loss (%)	Precision (%)	Recall (%)
CNN with Aug	61.61	1.19	75.48	46.78	67.08	1.00	85.07	48.62
CNN without Aug	91.9	0.23	93.63	90.49	90.02	0.34	91.68	88.98

Table 3 shows the performance results of the proposed model. The data augmentation techniques applied according to the performance results have negatively affected the performance metrics. There is no conclusive evidence that data augmentation techniques will give good results in every study. It did not give good results in this study either.

V. CONCLUSION

The increase in data sets in the field of object identification, segmentation and classification has increased the need for CNN methods. Technological investments are needed in CNN models, which is one of the deep learning-based models, in order to classify and make sense of the data in the increasing data stack. At this point, an effective classification has been made with layers that can obtain distinctive features from images in order to achieve the stated goals and objectives. Benchmark datasets such as CIFAR-10 allow comparison of performance measures of proposed models. In this study, it was determined whether data augmentation techniques had a positive or negative effect on the classification performance of the proposed model. Although it is stated in the literature that data augmentation processes have a positive effect on classification, the desired performance could not be achieved in this study.

REFERENCES

[1] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, “RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images,” *IEEE Access*, vol. 7, pp. 21420–21428, 2019, doi: 10.1109/ACCESS.2019.2896920.

[2] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, “Object detection with deep learning: A review,” *IEEE Trans. neural networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019.

[3] S. H. S. Basha, S. R. Dubey, V. Pulabaigari, and S.

Mukherjee, “Impact of fully connected layers on performance of convolutional neural networks for image classification,” *Neurocomputing*, vol. 378, pp. 112–119, 2020, doi: https://doi.org/10.1016/j.neucom.2019.10.008.

[4] 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.

[5] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” 2017.

[6] M. Choudhary, S. S. Chouhan, E. S. Pilli, and S. K. Vipparthi, “BerConvoNet: A deep learning framework for fake news classification,” *Appl. Soft Comput.*, vol. 110, p. 107614, 2021, doi: https://doi.org/10.1016/j.asoc.2021.107614.

[7] H. Çetiner, “Multi-Label Text Analysis with a CNN and LSTM Based Hybrid Deep Learning Model,” *J. Eng. Sci. Adiyaman Univ.*, vol. 9, no. 17, pp. 15–16.

[8] D. T. Çetin and S. Metlek, “Forecasting of Islamic Stock Index in Turkey with Deep Learning Using Index-Based Features,” *Acta Infologica*, vol. 5, no. 2, pp. 287–298, 2021.

[9] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE Trans. image Process.*, vol. 26, no. 7, pp. 3142–3155, 2017.

[10] T. Kobayashi, “Flip-invariant motion representation,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5628–5637.

[11] Q. Zheng, X. Tian, M. Yang, and H. Wang, “Differential Learning: A Powerful Tool for Interactive Content-Based Image Retrieval.,” *Eng. Lett.*, vol. 27, no. 1, 2019.

[12] Q. Zheng, M. Yang, X. Tian, N. Jiang, and D. Wang, “A Full Stage Data Augmentation Method in Deep Convolutional Neural Network for Natural Image Classification,” *Discret. Dyn. Nat. Soc.*, vol. 2020, p. 4706576, 2020, doi: 10.1155/2020/4706576.

[13] H. Inoue, “Data augmentation by pairing samples for images classification,” *arXiv Prepr. arXiv1801.02929*, 2018.

[14] J. Lemley, S. Bazrafkan, and P. Corcoran, “Smart augmentation learning an optimal data augmentation strategy,” *Ieee Access*, vol. 5, pp. 5858–5869, 2017.

[15] J. Wang and L. Perez, “The effectiveness of data augmentation in image classification using deep learning,” *Convolutional Neural Networks Vis. Recognit*, vol. 11, no. 2017, pp. 1–8, 2017.

[16] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification,” *Neurocomputing*, vol. 321, pp. 321–331, 2018.

[17] M. Zhu and Q. Chen, “Big data image classification based on distributed deep representation learning model,” *IEEE Access*, vol. 8, pp. 133890–133904, 2020.

[18] S. Tulasi Krishna and H. kumar Kalluri, “Deep Learning and Transfer Learning Approaches for Image Classification,” Jun. 2019.

[19] D. Masters and C. Luschi, “Revisiting small batch training for deep neural networks,” *arXiv Prepr. arXiv1804.07612*, 2018.

[20] A. Krizhevsky and G. Hinton, “Learning multiple layers of features from tiny images,” 2009.

[21] F. Giuste and J. Vizcarra, *CIFAR-10 Image Classification Using Feature Ensembles*. 2020.

[22] S.-J. Lee, T. Chen, L. Yu, and C.-H. Lai, “Image Classification Based on the Boost Convolutional Neural Network,” *IEEE Access*, vol. 6, pp. 12755–12768, 2018, doi: 10.1109/ACCESS.2018.2796722.

[23] M. Hussain, J. J. Bird, and D. R. Faria, “A study on cnn transfer learning for image classification,” in *Advances in Computational Intelligence Systems: Contributions Presented at the 18th UK Workshop on Computational Intelligence, September 5-7, 2018, Nottingham, UK*, 2019, pp. 191–202.

[24] N. Sharma, V. Jain, and A. Mishra, “An analysis of convolutional neural networks for image classification,” *Procedia Comput. Sci.*, vol. 132, pp. 377–384, 2018.