

Dog-Species Classification through CycleGAN and Standard Data Augmentation

Chan Park and Nammee Moon*

Abstract

In the image field, data augmentation refers to increasing the amount of data through an editing method such as rotating or cropping a photo. In this study, a generative adversarial network (GAN) image was created using CycleGAN, and various colors of dogs were reflected through data augmentation. In particular, dog data from the Stanford Dogs Dataset and Oxford-IIIT Pet Dataset were used, and 10 breeds of dog, corresponding to 300 images each, were selected. Subsequently, a GAN image was generated using CycleGAN, and four learning groups were established: 2,000 original photos (group I); 2,000 original photos + 1,000 GAN images (group II); 3,000 original photos (group III); and 3,000 original photos + 1,000 GAN images (group IV). The amount of data in each learning group was augmented using existing data augmentation methods such as rotating, cropping, erasing, and distorting. The augmented photo data were used to train the MobileNet_v3_Large, ResNet-152, InceptionResNet_v2, and NASNet_Large frameworks to evaluate the classification accuracy and loss. The top-3 accuracy for each deep neural network model was as follows: MobileNet_v3_Large of 86.4% (group I), 85.4% (group II), 90.4% (group III), and 89.2% (group IV); ResNet-152 of 82.4% (group I), 83.7% (group II), 84.7% (group III), and 84.9% (group IV); InceptionResNet_v2 of 90.7% (group I), 88.4% (group II), 93.3% (group III), and 93.1% (group IV); and NASNet_Large of 85% (group I), 88.1% (group II), 91.8% (group III), and 92% (group IV). The InceptionResNet_v2 model exhibited the highest image classification accuracy, and the NASNet_Large model exhibited the highest increase in the accuracy owing to data augmentation.

Keywords

CycleGAN, Data Augmentation, DNN, GAN, Image Classification

1. Introduction

In computer vision, image classification represents a basic and key task of visual recognition [1]. Various researchers have focused on image classification [2–6], for instance, to extract and classify regions of interest from satellite images [2]; to diagnose melanoma by learning various skin pictures in the medical domain [3]; and to classify animal species by learning animal images [4–6]. In particular, image classification aims to classify or predict the category of the test object by learning various photos through artificial intelligence. Several methods have been reported to enhance the accuracy of image classification. A promising method is to collect a large dataset. Notably, when classifying animal species, the collection of large datasets is challenging as it is difficult to record animals in an adequately static manner [7,8].

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received January 21, 2022; first revision March 21, 2022; accepted March 25, 2022.

*Corresponding Author: Nammee Moon (nammee.moon@gmail.com)

Dept. of Computer Science and Engineering, Hoseo University, Asan, Korea (chan.park941003@gmail.com, nammee.moon@gmail.com)

Consequently, researchers have attempted to dehaze photos using artificial intelligence [9].

However, even for dehazing, a dataset of hazy photos must be used to train artificial intelligence methods [10]. In this context, data augmentation can be performed to supplement insufficient datasets and enhance the accuracy of image classification. In the field of image processing, standard data augmentation is based on an image editing method. However, when standard data augmentation is applied to animal species classification, characteristics such as spots and mixed colors, which are extroverted characteristics in animals, cannot be reflected.

Thus, data augmentation methods based on generative adversarial networks (GANs) have been established [11,12]. GAN models can generate new images and have been proven to be effective in data generation tasks [12]. For instance, certain researchers generated images using GANs to obtain illustrations based on the text of children's books [13]. In data augmentation through GANs, a GAN image is created based on the original dataset and used as an image classification learning dataset.

Among other methods to increase the image classification accuracy, image classification has been performed using a deep neural network (DNN) model [14–16]. With progress in research on DNNs, various DNN models have been established [17–20].

In this study, 10 dog breeds are selected, and we attempt to supplement an inadequate dataset by reflecting the extroverted features of animals by using a GAN to increase the image classification accuracy. The performance of different DNN models in terms of image classification is evaluated.

The remaining paper is organized as follows: Section 2 introduces the related research on image classification through DNN models and data augmentation using a GAN. Section 3 describes the experiment setup and environment. Section 4 describes the datasets and process of selecting dog breeds, along with the GAN image creation through CycleGAN and data augmentation. Section 5 describes the composition of the learning group and performance evaluation process for each DNN model. Section 6 presents the concluding remarks.

2. Related Work

2.1 Image Classification via DNN Models

Various DNN models have been used to increase the accuracy of image classification [14–16]. Certain researchers attempted to learn malaria-infected and normal cells through ResNet and classify the infected cells [14]. Other researchers attempted to classify food categories by learning food photos through MobileNet or disease types by learning the leaves of diseased plants through NASNet [15,16]. With the emergence of image classification methods based on DNN models, the performance of these models has been extensively evaluated [21,22]. In 2018, the image classification and prediction performances of DNN models were evaluated using the ImageNet-1k validation set, which stores 14 million images. Moreover, the top-1 and top-5 accuracies for the models were specified against the number of operations in terms of giga-floating point operations per second (G-FLOPs) [22].

Notably, the top-N accuracy refers to the ratio of correct predictions when N classes are predicted through softmax in comparison with the true class. Fig. 1(a) shows that the top-1 accuracy for MobileNet_v2, ResNet-152, InceptionResNet_v2, and NASNet_Large are 71.81%, 78.25%, 80.28%, and 82.5%, respectively. Fig. 1(b) shows the top-5 accuracy for the models. Because the exact values were

not presented in the abovementioned study [22], approximate values for MobileNet_v2, ResNet-152, InceptionResNet_v2, and NASNet_Large are presented as 90.5%, 94.5%, 95.5%, and 96%, respectively.

Based on the abovementioned performance evaluation study, four DNN models are selected: MobileNet_v3_Large [17], ResNet-152 [18], InceptionResNet_v2 [19], and NASNet_Large [20]. MobileNet_v3_Large exhibits a 3.2% higher accuracy than that of MobileNet_v2 owing to the introduction of a nonlinear function, h-swish [17].

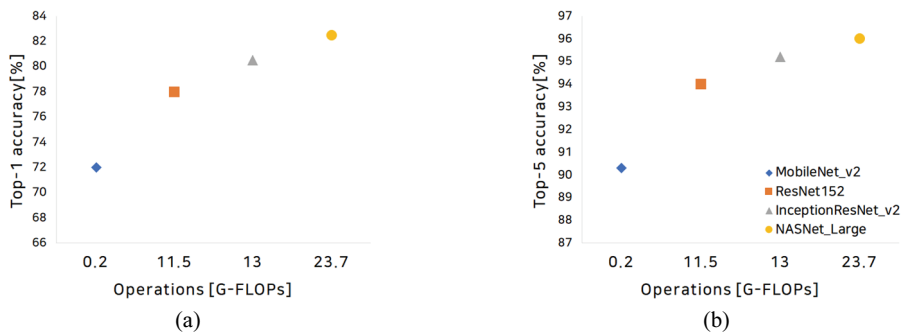


Fig. 1. Performance evaluation of various DNN models: (a) top-1 accuracy and (b) top-5 accuracy.

The ResNet model introduces residual learning, a “shortcut” concept, to solve the degradation problem that occurs as the learning layers are stacked and the problems of decreasing accuracy and increasing loss encountered in multilayered structures [18]. InceptionResNet_v2 exhibits a 3% higher accuracy than that of Inception_v3 owing to the addition of a residual block to the Inception_v3 model composed of Inception modules [19]. NASNet_Large can be applied to various datasets by designing blocks through a recurrent neural network and reinforcement learning [20]. In the abovementioned study [22], NASNet_Large exhibit the highest top-1 and top-5 accuracies. In this study, we compare the image classification performance of the DNN models considering these metrics.

2.2 Data Augmentation with GAN

In the field of image processing, data augmentation is an image editing method. For example, the amount of photo data can be increased through resizing, rotating, cropping, and random erasing methods [23,24]. Certain researchers attempted to classify images using data augmentation and evaluated the classification accuracy [25,26]. Based on random image cropping, the image classification test error rate of a dataset subjected to data augmentation was approximately 23% lower than that of the existing dataset without data augmentation [25]. Through random erasing, the test error rates could be decreased by approximately 9% [26] (Fig. 2).

However, standard data augmentation generates limited data [11]. For example, when augmenting an animal image using standard data augmentation techniques, characteristics such as spots and mixed colors, which are extroverted characteristics of animals, may not be reflected. Therefore, data augmentation methods based on GANs have been proposed [11,12]. Various GAN models such as CycleGAN, progressive growing GAN (PGGAN), unsupervised image-to-image translation (UNIT), and multimodal UNIT (MUNIT) have been proposed [27]. Recently, in the medical field, data augmentation was performed through GAN frameworks [28,29].

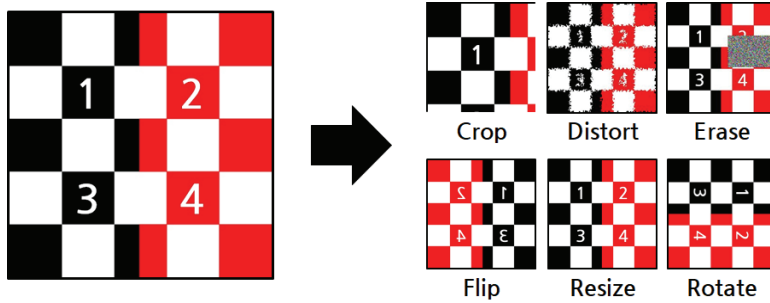


Fig. 2. Example of standard data augmentation.

Moreover, MUNIT and PGGAN were used to create and learn GAN brain images with tumors and detect tumors [28]. The medical image classification accuracy was attempted to be enhanced using CycleGAN and UNIT [29] (Fig. 3).



Fig. 3. Fake dog image creation using GAN.

Notably, vanilla GAN exhibits a limitation that the datasets are generated in pairs. In a paired dataset, the resolution and shape are identical. In the case of a GAN using a paired dataset, learning must be performed by mapping the input and output images. If the resolution and shape do not match, different images may be derived [30]. CycleGAN can unpair the datasets [31]. Therefore, CycleGAN is more suitable for dog photos, for which paired datasets cannot be generated owing to the various angles and appearances in the images. Fig. 4 shows an example of paired and unpaired datasets.

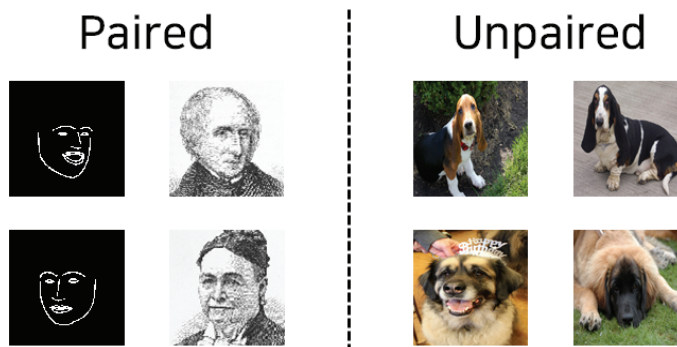


Fig. 4. Examples of paired and unpaired datasets.

In addition to the requirement of constructing paired datasets, vanilla GANs exhibit a key problem. A GAN consists of two networks: generator and discriminator. The generator transforms the data to pass the discriminator. In this scenario, the label or distribution of the converted data is biased toward a specific mode, known as mode collapse [32,33]. CycleGAN applies cycle-consistency and adversarial losses to solve the mode collapse problem. The cycle-consistency loss function is defined in Eq. (1).

$$L_{cyc}(G, F) = E_{x \sim p_{data}}(x)[\|F(G(x)) - x\|_1] + E_{y \sim p_{data}}(y)[\|G(F(y)) - y\|_1] \quad (1)$$

In Eq. (1), X and Y denote domains, x denotes the samples belonging to X , y denotes the samples belonging to Y , and G and F denote the mapping functions indicating translators or generators, respectively.

In this study, we perform data augmentation by using a GAN in an image classification task by fusing the dog photos generated through CycleGAN and standard data augmentation.

3. Experiment Overview

First, a dataset is constructed by selecting 10 dog breeds from the Stanford Dogs Dataset and Oxford-IIIT Pet Dataset. Subsequently, the dataset is separated and used for GAN image generation. Next, GAN images are generated through CycleGAN. Four learning groups are created using the original datasets and GAN image datasets. The number of images is increased through standard data augmentation. Finally, we train the four DNN models and evaluate the classification accuracy (Fig. 5). The experimental environment is summarized in Table 1.

4. CycleGAN-based Data Augmentation

4.1 Datasets

The Stanford Dogs Dataset and Oxford-IIIT Pet Dataset are used. Ten dog breeds (Basset Hound, Beagle, Boxer, English Setter, German Shorthaired, Keeshond, Leonberger, Miniature Pinscher, Pomeranian, and Pug), which are included in both datasets, are selected. Breeds with only a single color are not included. Moreover, pictures involving a person with a dog, pictures involving two dogs of different breeds, and pictures with effects such as sepia and grayscale are excluded. The resulting dataset has 3,000 photos, with 300 photos per dog breed. The contents of the dataset are presented in Table 2.

4.2 GAN image creation using CycleGAN

The learning configuration of CycleGAN is divided into TrainX and TrainY. GAN image generation is performed in $X \rightarrow Y$ style and $Y \rightarrow X$ style based on the two trained groups. In this study, single-color and mixed color images are configured as TrainX and TrainY, respectively, and 1,000 GAN images are generated, as shown in Fig. 6.

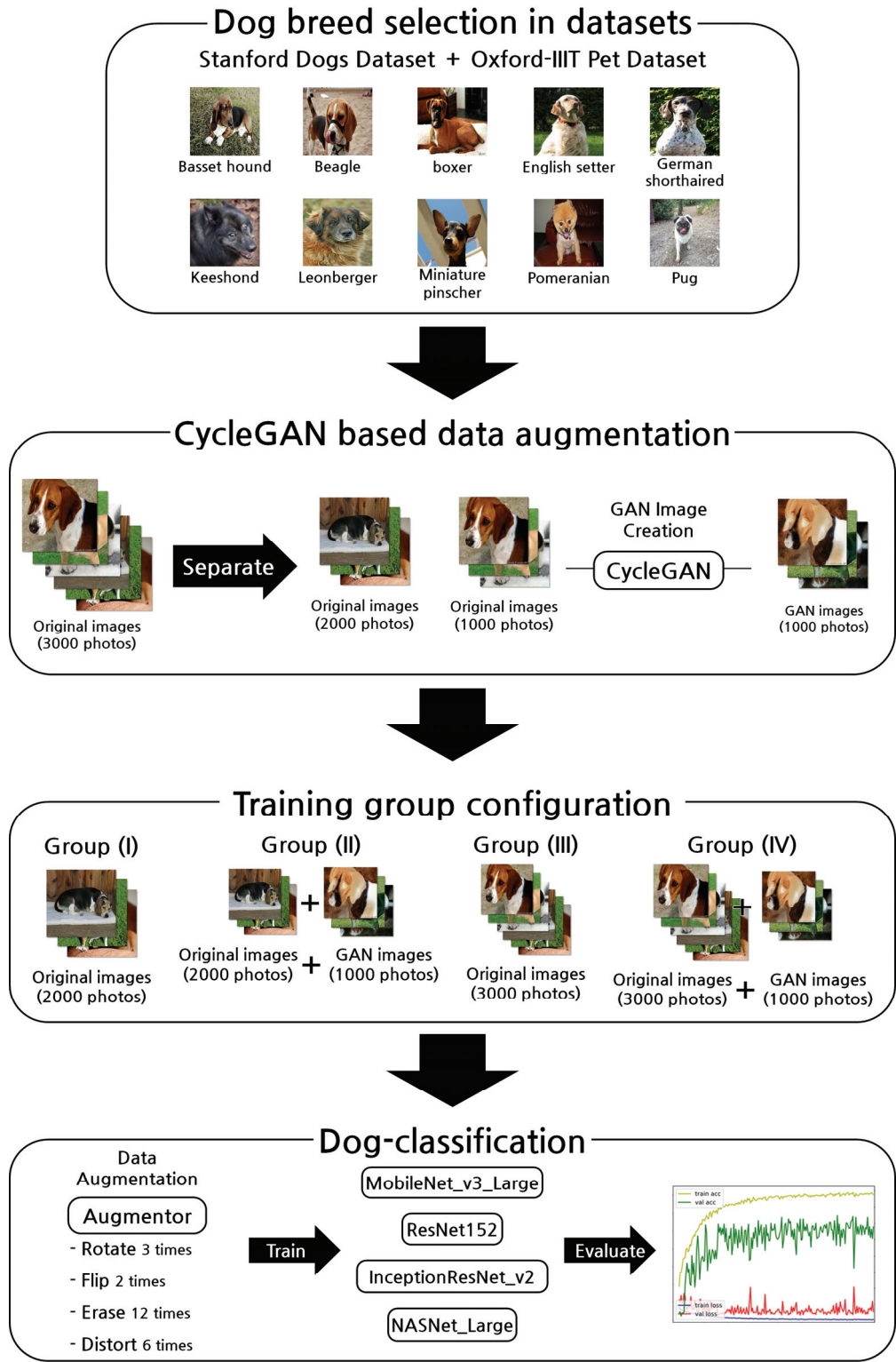


Fig. 5. Process flow of the experiment.

Table 1. Experimental environment

Type	Content
CPU	AMD Ryzen 9 5950X
GPU	Nvidia GeForce RTX 3090
RAM	128 GB
Python	3.9.6
TensorFlow	2.5.0

Table 2. Datasets used in the experiment

	Stanford Dogs Dataset	Oxford-IIIT Pet Dataset
Number of categories	120	37 (dogs and cats)
Number of images	20,580	7,400

4.3 Image data augmentation using Augmentor

Augmentor is a package that supports Python and Julia and provides image editing functions required for data augmentation [34]. In the dataset subjected to GAN image creation and grouping (described in a Section 5), the number of images is increased using Augmentor. The image is resized to 256×256 , rotated 90° three times, and reversed left and right and up and down. For each rotation and reversal, two erasing steps and one distortion are performed. Random image cropping is not performed. Because the size of the dog in each image is different, when random image cropping is applied, only the background is cropped or only specific areas such as eyes, feet, and the nose are cropped. Unfortunately, this process leads to a decreased learning accuracy. Through this process, the amount of image data can be augmented by 24 times. The results of data augmentation based on Augmentor are shown in Fig. 7.

**Fig. 6.** Generated images via CycleGAN.

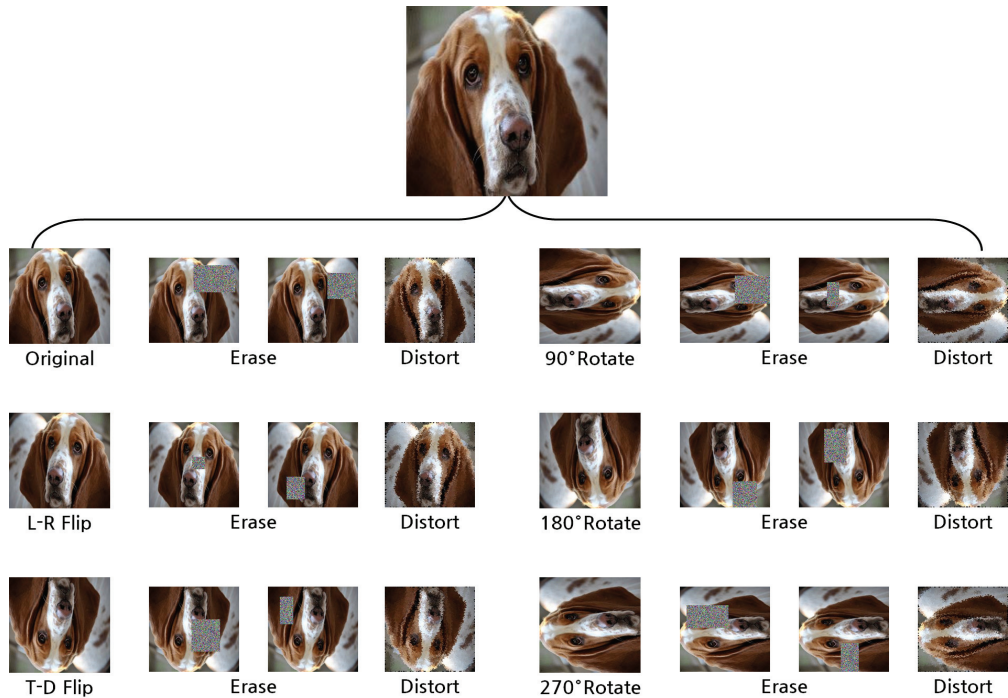


Fig. 7. Results of standard data augmentation using Augmentor.

5. Dog Classification

5.1 Training Group Configuration

One-third of the dataset of 10 dog breeds consisting of 3,000 pictures is selected. GAN images are created by implementing CycleGAN over the 1,000 photos, and six groups are formed. In the non-standard data augmentation datasets (NDA-I and NDA-II), standard data augmentation is not applied. Groups I and III are composed of 2,000 and 3,000 original images, respectively, and used to evaluate the basic performance of the DNN model. Groups II and IV contain 2,000 original images + 1,000 GAN images and 3,000 original images + 1,000 GAN images, respectively, and are used to compare the performance of the DNN model when the GAN images are added to the dataset. Subsequently, the amount of data is augmented through Augmentor. The composition of each learning group is shown in Table 3.

Table 3. Composition of training groups

	NDA-I	NDA-II	Group I	Group II	Group III	Group IV
Number of original images	2,000	2,000	2,000	2,000	3,000	3,000
Number of GAN images	-	1,000	-	1,000	-	1,000
Total number of images	2,000	3,000	2,000	3,000	3,000	4,000
After data augmentation	-	-	48,000	72,000	72,000	96,000

5.2 Comparison and Performance Evaluation for Each Model

For each group, training is conducted over MobileNet_v3_Large, ResNet-152, InceptionResNet_v2, and NASNet_Large. Using the ImageGenerator function of TensorFlow, the pixel size is preprocessed in the range of -1 to 1, and the data augmentation function of ImageGenerator is not applied. For each DNN model, ImageNet weights are set as zero. The batch size is 16; 100 epochs are implemented; and the learning is terminated when the validation accuracy does not increase within five epochs, through the early_stopping function. The graph of the learning result for each model is shown in Fig. 8.

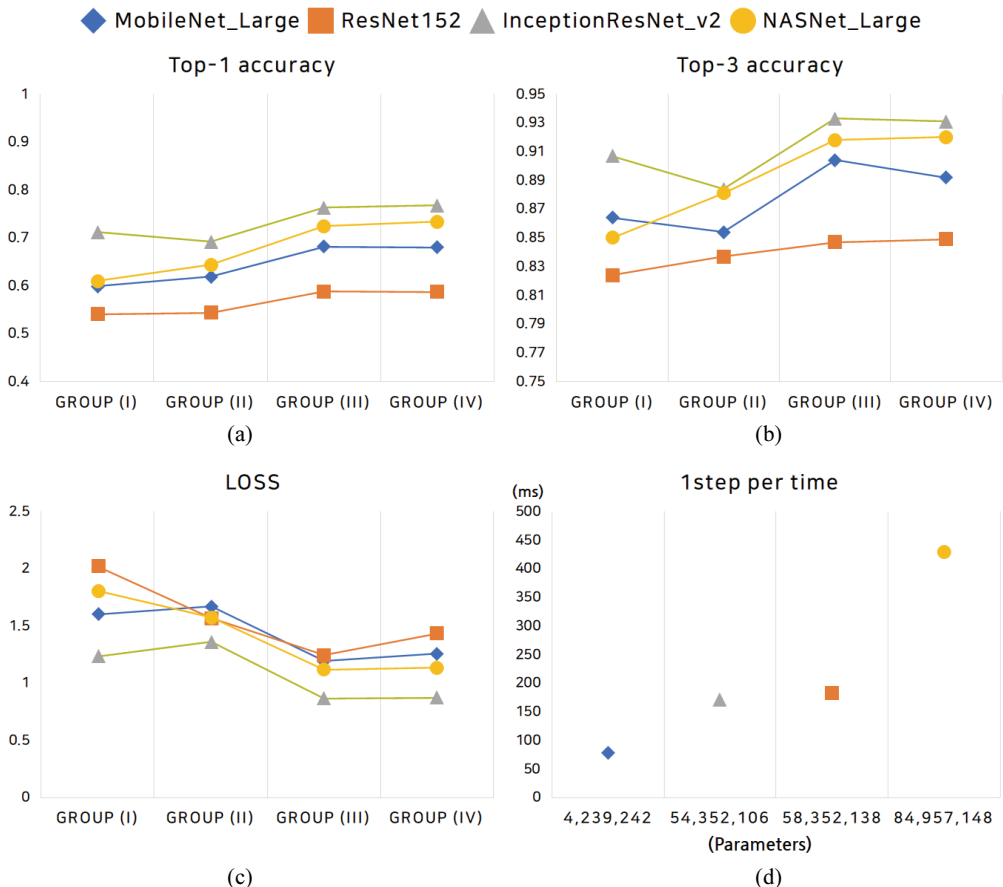


Fig. 8. Performance evaluation of different models. (a) and (b) show the top-1 and top-3 accuracies for each model, respectively. (c) shows the final loss when the highest prediction accuracy is achieved. (d) shows the number of parameters for each model and learning rate per epoch.

In Table 4, among the four DNN models, the InceptionResNet_v2 model exhibits a satisfactory classification performance. When data augmentation is performed for MobileNet_v3_Large, the top-3 accuracy decreases. ResNet-152 exhibits the lowest overall performance; however, the accuracy is enhanced when data augmentation is performed. InceptionResNet_v2 exhibits the highest classification accuracy. However, when data augmentation is applied, the accuracy is not enhanced. NASNet_Large exhibits the second-highest accuracy among the four models, and the accuracy is enhanced through data augmentation. The top-3 accuracies MobileNet_v3_Large and InceptionResNet_v2 decrease by 1.2%

and 1.36%, respectively and those of ResNet-152 and NASNet_Large increase by 0.91% and 1.92%, respectively. We compared the accuracy of NDA-I and NAD-II, which unperformed standard data augmentation and Groups I and II, which performed standard data augmentation. The top-3 accuracies NDA-I and Group I showed a difference in accuracy of about 1.88 times. NDA-II and Group II showed a difference in accuracy of about 1.68 times.

Table 4. Training result of different DNN models

	MobileNet_v3_Large	ResNet-152	InceptionResNet_v2	NASNet_Large
NDA-I				
Top-1 accuracy	0.051	0.094	0.108	0.105
Top-3 accuracy	0.243	0.316	0.318	0.335
Loss	2.491	2.382	2.806	2.352
NDA-II				
Top-1 accuracy	0.109	0.120	0.124	0.111
Top-3 accuracy	0.322	0.325	0.320	0.321
Loss	2.450	2.330	3.004	2.463
Group I				
Top-1 accuracy	0.599	0.541	0.712	0.61
Top-3 accuracy	0.864	0.824	0.907	0.850
Loss	1.601	2.023	1.235	1.806
Group II				
Top-1 accuracy	0.619	0.544	0.692	0.644
Top-3 accuracy	0.854	0.837	0.884	0.881
Loss	1.671	1.566	1.359	1.573
Group III				
Top-1 accuracy	0.682	0.588	0.763	0.724
Top-3 accuracy	0.904	0.847	0.933	0.918
Loss	1.194	1.245	0.868	1.120
Group IV				
Top-1 accuracy	0.680	0.587	0.768	0.734
Top-3 accuracy	0.892	0.849	0.931	0.92
Loss	1.257	1.435	0.872	1.136

The bold font indicates the best performance model in each test.

6. Conclusion

We compare the image classification accuracy of four DNN models after generating GAN images through CycleGAN, implementing the standard data augmentation process, and performing training on the DNN models. Most of the existing image classification studies are focused on classifying images pertaining to a large category. In contrast, this study is focused on classifying images of dogs pertaining to a detailed category. GAN images are generated using CycleGAN to reflect dog characteristics such as spots and mixed color, which cannot be reflected through standard data augmentation techniques. Subsequently, standard data augmentation is performed to increase the amount of image data. The augmented data are applied for DNN model training, and the image classification accuracy is compared.

When data augmentation based on GAN is performed, the accuracy for certain models increases. However, the accuracy enhancement based on GAN image data is not adequate for these data to replace the original data, and the accuracy of several DNN models decreases. To address this problem and increase the accuracy of image classification, it is necessary to use other GAN models or enhance the standard data augmentation process.

Acknowledgement

This research is supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C2011966).

References

- [1] A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," in *Proceedings of 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, Jakarta, Indonesia, 2017, pp. 26-31.
- [2] D. Arezki and H. Fizazi, "Alsat-2B/Sentinel-2 imagery classification using the hybrid pigeon inspired optimization algorithm," *Journal of Information Processing Systems*, vol. 17, no. 4, pp. 690-706, 2021.
- [3] T. Akram, H. M. J. Lodhi, S. R. Naqvi, S. Naeem, M. Alhaisoni, M. Ali, S. A. Haider, and N. N. Qadri, "A multilevel features selection framework for skin lesion classification," *Human-centric Computing and Information Sciences*, vol. 10, article no. 12, 2020. <https://doi.org/10.1186/s13673-020-00216-y>
- [4] H. Liu, "Animal image classification recognition based on transfer learning," *International Core Journal of Engineering*, vol. 7, no. 8, pp. 135-140, 2021.
- [5] A. Vecvanags, K. Aktas, I. Pavlovs, E. Avots, J. Filipovs, A. Brauns, G. Done, D. Jakovels, and G. Anbarjafari, "Ungulate detection and species classification from camera trap images using RetinaNet and Faster R-CNN," *Entropy*, vol. 24, no. 3, article no. 353, 2022. <https://doi.org/10.3390/e24030353>
- [6] M. Hu and F. You, "Research on animal image classification based on transfer learning," in *Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen, China, 2020, pp. 756-761.
- [7] S. Schneider, S. Greenberg, G. W. Taylor, and S. C. Kremer, "Three critical factors affecting automated image species recognition performance for camera traps," *Ecology and Evolution*, vol. 10, no. 7, pp. 3503-3517, 2020.
- [8] Y. Guo, T. A. Rothfus, A. S. Ashour, L. Si, C. Du, and T. F. Ting, "Varied channels region proposal and classification network for wildlife image classification under complex environment," *IET Image Processing*, vol. 14, no. 4, pp. 585-591, 2020.
- [9] L. Zhao, Y. Zhang, and Y. Cui, "A multi-scale U-shaped attention network-based GAN method for single image dehazing," *Human-centric Computing and Information Sciences*, vol. 11, article no. 38, 2021. <https://doi.org/10.22967/HICIS.2021.11.038>
- [10] C. O. Ancuti, C. Ancuti, and R. Timofte, "NH-HAZE: an image dehazing benchmark with non-homogeneous hazy and haze-free images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, WA, 2020, pp. 1798-1805.
- [11] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," 2017 [Online]. Available: <https://arxiv.org/abs/1711.04340>.

- [12] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017 [Online]. Available: <https://arxiv.org/abs/1712.04621>.
- [13] J. Cho and N. Moon, "Design of image generation system for DCGAN-based kids' book text," *Journal of Information Processing Systems*, vol. 16, no. 6, pp. 1437-1446, 2020.
- [14] A. S. B. Reddy and D. S. Juliet, "Transfer learning with ResNet-50 for malaria cell-image classification," in *Proceedings of 2019 International Conference on Communication and Signal Processing (ICCCSP)*, Chennai, India, 2019, pp. 945-949.
- [15] S. Phiphatphaisit and O. Surinta, "Food image classification with improved MobileNet architecture and data augmentation," in *Proceedings of the 3rd International Conference on Information Science and Systems*, Cambridge, UK, 2020, pp. 51-56.
- [16] A. Adedoja, P. A. Owolawi, and T. Mapayi, "Deep learning based on NASNet for plant disease recognition using leave images," in *Proceedings of 2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*, Winterton, South Africa, 2019, pp. 1-5.
- [17] A. Howard, M. Sandler, G. Chu, L. C. Chen, B. Chen, M. Tan, M., ... & Adam, H. (2019). "Searching for MobileNetV3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, South Korea, 2019, pp. 1314-1324.
- [18] 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*, Las Vegas, NV, 2016, pp. 770-778.
- [19] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, CA, 2017, pp. 4278, 4284.
- [20] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 2018, pp. 8697-8710.
- [21] W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li, "Development of convolutional neural network and its application in image classification: a survey," *Optical Engineering*, vol. 58, no. 4, article no. 040901, 2019. <https://doi.org/10.1117/1.OE.58.4.040901>
- [22] S. Bianco, R. Cadene, L. Celona, and P. Napolitano, "Benchmark analysis of representative deep neural network architectures," *IEEE Access*, vol. 6, pp. 64270-64277, 2018.
- [23] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," *Artificial Intelligence Review*, vol. 55, pp. 2351-2377, 2022.
- [24] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, article no. 60, 2019. <https://doi.org/10.1186/s40537-019-0197-0>
- [25] R. Takahashi, T. Matsubara, and K. Uehara, "Data augmentation using random image cropping and patching for deep CNNs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 9, pp. 2917-2931, 2019.
- [26] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," in *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, NY, 2020, pp. 13001-13008.
- [27] A. Jabbar, X. Li, and B. Omar, "A survey on generative adversarial networks: variants, applications, and training," *ACM Computing Surveys*, vol. 54, no. 8, article no. 157, 2022. <https://doi.org/10.1145/3463475>
- [28] C. Han, L. Rundo, R. Araki, Y. Nagano, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Combining noise-to-image and image-to-image GANs: Brain MR image augmentation for tumor detection," *IEEE Access*, vol. 7, pp. 156966-156977, 2019.
- [29] D. H. Lee, Y. Li, and B. S. Shin, "Generalization of intensity distribution of medical images using GANs," *Human-centric Computing and Information Sciences*, vol. 10, article no. 17, 2020. <https://doi.org/10.1186/s13673-020-00220-2>

- [30] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 2017, pp. 5967-5976.
- [31] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017, pp. 2242-2251.
- [32] W. Li, L. Fan, Z. Wang, C. Ma, and X. Cui, "Tackling mode collapse in multi-generator GANs with orthogonal vectors," *Pattern Recognition*, vol. 110, article no. 107646, 2021. <https://doi.org/10.1016/j.patcog.2020.107646>
- [33] H. De Meulemeester, J. Schreurs, M. Fanuel, B. De Moor, and J. A. Suykens, "The Bures metric for generative adversarial networks," in *Machine Learning and Knowledge Discovery in Databases: Research Track*. Cham, Switzerland: Springer, 2021, pp. 52-66.
- [34] M. D. Bloice, C. Stocker, and A. Holzinger, "Augmentor: an image augmentation library for machine learning," 2017 [Online]. Available: <https://arxiv.org/abs/1708.04680>.



Chan Park <https://orcid.org/0000-0002-9521-8939>

He received his B.S. degree from the School of Computer Science and Engineering at Hoseo University. He is currently pursuing an M.S. degree at the Department of Computer Science and Engineering at Hoseo University, Korea. His research interests include image processing, artificial intelligence, machine learning, and big data processing and analysis.



Nammee Moon <https://orcid.org/0000-0003-2229-4217>

She received her B.S., M.S., and Ph.D. degrees from the School of Computer Science and Engineering at Ewha Womans University in 1985, 1987, and 1998, respectively. She served as an assistant professor at Ewha Womans University from 1999 to 2003 and as a professor of Digital Media, Graduate School of Seoul Venture Information, from 2003 to 2008. Since 2008, she has been a professor at the Department of Computer Science and Engineering at Hoseo University. Her research interests include social learning, HCI, user-centric data, artificial intelligence, and big data processing and analysis.