

전이 학습과 데이터 증강을 이용한 너구리와 리쿤 분류

박동민¹, 조영석², 염석원^{1*}

¹대구대학교 IT융합공학과, ²대구대학교 생물교육학과

Classification of Raccoon dog and Raccoon with Transfer Learning and Data Augmentation

Dong-Min Park¹, Yeong-Seok Jo², Seokwon Yeom^{1*}

¹Department of IT Convergence Engineering, Daegu University

²Department of Biology Education, Daegu University

요약 최근 인간의 활동 범위가 증가함에 따라 외래종의 유입이 잦아지고 있고 환경에 적응하지 못해 유기된 외래종 중 2020년부터 유해 지정 동물로 지정된 리쿤이 문제가 되고 있다. 리쿤은 국내 토종 너구리와 크기나 생김새가 유사하여 일반적으로 포획하는데 있어서 구분이 필요하다. 이를 해결하기 위해서 이미지 분류에 특화된 CNN 딥러닝 모델인 VGG19, ResNet152V2, InceptionV3, InceptionResNet, NASNet을 사용한다. 학습에 사용할 파라미터는 많은 양의 데이터인 ImageNet으로 미리 학습된 파라미터를 전이 학습하여 이용한다. 너구리와 리쿤 데이터셋에서 동물의 외형적인 특징으로 분류하기 위해서 이미지를 회색조로 변환한 후 밝기를 정규화하였으며, 조정된 데이터셋에 충분한 학습을 위한 데이터를 만들기 위해 좌우 반전, 회전, 확대/축소, 이동을 이용하여 증강 기법을 적용하였다. 증강하지 않은 데이터셋은 FCL을 1층으로, 증강된 데이터셋은 4층으로 구성하여 진행하였다. 여러 가지 증강된 데이터셋의 정확도를 비교한 결과, 증강을 많이 할수록 성능이 증가함을 확인하였다.

• 주제어 : 딥러닝, 머신러닝, 전이 학습, 데이터 증강, 동물 분류

Abstract In recent years, as the range of human activities has increased, the introduction of alien species has become frequent. Among them, raccoons have been designated as harmful animals since 2020. Raccoons are similar in size and shape to raccoon dogs, so they generally need to be distinguished in capturing them. To solve this problem, we use VGG19, ResNet152V2, InceptionV3, InceptionResNet and NASNet, which are CNN deep learning models specialized for image classification. The parameters to be used for learning are pre-trained with a large amount of data, ImageNet. In order to classify the raccoon and raccoon dog datasets as outward features of animals, the image was converted to grayscale and brightness was normalized. Augmentation methods were applied using left and right inversion, rotation, scaling, and shift to create sufficient data for transfer learning. The FCL consists of 1 layer for the non-augmented dataset while 4 layers for the augmented dataset. Comparing the accuracy of various augmented datasets, the performance increased as more augmentation methods were applied.

• Key Words : Deep learning, Machine Learning, Transfer Learning, Data Augmentation, Animal Classification

Received 24 November 2022, Revised 28 February 2023, Accepted 5 March 2023

* Corresponding Author Seokwon Yeom, Department of IT Convergence Engineering, Daegu University, 201, Daegudae-ro, Jillyang-eup, Gyeongsan-si, Gyeongsangbuk-do, Republic of Korea. E-mail: yeom@daegu.ac.kr

I. 서론

최근 무역, 교통이 발전하고 전 지구적으로 세계화가 되면서 인간의 활동 범위의 증가로 외래종 침입에 대한 위협이 증가하고 있다[1]. 무역과 교통이 발전하면서 외래종을 유입한 사례로 황소개구리, 큰입배스, 뉴트리아, 붉은귀거북 등이 있다. 이러한 침입종들은 대체로 번식력이 높고 포식성이 강하여 국내 토착종과 같은 환경에 살 경우 경쟁, 질병 등의 문제를 일으켜 국내 생물의 다양성에 문제를 줄 뿐만 아니라 생태계 전반에 위협을 가한다[2]. 이에 생물다양성협약에서는 외래종의 침입을 생물다양성의 가장 큰 위협이라고 지목하고, 외래종에 대한 관리 대책이 필요함을 알리고 있다[3].

본 연구는 토착종의 생태계를 보호하기 위해서 외래종 확산 방지를 목표로 한다. 여러 종의 동물 중에서 한국 토착종인 너구리와 유사한 외래종 라쿤을 구분하여 식별하는 것을 목표로 한다. 두 동물은 생김새가 유사하지만, 라쿤은 아메리카 너구리 과에 속하는 다른 종이며, 너구리와 달리 꼬리에 링 무늬가 있다. 라쿤은 직·간접적으로 너구리와 경쟁 상대이며, 생태계에 부정적인 영향을 끼치고[4] 오래된 목조건물과 농작물에 피해를 준다. 한국에서는 2019년 진행된 라쿤 실태조사를 통하여, 2020년 환경부에서 라쿤을 생태계 위해성 2급 동물로 지정하여 관리를 시작하였다[5].

유사한 생김새의 두 동물을 구분하기 위해서 여러 딥러닝 네트워크 중에서 이미지 인식에 활용되고 있는 CNN을 이용한다[6]. 많은 CNN 모델 중에서 본 논문은 이미지 분류에 높은 성능을 보인 ResNet, Inception, VGG 등과 같은 CNN 모델을 사용한다[7-11]. 각 모델은 많은 양의 데이터를 가진 ImageNet을 사용하여 미리 학습된 파라미터를 가진 모델로 특징을 추출하는 전이 학습을 이용한다. 많은 양의 다양한 데이터를 이용하여 학습된 파라미터는 각 이미지의 경계선이나 색과 같은 보편적인 특징을 학습했기 때문에 새로운 이미지가 들어와도 빠르게 학습하여 결과를 낼 수 있다 [12].

본 연구에서 사용한 동물의 원본 데이터의 수는 자체적으로 모델에 학습하여 진행하기에 부족하여 데이터 증강을 사용한다. 데이터 증강을 통하여 늘어난 이미지들은 모델이 동물에 대한 특징을 학습하는 데 도움을 주어 성능에 영향을 줄 뿐만 아니라 적은 데이터

에 의한 과소 적합 문제도 해결할 수 있다[13].

본 연구는 동물의 외형적인 특징으로 분류하기 위해서 두 동물의 데이터셋의 환경적인 요소의 의존도를 줄이고 실험을 통해서 제한된 데이터의 증강을 위하여 적합한 증강 기법을 찾는다. 증강된 데이터셋에서 각 모델의 학습 데이터와 검증 데이터의 정확도를 비교한 결과, 증강이 많이 될수록 성능이 증가하였다.

II장에서는 여러 CNN 모델의 특징과 전이 학습과 데이터 증강에 관해서 기술하고 III장에서는 실험을 진행하고 IV장에서는 결론을 제시한다.

II. 본론

2.1. CNN

CNN은 직접 특징을 추출할 필요 없이 데이터로부터 모델이 직접 학습하는 딥러닝 모델 구조이다. 수십, 수백 개의 계층을 통과하며 각 계층에서 특징을 검출한다. 특징을 추출할 영상에서는 서로 다른 크기의 필터가 적용되고, 필터를 통하여 매우 간단한 특징부터 물체 고유의 특징까지 추출한다. 컨볼루션 과정 후, 효과적으로 물체의 특징을 추출하기 위하여 활성화 함수를 사용하며, 풀링을 통하여 다음 신경망 단계에서 학습할 파라미터 수를 줄인다. 계층의 수가 증가할수록 과적합 해결과 연산을 효율적으로 수행하는 방법이 추가되어 발전되고 있다[6].

본 연구에서는 다양한 CNN 모델 중에서 최근 각광 받는 ResNet, VggNet, Inception, NASNet의 결합 또는 변형된 모델인 VGG19, ResNet152V2, InceptionV3, InceptionResNetV2, NASNetLarge를 채용한다.

VGG 모델은 간단한 구조로 이루어져 있고, 7×7이나 5×5 같은 큰 컨볼루션 연산을 3×3 컨볼루션을 여러 층으로 쌓은 형태를 이용하여 연산 속도와 정확도를 향상시켰다[7]. VGG 모델의 입력 이미지의 크기는 224×224×3화소이다. ResNet과 같은 크기로 학습을 진행하나, 간단한 구조로 이루어져 있어서 ResNet 모델보다 빨리 결과를 도출한다. 본 실험에서 사용하는 VGG19 모델은 그림 1과 같이 간단한 구조로 컨볼루션 연산을 총 19번을 진행하게 된다.

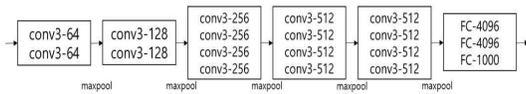


Fig. 1. VGG19 Model.

ResNet 모델은 그림 2와 같이 입력값을 출력값에 연산해주는 Residual block 형태를 사용하는 점과 컨볼루션 층에서 1×1 , 3×3 , 1×1 3개의 층을 사용하여 특징 맵의 차원을 줄여 실행 시간을 단축시켰으며, 152개의 층까지 모델이 구축되면서 오류율을 줄여서 성능을 향상시켰다[7]. ResNet 모델의 입력 이미지의 크기는 $224 \times 224 \times 3$ 화소이다.

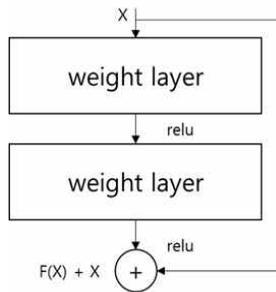


Fig. 2. Residual Block.

Inception 모델은 컨볼루션 망에서 fully connected로 연산할 경우 overfitting 현상과 계산량이 많아지는 점을 보완하기 위해 그림 3과 같이 이미지 특징의 크기에 따라 다른 크기의 컨볼루션 망으로 연산하는 sparsely connected를 사용하며, 한 단계 발전시켜 VGG에서 설명한 것과 같은 방식으로 큰 컨볼루션을 여러 차례의 3×3 연산으로 대체하여 진행한다[9]. Inception 모델의 입력 이미지의 크기는 $299 \times 299 \times 3$ 화소이다.

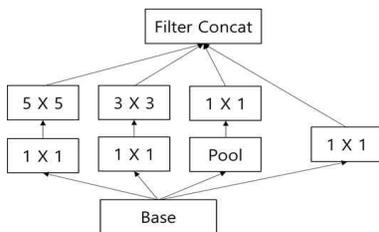


Fig. 3. InceptionV3 Convolution.

InceptionResNet 모델은 컨볼루션 망이 Inception과 유사하지만, 컨볼루션 망에서 1×1 을 추가하여 ResNet

컨볼루션 망과 같은 구조로 만들어 사용하여 기존의 Inception보다 적은 epoch 수에도 결과를 도출한다[10]. InceptionResNet 모델의 입력 이미지의 크기는 $299 \times 299 \times 3$ 화소이다. 그림 4는 본 실험에 사용한 InceptionResNetV2의 컨볼루션 망을 보여준다.

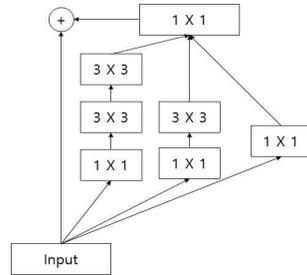


Fig. 4. InceptionResNetV2 Convolution.

NASNetLarge 모델은 그림 5와 같이 Reduction cell과 Normal cell의 반복으로 학습한다. Reduction cell은 특징 맵을 $1/2$ 로 줄이고, Normal cell은 입력된 크기와 같은 크기로 특징 맵을 출력하며 Cell 안의 컨볼루션 과정은 ReLU 함수가 적용된다[11]. NASNet 모델의 입력 이미지의 크기는 $331 \times 331 \times 3$ 화소이다.

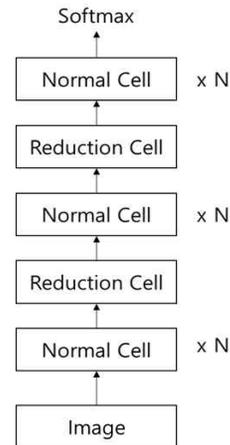


Fig. 5. NASNet Model.

2.2. 전이 학습

전이 학습은 140만 개의 이미지와 1,000개의 라벨을 가지고 있는 이미지넷[14]과 같은 큰 데이터셋으로 모델을 학습한다. 이미지넷 데이터를 이용하여 학습된

가중치는 CNN 모델에서 컨볼루션 망 일부 혹은 전체 까지의 학습에 사용된다. 학습할 데이터가 많고 사전 학습된 모델과 유사성이 작으면 모델 전체를 재학습시키지만 데이터셋이 작고 유사성이 크면 완전연결층(FCL: Fully Connected Layer)만 학습시킨다. 이미지넷으로 미리 학습된 모델의 정확도가 좋을수록 전이 학습의 성능이 좋아진다. 그림 6과 같이 본 논문은 적은 양의 데이터셋을 이용하기 때문에 모델 내에서 충분한 학습을 위해서 많은 데이터셋으로 미리 학습된 이미지넷의 가중치가 필요하여 전이 학습을 사용하였다. 그림 6에서 빨간색과 초록색 부분은 재사용하는 컨볼루션 망과 FCL을 나타내고 파란색은 추가한 FCL이며, 점선 박스는 전체 FCL이다.

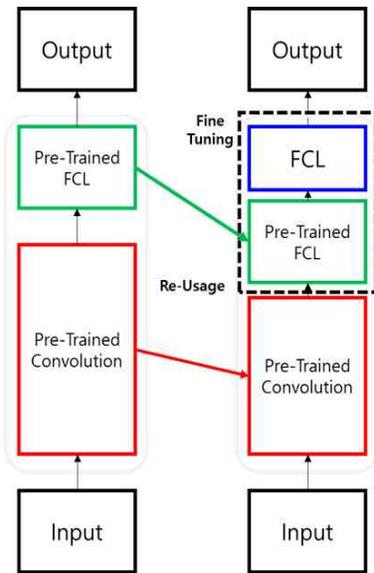


Fig. 6. Transfer Learning.

2.3. 데이터 증강

많은 CNN 모델은 이미지를 사용하는 컴퓨터 비전 업무에 탁월한 효과를 내고 있다. CNN 모델이 과적합, 과소 적합의 문제없이 학습을 하기 위해서는 많은 양의 이미지 데이터가 필요하다. 하지만 특정한 분야에서 활용할 수 있는 이미지 데이터는 한정적일 수 있다. 이런 문제를 해결하기 위해서 데이터 증강 기법을 사용한다. 증강 기법에는 형태 변환, 색 변환, 이미지 혼합, 커널 필터 등 다양한 방법이 있다[12].

본 논문에서는 형태 변환을 이용하여 데이터를 증강한 뒤 학습을 진행하였다. 다양한 형태 변환 기법을 적용하여 증강 기법을 통하여 성능이 향상하는지 확인해본다.

III. 실험 방법 및 결과

두 동물의 형태적 특징을 활용하기 위해서 컬러 이미지를 640×480화소로 크기 변환 후 회색조로 변환시킨다. 변환시킨 후 식 1을 사용하여 회색조 이미지들의 명도를 정규화(stretching)한다. 식 1에서 t_{max} 는 출력 영상의 최대 명도 값으로 255이며, o_{max} 는 원본 이미지 최대 명도 값이다. v 는 원본 이미지의 각 화소의 명도 값이며 v_t 는 정규화된 명도를 나타낸다.

$$v_t = \frac{t_{max}}{o_{max}} \times v \quad (1)$$

다음으로는 형태 변환 증강 기법인 좌우 반전, 30°, 45°, -30°, -45° 회전, 확대(1.2배)와 축소(0.8배), 이동(가로방향 ±128, ±192, 세로방향 ±96, ±144)을 이용하여 데이터셋을 증강하였다. 그러므로 하나의 영상은 270(=2×5×3×9)개로 증강되었다.

그림 7은 본 실험에서 이미지를 딥러닝 하는 순서도이다. 그림 7(a)는 증강되지 않은 데이터셋으로 실험하여 1층으로 구성된 FCL, 그림 7(b)는 다양한 형태 변환 증강 기법이 적용된 데이터셋으로 실험을 진행하여 4층으로 구성된 FCL로 구성되어있다. 은닉층의 활성화 함수는 ReLU 함수이고 출력층에서 두 개의 클래스를 분류하기 위하여 Softmax 함수를 사용하였다. 손실 함수는 Categorical Cross-Entropy, 옵티마이저는 Adam을 사용하여 학습을 진행하였다.

그림 8(a)는 본 실험에서 사용한 너구리 이미지이고, 그림 8(b)는 라쿤 이미지이다. 그림 9는 너구리와 라쿤의 이미지를 640×480 해상도로 크기 변환 후 회색조 영상으로 나타낸 이미지이다. 그리고 두 데이터셋의 명도 차이를 줄이기 위해서 식 1을 적용한 결과는 그림 10에서 볼 수 있다.

본 실험의 딥러닝 학습 환경은 RAM: 64GB, CPU: Intel i9, GPU: NVIDIA RTX 3090인 데스크탑을 사용하였다. 실험에 사용한 증강하기 전의 너구리와 라쿤의

이미지 수는 각각 100개이며 학습 이미지와 검증 이미지의 비율을 7:3으로 진행하였다. 표 1은 증강기법을 적용하지 않은 실험으로 batch size 15, epoch 1000으로 설정하여 나온 결과이다. 이 표는 모델별 학습 데이터, 검증 데이터의 정확도와 재현율을 보여준다. 증강하지 않은 데이터셋은 이미지넷으로 학습한 사전 학습모델의 성능 순서와 유사함을 보였다[15].



(a) Raccoon Dog (b) Raccoon

Fig. 10. Normalized Image.

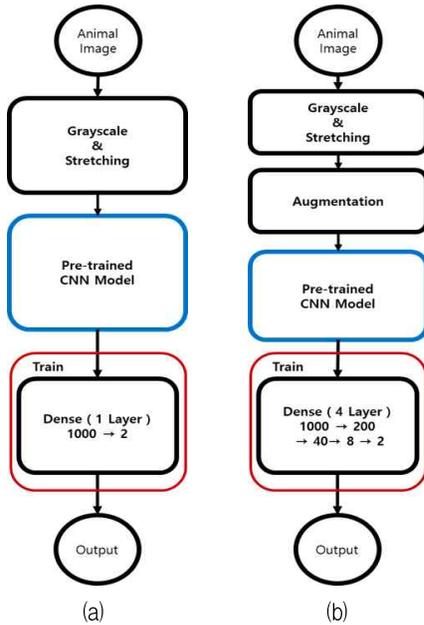


Fig. 7. Block Diagram of Deep Learning,
 (a) Block Diagram of Small Dataset Experiment,
 (b) Block Diagram of Large Dataset Experiment.



(a) Raccoon Dog (b) Raccoon

Fig. 8. Animal Image.



(a) Raccoon Dog (b) Raccoon

Fig. 9. Grayscaled Image.

그림 11은 명도를 정규화한 이미지 데이터셋에 형태 변환 증강 기법을 적용하여 진행한 실험 결과로 훈련 데이터와 검증 데이터의 정확도를 보여준다. 데이터셋 A는 좌우 반전, 회전, 이동, 확대/축소가 적용되었다. 데이터셋 B는 좌우 반전, 회전, 확대/축소가 적용되었다. 데이터셋 C는 이동, 확대/축소가 적용되었다. 각 A, B, C의 훈련 데이터와 검증 데이터는 전체 데이터 수 54,000, 6,000, 5,400개를 7:3으로 나누고 batch size는 100, epoch는 75로 실험을 진행하였다. 그림 11에서 검은색은 VGG19, 빨간색은 ResNet152V2, 노란색은 InceptionV3, 초록색은 InceptionResNetV2, 파란색은 NASNetLarge를 표시하였다.

각 실험에서 학습 데이터셋에 대한 결과는 모든 경우에서 97.7~100%로 높게 나왔다. 검증 데이터셋에 대한 결과는 데이터셋 A의 경우, InceptionV3, VGG19, NASNetLarge 모델의 순서로 정확도가 높았다. 데이터셋 B의 경우, NASNetLarge, InceptionV3, InceptionResNetV2 순서로 높았다. 데이터셋 C의 경우, InceptionV3, VGG19, NASNetLarge 순서로 정확도가 높았다. 회전 증강의 점유율이 큰 데이터셋 B의 경우 VGG19의 성능이 낮았는데 이는 VGG19의 입력 이미지 크기가 다른 사전 학습 모델보다 작기 때문인 것으로 보인다. 가장 많은 증강 기법을 사용한 데이터셋 A의 결과가 가장 좋았으며, 데이터셋 A와 C에서 가장 좋은 성능을 보인 InceptionV3 모델의 정확도는 각각 97.4%와 96.8%이었고 데이터셋 B에서 NASNetLarge의 정확도는 96.2%이었다.

증강 기법이 적용되기 전에는 학습과 검증 데이터의 최대 정확도가 각각 98.57%와 91.67%이었고 증강 기법을 적용하고 FCL층을 늘린 결과, 학습 데이터에 대해 97.7~100%, 검증 데이터에 대하여 96.2~97.4%의 정확도를 보였다.

Table 1. Normalized Image without Augmentation Results.

	Training Accuracy	Validation Accuracy	Raccoon Dog		Raccoon		Time rate
			Recall		Recall		
			Train	Valid	Train	Valid	
VGG19	87.14	80.00	92.86	86.67	81.43	73.33	0.523
ResNet152V2	95.71	91.67	100	96.67	91.43	86.67	0.528
InceptionV3	98.57	86.67	100	93.33	97.14	80.00	0.655
InceptionResNetV2	91.43	91.67	94.29	93.33	90.00	88.57	0.678
NASNetLarge	98.57	85.00	97.14	83.33	100	86.67	1.0

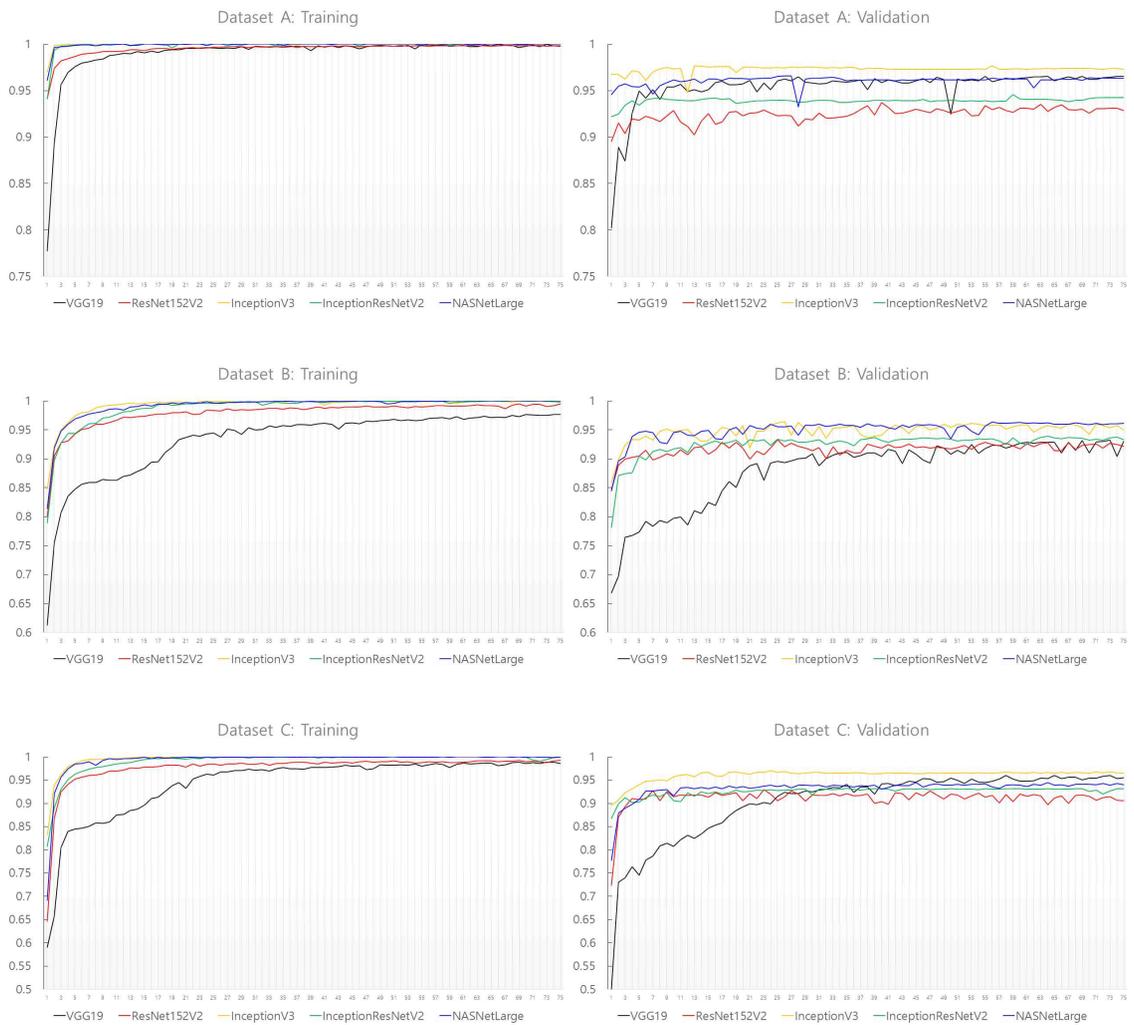


Fig. 11. Accuracy of Each Model.

IV. 결론

본 논문에서는 너구리와 라쿤의 형태적 특징을 이용하여 분류하는 연구를 수행하였다. 형태적 특징을 이용하기 위해서 컬러 이미지를 회색조로 변환하고 명도 정규화를 하였다. 증강 기법이 적용될 경우, 성능에 얼마나 영향을 미치는지 학습 데이터와 검증 데이터의 정확도를 통하여 확인하였고, 추가적으로 증강된 양에 따라 성능에 얼마나 작용하는지 확인해보았다.

이를 활용하여 국내 토착종과 유사한 생김새를 가진 외래종 생물들을 식별하여 포획하는 데 사용할 수 있으며 국내 생물의 생태계를 보호할 수 있다. 앞으로 본 연구의 데이터와 같이 식별이 필요한 다양한 외래종 동물 데이터를 이용하여 연구를 진행할 계획이다.

ACKNOWLEDGMENTS

본 결과물은 환경부의 재원으로 한국환경산업기술원의 생물다양성위협 외래생물관리기술개발사업의 지원을 받아 연구되었습니다.(KEITI 2021002270003).

REFERENCES

- [1] Kim, D. E, "Management System of Invasive Alien Species Threating Biodiversity in Korea and Suggestions for the Improvement", Journal of Environmental Impact Assessment, 27(1), 33-55(2018). <https://doi.org/10.14249/EIA.2018.27.1.33>
- [2] Ministry of Environment. 2014. 1st Management Plan of Alien Species (2014-2018).
- [3] Da Bin Kim · Kyung Ah Koo. "A Study on the Current Status and Improvement of Ecosystem Disturbance Species", 29(4), 59-81(2021).
- [4] Ikeda T, Asano M, Matoba Y, Abe G, "Present status of invasive alien raccoon and its impact in Japan", Global environmental research, 8(2), pp. 125-131, 2004.
- [5] National Institute of Ecology, "Ecological studies of alien species (VI)", Seocheon National Institute of Ecology
- [6] Jihae Kim et al, "Comparison of Fine Grained Classification of Pet Images Using Image Processing and CNN", Journal of Broadcast Engineering (JBE) Vol. 26, No. 2, pp. 175-183, 2021.
- [7] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, Quoc V. Le; "Learning Transferable Architectures for Scalable Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8697-8710
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun; "Deep Residual Learning for Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778
- [9] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, Zbigniew Wojna; "Rethinking the Inception Architecture for Computer Vision", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818-2826
- [10] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alexander A Alemi; "Inception-v4, inception-resnet and the impact of residual connections on learning", Thirty-first AAAI conference on artificial intelligence. 2017.
- [11] Simonyan, Karen, and Andrew Zisserman; "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556 (2014).
- [12] Kaiming He Ross Girshick Piotr Dollár, "Rethinking ImageNet Pre-training" arXiv:1811.08883v1 (2018).
- [13] Shorten, C., Khoshgoftaar, T.M., "A survey on Image Data Augmentation for Deep Learning", J Big Data 6, 60 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
- [14] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database", 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848
- [15] "Keras Applications," Keras, <https://keras.io/api/applications/>

저자소개

박 동 민 (Dong-Min Park)



2021년 : 대구대학교
정보통신공학과 (공학사)
2022년~현재 : 대구대학교
IT융합공학과 (공학석사)
관심분야 : Intelligent Image
Processing, AI

조 영 석 (Yeong-Seok Jo)



2004년 : 동국대학교
동물 분류생태학 (이학석사)
2014년 : Texas Tech University
야생동물관리학 (이학박사)
2020년 ~ 현재 : 대구대학교
생물교육과 교수
관심 분야 : 야생동물 분류/생태학,
보전생물학

염 석 원 (Seokwon Yeom)



1997년 : 한국전력사원
1999년 : 고려대학교 대학원
전자공학 (공학석사)
2006년 : 코네티컷 대학교
전자컴퓨터공학 (공학박사)
2007년~현재 : 대구대학교
AI학부, IT융합공학과 교수

2014년 : 메릴랜드 대학교 방문교수

2016년~현재 : 한국지능시스템학회 이사

2018년~현재 : 한국융합신호처리학회 부회장

관심분야 : Intelligent Image Processing,
Pattern Recognition, Target Tracking,
Optical Information Processing