

Attention layer를 활용한 이미지 기반 피부암 분류 시스템 Image-Based Skin Cancer Classification System Using Attention Layer

이규원, 우성희*

한국교통대학교 바이오메디컬융합학과

GyuWon Lee, SungHee Woo*

Department of Biomedical Engineering, Korea National University of Transportation, Chungju 27469, Korea

[요약]

고령화가 심화되면서 암 발병률이 증가하고 있다. 피부 암은 외적으로 보이지만 사람들이 알아채지 못하거나 가볍게 간과하는 경우가 많다. 이에 초기 발견 시기를 놓쳐 말기의 경우 생존율이 7.5~11%로 사망에 이를 수 있다. 하지만 피부 암을 진단함에 있어 육안으로 진단하는 것이 아닌 정밀검사, 세포 검사 등 시간과 비용이 많이 든다는 단점이 있다. 따라서 본 연구에서는 이러한 단점을 해결하기 위해 Attention CNN 모델 기반 피부암 분류 시스템을 제안한다. 이 시스템은 전문의로 하여금 피부 암을 초기에 발견하여 신속한 조치를 취할 수 있도록 하는데 도움을 줄 수 있다. 피부암 종류에 따른 이미지 데이터 불균형 문제에서 분포 비율이 낮은 데이터에는 Over Sampling 기법을, 분포 비율이 높은 데이터에는 Under Sampling 기법을 적용하여 완화하고 Attention layer가 없는 모델과 있는 모델을 비교하여 Attention layer가 없는 사전학습 모델에 추가한 피부암 분류 모델을 제안한다. 또한, 특정 클래스에 대하여 데이터 증강 기법을 강화하여 데이터 불균형 문제를 해결할 계획이다.

[Abstract]

As the aging population grows, the incidence of cancer is increasing. Skin cancer appears externally, but people often don't notice it or simply overlook it. As a result, if the early detection period is missed, the survival rate in the case of late stage cancer is only 7.5-11%. However, the disadvantage of diagnosing, serious skin cancer is that it requires a lot of time and money, such as a detailed examination and cell tests, rather than simple visual diagnosis. To overcome these challenges, we propose an Attention-based CNN model skin cancer classification system. If skin cancer can be detected early, it can be treated quickly, and the proposed system can greatly help the work of a specialist. To mitigate the problem of image data imbalance according to skin cancer type, this skin cancer classification model applies the Over Sampling, technique to data with a high distribution ratio, and adds a pre-learning model without an Attention layer. This model is then compared to the model without the Attention layer. We also plan to solve the data imbalance problem by strengthening data augmentation techniques for specific classes.

Key Words: Skin Cancer, Image Classification, Attention, Imbalanced Data

<http://dx.doi.org/10.14702/JPEE.2024.059>



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.

Received 23 January 2024; **Revised** 12 February 2024

Accepted 21 February 2024

***Corresponding Author**

E-mail: shwoo@ut.ac.kr

I. 서론

고령화 시대의 심화로 인해 노인들의 피부암 발병률이 높아지고 있다. 2023년 5월에 보고된 국가 암 등록사업 연례 보고서에 따르면 2020년 우리나라에 암 발생 환자는 총 247,952명으로, 그중 악성 흑색종은 625건을 차지하였다. 그중 65% 이상이 60대 이상에서 발생하였으며 계속 증가하는 추세이다.

악성 흑색종은 초기에 발견된 경우 생존율이 97%로 높지만 말기에 발견된 경우 피부나 피하조직보다 내부 장기로 전이되어 생존율이 7.5~11%로 초기 발견이 중요하다. 하지만 일반인은 시각적으로 어떤 피부 질환인지 분류하는 데 어려움이 있다.

본 연구에서 제안하는 피부암 검출 시스템을 통해 질환을 초기에 진단하여 환자의 생존 확률을 높이고 전문의의 질환 분류를 지원하여 오진을 최소화하고 오랜 진료에 따른 피로와 진료 시간을 단축할 수 있다.

II. 관련연구

A. 이미지 분류

HAM 10000은 Human Against Machine with 10,000 training images의 약자로 10,015개의 더 모스 코피 이미지로 구성된다. 제공된 데이터의 50% 이상은 조직병리학에 따라 확인되었고, 나머지에 대한 근거는 전문가 합의 또는 생체 내 현미경 검사를 통해 검증되었다.

S. S. Chaturved 외 2인의 연구에 따르면[1] 사전학습 모델로 MobileNet을 사용하였으며 600*450 픽셀의 이미지를 입력 사이즈 224*224로 변환하였다. 또한 훈련과 검증 데이터의 비율을 90:10으로 설정하여 학습을 진행하였다. MobileNet은 비교적 처리 속도가 빠르지만 정확도가 낮은 단점이 있다. C. H. Bae 외 3인의 연구[2]에서 CNN 기반 딥러닝 알고리즘 3종을 대상으로 미용 관련 피부질환 이미지 학습과 인식에 대한 연구를 진행하였다[2]. 피부질환 이미지 데이터는 치료 전후 결과를 비교하기 위해 발병 부위별로 수집하였다. 모델로는 Custom CNN Model, ResNet, SE-ResNet을 사용하였으며, 학습 깊이를 50, 101, 152로 구분하여 진행하였다. 50에서 101로 증가시켰을 때 2%의 정확도 향상 효과를 보였지만, 152로 증가시켰을 때는 정확도가 5.3% 감소하였다. 이를 통해 학습 깊이가 일정 부분 이상 증가하면 학습률이 감소한다는 것을 알 수 있다. Custom CNN Model은

71.5%, ResNet 알고리즘 기반의 학습 모델이 90.6%, 그리고 SE-ResNet 알고리즘 기반의 학습 모델은 95.3%의 정확도를 보였다[2].

T. H. H. Aldhyani 외 3인의 연구[3]는 28*28 픽셀로 변환하여 7가지 클래스(그림 1)를 도출하였다. 데이터 셋의 분포를 조정하는 과정에서 훈련 데이터 셋은 데이터 증강을 통해 클래스당 약 8000개의 데이터를 생성하였고 테스트 데이터

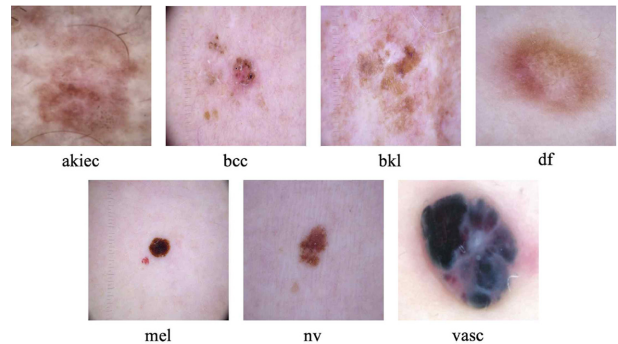


그림 1. 클래스 별 이미지

Fig. 1. Image by Class.

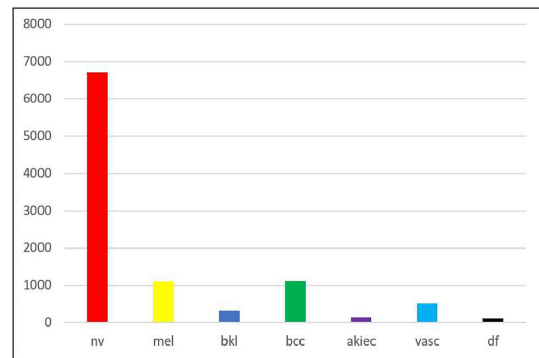


그림 2. 참고문헌[3]의 증강 전 데이터 분포

Fig. 2. Data distribution before augmentation in reference [3].

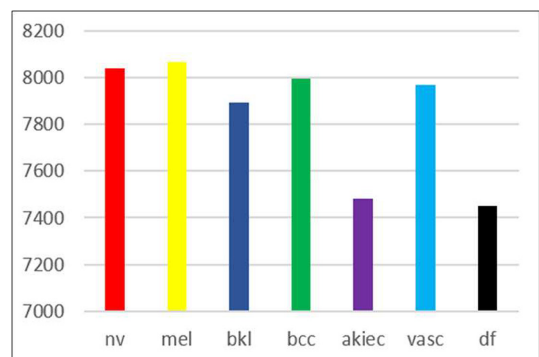


그림 3. 참고문헌[3]의 증강 후 데이터 분포

Fig. 3. Data distribution after augmentation in reference [3].

셋에도 데이터 증강 기법을 적용하였다. 그림 2과 그림 3에서 확인 가능한 이 방법은 모델의 성능을 보장할 수 없으며 약 100장의 이미지를 가진 클래스를 8000개 가까이 증강하여 모델 평가의 신뢰성이 떨어진다.

III. SE-DenseNet 121

본 연구에서 Attention layer가 있는 모델과 없는 모델을 비교한 뒤 “Attention layer가 피부 병변을 분류하는데 효과적이다.”를 증명한다. 또한 Attention layer가 존재하지 않는 사전 학습 모델에 추가하여 성능을 비교한다.

A. Attention layer

1) SE-block

Attention 기법 중 하나인 SE(Squeeze-and-Excitation) block은 채널 간의 상호 의존성을 강조하고, 이를 통해 강화된 특징을 모델 전반에 통합하여 성능을 향상시키는 역할을 한다. SE-block의 구조(그림 4)는 다음과 같다.

참조 문헌 [4]에 따르면, SE block은 네트워크 내에서 중요한 특징 맵을 강조함으로써 보다 효과적인 학습을 가능하게 한다. 다음은 SE block에 대한 보다 구체적인 설명이다.

특정 채널을 선택했을 때, 우리가 얻을 수 있는 정보는 H/W 차원의 2D 이미지로, 이는 주로 해당 픽셀의 local receptive field에 관한 정보를 담고 있다. 그러나 채널 전체의 종합적인 정보를 얻기 위해서는 모든 픽셀을 고려해야 한다. 이를 위해 Channel descriptor라 불리는 채널을 대표하는 스칼라 값 z 를 생성한다. 다음으로, 추출된 채널 별 descriptor를 활용해 실제로 채널 간의 의존성을 고려해야 한다. 이 과정에서 채널 간의 비선형 관계를 파악할 수 있어야 하며, 단순히 하나의 채널에만 가중치를 높게 주는 것이 아니라, 여러 채널에 동시에 Attention을 부여할 수 있어야 한다. 참고문헌 [4]의 저자들은 이를 위해 ReLU 함수와 Sigmoid 함수를 사

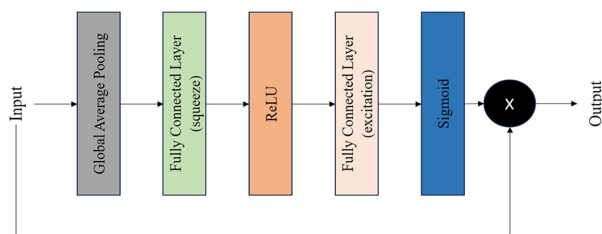


그림 4. SE block의 구조

Fig. 4. Structure of SE block.

용하였다. Channel descriptor vector를 선형 변환하여 ReLU 함수를 거치게 함으로써 필요한 정보만을 추출하고, 이후 Excitation 과정에서 Sigmoid 함수를 통과시켜 0과 1 사이의 값을 가지게 한다. 이 과정에서 각 특징들이 그 중요도에 따라 스케일링 되어, 모델이 주요한 특징에 더 집중할 수 있게 돕는다.

2) 사전 훈련된 신경망 입출력 계층 설계

사전학습 모델은 torchvision.models 제공되며 사용한 모델은 다음과 같다.

ResNet50, DenseNet, VGG-16, EfficientNet_b0, EfficientNet_b1, EfficientNet_b2

Attention 기법이 적용된 모델(EfficientNet_b0, EfficientNet_b1, EfficientNet_b2)과, 적용되지 않은 모델(ResNet50, DenseNet, VGG-16)의 비교를 통해 Attention 기법이 피부 병변 분류에 효과적임을 증명한다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = \frac{2 * precision * recall}{precision + recall} \quad (4)$$

Accuracy는 True와 False를 가리지 않고 모델이 바르게 분류한 비율을 의미한다. Precision은 모델이 True로 분류한 것 중 실제 값이 True인 비율이며, Recall은 실제 값이 True인 것 중 모델이 True로 분류한 비율을 의미한다. F1 score는 Precision과 Recall을 모두 고려하기 때문에 클래스의 label이 불균형할 때 모델의 성능을 정확히 평가할 수 있다.

표 1을 통해 Attention 기법이 적용된 모델이 적용되지 않

표 1. Attention기법에 따른 모델 성능 비교

Table 1. Comparison of model performance according to attention techniques

Model	Validation loss	Accuracy (1)	Precision (2)	Recall (3)	F1-Score (4)
EfficientNet_b2	0.3419	0.9171	0.9188	0.9184	0.9141
EfficientNet_b1	0.3875	0.9100	0.9116	0.9098	0.9099
EfficientNet_b0	0.4597	0.9043	0.9099	0.9038	0.9053
DenseNet-121	0.3744	0.8900	0.8962	0.8925	0.8938
ResNet-50	0.4779	0.8757	0.8408	0.8335	0.8336
VGG-16	0.5876	0.7986	0.8113	0.8121	0.8087

은 모델보다 Accuracy(1), F1-Score(4)가 높다는 것을 알 수 있다. Attention 기법이 적용되지 않은 모델 중 가장 성능이 좋은 DenseNet-121 모델에 SE-block을 추가하여 실험을 진행하였다.

B. SE-DenseNet 121

본 연구에서는 Attention 기법이 적용되지 않은 모델을 선정하고 Attention layer를 추가하여 모델의 성능을 향상시키려고 한다. ResNet과 DenseNet 중 DenseNet에 Attention layer를 추가한 이유는 DenseNet은 이전 모든 층에서 나온 출력을 현재 층에 직접 연결한다. 짧고 직접적인 경로를 통해 gradient가 흐르므로, 기울기 소멸 문제가 완화된다. 또한 Dense connection을 사용함으로써 각 층에서 생성된 특징들이 모두 연결되어 특성을 모두 공유하게 된다. 그래서 모델의 매개변수 수가 줄어들고, 더 효율적으로 학습할 수 있다. 최종적으로 Attention layer를 도입함으로써, 필요하지 않은 부분을 0으로 만들고, 중요한 부분만 특성 맵에 반영함으로써 피부 병변을 분류하는데 성능이 향상되었다. 또한 작은 모델 크기에 비해 ResNet-50과 VGG-16에 비해 상대적으로 크기가 작지만 높은 성능을 제공한다. 따라서 SE block을 추가하여 모델의 표현과 성능을 향상시킬 수 있으면서도 모델의 크기를 크게 증가시키지 않아도 된다. SE-DenseNet의 구

표 2. SE-DenseNet 검증

Table 2. SE-DenseNet Validation

Model	Validation loss	Accuracy	F1-Score
SE-DenseNet 121	0.3889	0.9043	0.9055

조(그림 5)와 Validation(표 2)에 대한 결과는 다음과 같다.

VI. 실험결과

A. 데이터 셋 구성

본 연구에 사용되는 데이터 셋은 데이터의 분포가 치중되어 있는 클래스가 존재한다. 예를 들면 현재 데이터 셋은 nv 클래스가 약 67%로 구성되며, 치중된 데이터 분포는 모델의 과적합을 불러오며, 해당 데이터를 통한 학습은 nv에 집중된 특징 추출을 진행할 가능성이 크다. 또한 이미지에 정상 피부 데이터와 질환 부위가 같이 있어 모델 학습률을 저하시킨다.

1) Over Sampling

Over Sampling은 Augmentation 기법을 이용하여 데이터 셋에서 분포가 적은 데이터를 회전은 90도, 180도, 270도 중 랜덤하게 적용했고 가우시안 블러, 플립 등을 랜덤 한 확률을 기반으로 각 클래스마다 500개를 생성한다.

2) Under Sampling

Under Sampling은 불균형한 클래스 샘플을 줄이기 위해 사용한다. 각 데이터에서 500개를 랜덤하게 뽑은 후 나머지 샘플을 제거함으로써 데이터 셋의 균형을 맞춘다.

B. 실험 방법

본 연구는 클래스 당 500개의 이미지로 Train 및 Validation

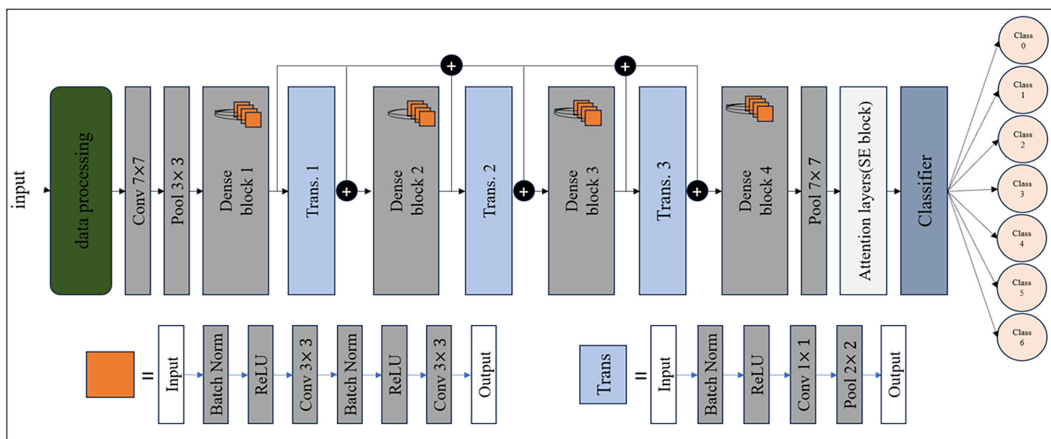


그림 5. SE-DenseNet의 구조

Fig. 5. Structure of SE-DenseNet.

단계를 진행하였고, Test 단계에서 클래스 당 15개로 총 105개의 이미지로 실험을 진행하였다.

1) 실험 목적 및 절차

피부암 데이터 중에서 특징이 잘 드러나지 않거나, 노이즈가 많은 데이터는 학습 성능을 저하시킨다. 이를 방지하고자 Attention Layer를 사전학습된 모델에 추가함으로써 모델의 성능을 높이고자 한다. 이미지 데이터는 전처리를 통해 준비하며 사전 훈련된 신경망을 불러와 실험의 목적에 맞게 관련 계층을 강화한다.

2) 데이터 전처리

이미지 샘플링을 통해 클래스 당 500개의 이미지 데이터를 준비하였다. Train 80%와 Validation 20%로 분리 후, 이미지 크기를 600*450 pixels에서 224*224 pixels로 조정하였다.

3) 모델 설계

본 논문은 7가지의 사전학습 모델과 2가지의 개선된 모델을 사용하여 실험을 진행하였다. Test Data는 각 클래스 별 15장의 이미지 포함하며, 총 105장의 이미지로 실험을 진행하였다.

Attention 기법을 적용한 모델 SE-DenseNet 121, EffcientNet_b1, EffcientNet_b2 모두 Attention 기법이 적용되지 않은 모델보다 높게 도출되었다. 표 1에서 확인 가능한 Validation Dataset에 비해 표 3의 Test Dataset 대한 모델 성능 지표는 현저히 낮은 성능을 보였다. 또한 그림 6과 그림 7을 보았을 때 akiec와 df에 대해 분류하는데 어려움이 있었다. 이는 모델이 과적합(Overfitting) 되었을 가능성이 있다. 모델이 훈련 데이터에 너무 많이 적합 되면 훈련 데이터 셋에 대한 예측 정확도는 높지만, 실제 데이터에 대한 예측 정확도는 낮다.

표 3. 실험 결과

Table 3. Experiment Result

Model	Test loss	Accuracy	precision	Recall	F1-Score
SE-DenseNet 121	2.2570	0.5048	0.4015	0.5048	0.4369
EffcientNet_b2	2.3475	0.4667	0.5606	0.4667	0.4032
EffcientNet_b1	2.5201	0.4095	0.5437	0.4095	0.3744
EffcientNet_b0	3.8258	0.3905	0.3823	0.3905	0.3307
VGG-16	5.5829	0.3810	0.3820	0.3810	0.3300
ResNet-50	3.6476	0.3810	0.4139	0.3810	0.3081
DenseNet-121	3.6545	0.3714	0.3568	0.3714	0.3046

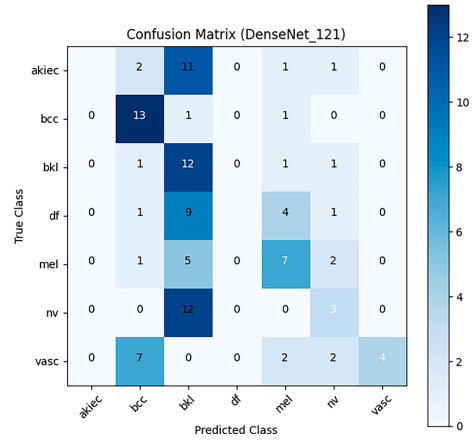


그림 6. DenseNet-121의 혼동행렬

Fig. 6. Confusion Matrix of DenseNet-121.

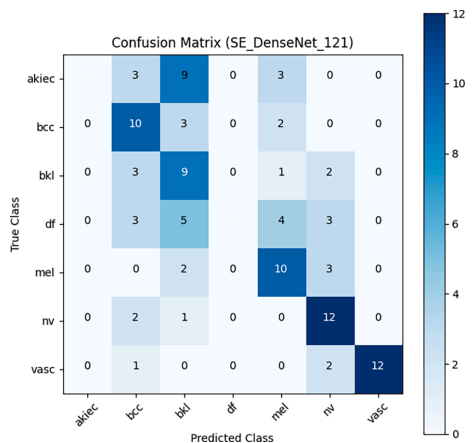


그림 7. SE-DenseNet-121의 혼동행렬

Fig. 7. Confusion Matrix of SE-DenseNet-121.

V. 결론

본 연구에서는 피부암 분류를 위해 Attention 기법이 적용된 모델과 그렇지 않은 모델을 사용하였으며, 그 결과 Attention 기법이 피부 병변 분류에 유용하다는 결론을 얻었다. 또한 데이터 클래스 간의 비율 차이를 줄여 모델의 학습 성능을 높였다. Over Sampling과 Under Sampling을 활용하여 데이터 셋의 치중된 분포를 조정하고, 이를 통해 학습 성능을 향상시키려 하였다. Attention layer가 추가된 모델들이 피부 병변을 분류하는데 더 좋은 성능을 보였다. Validation에 비해 Test의 성능이 낮게 나왔지만 모델을 강화하고 데이터 과적합에 대한 문제를 해결한다면 우수한 성능을 낼 수 있다. 또한 Confusion Matrix를 통한 분석 결과, Test 성능이

Validation에 비해 낮게 나오는 원인으로 특정 클래스에서의 과적합을 파악하였다. 향후 특정 클래스에 대한 augmentation 기법을 강화하고, 데이터 불균형 문제가 발생하였을 때 Undersampling이 아닌 K-means 클러스터링을 통해 중심에 있는 데이터 위주로 선정하여 정확도를 향상시키는 방법과 클래스별 모델의 가중치를 조정함으로써 성능을 향상시킬 계획이다.

참고문헌

[1] S. S. Chaturvedi, K. Gupta, and P. S. Prasad, "Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet," in A. Abraham et al. (Eds.), *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, Springer Singapore, 2021.

[2] C. H. Bae, W. Y. Cho, H. J. Kim, and O. K. Ha, "An experimental comparison of CNN-based deep learning Algorithms for recognition of beauty-related skin disease," *Journal of the Korea Computer Information Society*, vol. 25, no. 12, pp. 25-34, 2020.

[3] T. H. H. Aldhyani, A. Verma, M. H. Al-Adhaileh, and D. Koundal, "Multi-class skin lesion classification using a

lightweight dynamic kernel deep-learning-based convolutional neural network," *Diagnostics*, vol. 12, no. 9, pp. 2048, August, 2022.

[4] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7132-7141, 2018.

[5] J. Park and Y. Han, "ResUNet++ for skin cancer lesion segmentation," *Journal of the Korean Institute of Electrical Engineers*, vol. 59, no. 2, pp. 95-100, 2022.

[6] Y. Hwang, "Multi-feature fusion-based classification of malignant melanoma and basal cell carcinoma in dermoscopy images," Ph.D. dissertation, Dongguk University, 2022.

[7] M. Kwak, K. Kim, and J. Choi, "Multi-scale attention and deep ensemble-based animal skin lesion classification technique," *Journal of the Korea Multimedia Society*, vol. 25, no. 8, pp.1212-1223, 2022.

[8] J. Oh, "Skin lesion classification using inception-resnet-V2," M.S. thesis, Chungbuk University, 2023.

[9] M. S. Akter, H. Shahriar, S. Sneha, and A. Cuzzocrea, "Multi-class skin cancer classification architecture based on deep convolutional neural network," in *Proceedings of the 2022 IEEE International Conference on Big Data (Big Data)*, 2022.



이 규 원 (GyuWon Lee) _정회원

2020년 2월 ~ 2022년 2월 : 한국교통대학교 컴퓨터공학과
2022년 2월 ~ 현재 : 한국교통대학교 바이오메디컬융합학과
관심분야 : 의료영상진단, 딥러닝, DICOM, 의료인공지능



우 성 희 (SungHee Woo) _종신회원

1995년 ~ 현재 : 한국교통대학교 컴퓨터공학과 교수
2020년 ~ 현재 : 한국교통대학교 바이오메디컬융합학과 교수
1999년 2월 : 충북대학교 전자계산학과 이학박사
관심분야 : 정보보안, 컴퓨터 네트워크, 의료정보보안, 데이터통신