

딥러닝을 이용한 마스크 착용 여부 검사 시스템

남충현¹ · 남은정¹ · 장경식^{2*}

Mask Wearing Detection System using Deep Learning

Chung-hyeon Nam¹ · Eun-jeong Nam¹ · Kyung-Sik Jang^{2*}

¹Graduate Student, Department of Computer Engineering, Korea University of Technology and Education, Cheonan 31253, Korea

^{2*}Professor, Department of Computer Engineering, Korea University of Technology and Education, Cheonan 31253, Korea

요 약

최근 COVID-19로 인해 마스크 착용 여부 자동 검사 시스템에 신경망 기술들을 적용하는 연구가 활발히 진행되고 있다. 신경망 적용 방식에 있어서 1단계 검출 방식 또는 2단계 검출 방식을 사용하며, 데이터를 충분히 확보할 수 없는 경우 사전 학습된 신경망에 대해 가중치 미세 조절 기법을 적용하여 학습한다. 본 논문에서는 얼굴 인식부와 마스크 검출부로 구성되는 2단계 검출 방식을 적용하였으며, 얼굴 인식부에는 MTCNN 모델, 마스크 검출부에는 ResNet 모델을 사용하였다. 마스크 검출부는 다양한 실 상황에서의 인식률과 추론 속도 향상을 위하여 5개의 ResNet 모델을 적용하여 실험하였다. 학습 데이터는 웹 크롤러를 이용하여 수집한 17,219개의 정지 영상을 이용하였으며, 1,913개의 정지 영상과 1분 동영상 2개에 대해 각각 추론을 실시하였다. 실험 결과 정지 영상인 경우 96.39%, 동영상인 경우 92.98%의 높은 정확도를 보였고, 동영상 추론 속도는 10.78fps임을 확인하였다.

ABSTRACT

Recently, due to COVID-19, studies have been popularly worked to apply neural network to mask wearing automatic detection system. For applying neural networks, the 1-stage detection or 2-stage detection methods are used, and if data are not sufficiently collected, the pretrained neural network models are studied by applying fine-tuning techniques. In this paper, the system is consisted of 2-stage detection method that contain MTCNN model for face recognition and ResNet model for mask detection. The mask detector was experimented by applying five ResNet models to improve accuracy and fps in various environments. Training data used 17,217 images that collected using web crawler, and for inference, we used 1,913 images and two one-minute videos respectively. The experiment showed a high accuracy of 96.39% for images and 92.98% for video, and the speed of inference for video was 10.78fps.

키워드: 딥러닝, 영상 처리, 객체 인식, 영상 분류

Keywords: Deep learning, Image processing, Object detection, Image classification

Received 5 October 2020, Revised 10 November 2020, Accepted 21 November 2020

* Corresponding Author Kyung-Sik Jang(E-mail:ksjang@koreatech.ac.kr, Tel:+82-41-560-1352)

Professor, Department of Computer Engineering, Korea University of Technology and Education, Cheonan, 31253, Korea

Open Access <http://doi.org/10.6109/jkiice.2021.25.1.44>

print ISSN: 2234-4772 online ISSN: 2288-4165

© 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.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

최근 영상 처리 분야에 신경망 기술이 적극적으로 활용되고 있다. 특히, 번호판 인식, 침입 탐지, 물체 유기, 연기 및 화재 감지 등 실시간 지능형 CCTV에서 이미지 내에 특정 객체가 어느 위치에 있는지 찾아내는 객체 탐지, 특정 객체가 어떤 분류에 속하는지 판별하는 이미지 분류 등 영상 처리 기술에 신경망을 접목하여 높은 성능을 내고 있다.

신경망 기술은 학습 과정에서 다양한 특징을 반영할 수 있는 데이터가 많을수록 높은 성능을 보이며[1], 시스템 운용을 위한 신경망은 학습 시간을 최소화 하고자 기존 높은 성능을 보이는 사전 학습된 신경망을 가중치 미세 조절 기법(Fine-Tuning)을 이용하여 소량의 데이터를 학습한다[2].

이러한 신경망 기술은 관리자가 필요한 CCTV 환경에서 사람의 개입을 최소화 할 수 있는 자동화 시스템에 적용되고 있으며[3][4], 현재 COVID-19로 인해 마스크 착용 여부를 자동으로 검사하는 시스템에 대해서도 활발히 연구되고 있다[5][6]. 마스크 착용 여부 검사 시스템에 적용되는 신경망 기술은 후보 얼굴 관심 영역 검출 단계와 마스크 착용 여부 검사 단계로 이루어진 2단계 검출 방법과 마스크 착용 영역을 한 번에 검출하는 1단계 검출 방법이 있으며, 2단계 검출 방식은 단계적 검출 방법을 사용하기 때문에 상대적으로 1단계 검출 방식보다 추론 속도가 늦지만 높은 정확도를 보인다는 장점을 가지고 있다[7].

본 논문에서는 2단계 검출 방식을 적용하여 건물 출입자에 대해 마스크 착용 여부를 식별하는 시스템을 제안한다. 자세히는 MTCNN 모델[8]을 사용한 얼굴 인식부와 ResNet(Residual Network) 모델[9]을 사용한 마스크 검출부로 구성되어 있으며, 웹 크롤러를 이용하여 수집한 학습 데이터 17,219장을 가중치 미세 조절 기법을 이용하여 학습하였다.

제안하는 마스크 착용 여부 검사 시스템을 실험하기 위해 ResNet 모델은 5가지(18, 34, 50, 101, 152)를 각각 학습하여 정지 영상과 및 동영상에 대해 성능을 평가하였다. 그 결과 정지 영상에 대해서는 ResNet101 모델이 96.39%로 가장 높은 정확도를 보였고, ResNet18 모델이 10.18ms로 가장 빠른 추론 속도를 보였다. 또한 동영상에 대해서는 ResNet152 모델이 92.58%의 가장 높은

정확도를 보였으며, ResNet18 모델이 가장 빠른 추론 속도인 10.73ms를 보였다.

II. 관련 연구

2.1. 이미지 분류

이미지가 어떤 범주에 속해 있는지 판별하는 이미지 분류에 많이 사용되는 신경망 모델에는 AlexNet[10], GoogLeNet[11], VGGNet[12] 등이 있다.

먼저 AlexNet은 5개의 합성곱 층(Convolutional layer)과 3개의 전 완전 연결 층(fully-connected layer)으로 이루어져있으며, 두 개의 GPU를 사용하여 병렬 연산을 수행하는 것이 특징이다.

두 번째로 GoogLeNet은 22개의 층과 9개의 인셉션 모듈을 가지고 있으며, 인셉션 모듈 내에는 특징을 효과적으로 추출하기 위한 다양한 크기의 필터를 적용하였다. 또한 합성곱 신경망(CNN, Convolutional Neural Network)의 네트워크가 깊어질수록 학습해야 하는 파라미터의 수가 늘어나 연산량이 많아진다는 점을 보완하기 위해 1x1 필터가 적용된 콘볼루션 층을 사용하였다.

마지막으로 VGGNet은 GoogLeNet에 비해 단순한 구조를 가지고 있으며, 3x3의 필터 크기를 적용한 콘볼루션 층을 여러 개 쌓아 네트워크의 깊이가 모델의 성능을 높이는 데 중요한 역할을 한다는 것을 보여주었다. 하지만 네트워크를 깊게 쌓은 만큼 연산량이 높아지며 이에 따라 메모리를 많이 사용한다는 단점이 있다.

2.2. 객체 탐지

입력 이미지 내에 관심 객체의 위치를 찾고, 찾은 객체가 어떤 범주에 속해 있는지 판별하는 객체 탐지는 2단계 검출 방법을 이용한 R-CNN[13](Region-based Convolutional Neural Network), Fast R-CNN(Fast Region-based Convolutional Neural Network)[14]등과 1단계 검출 방법을 이용한 Yolo(You Only Look Ones) [15] 등의 모델이 연구되어져 왔다.

R-CNN 모델은 입력 이미지에 대해 선택적 탐색(Selective Search) 알고리즘을 적용한 후보 관심 영역들을 찾아내는 관심 영역 제안 단계와 합성곱 신경망을 이용하여 각 후보 관심 영역에 대해 특징을 추출하는 단계, 추출된 특징이 어느 범주에 속해 있는지 분류하고

위치를 예측하는 단계로 나누어져 있다.

하지만 선택적 탐색 알고리즘을 이용해 찾은 후보 관심 영역들을 모두 합성곱 신경망을 이용해 특징을 추출하기 때문에 추론 속도가 낮다는 단점이 있다.

이러한 문제를 해결하기 위해 Fast R-CNN 모델은 입력 이미지에 대해 합성곱 신경망을 이용하여 추출된 특징들에 대해 선택적 탐색 알고리즘을 사용하는 방식으로 제안되었으며, 결과적으로 R-CNN보다 추론 속도를 향상시켰다.

관심 영역 제안 단계와 관심 영역이 어느 범주에 속해 있는지 분류하는 단계로 나누어 진행되는 R-CNN, Fast R-CNN 모델과는 달리 Yolo는 각 이미지를 SxS개의 서브 이미지로 분할하고 서브 이미지 내의 객체를 찾는 방식으로 관심 영역 제안 단계 없이 한 번에 객체 탐지 작업을 수행할 수 있기 때문에 상대적으로 빠른 추론 속도를 보여주었다. 하지만 한 번에 수행하는 만큼 R-CNN, Fast R-CNN 모델보다 성능이 낮다는 단점이 있다.

본 논문에서 마스크 착용 여부 검사 시스템을 제안하기 위해 기존 연구에서 후보 관심 영역 제안 단계와 후보 관심 영역 분류 단계로 2단계 검출 방식이 적용된 R-CNN 모델을 착안하여 얼굴 인식 단계와 마스크 검출 단계를 가진 시스템을 제안하였다.

III. 마스크 착용 여부 검사 시스템

본 논문에서 제안한 마스크 착용 여부 검사 시스템은 그림1과 같이 입력 이미지에 대해 후보 얼굴 관심 영역들을 추출하는 얼굴 인식 단계와 추출된 후보 얼굴 관심 영역들이 마스크를 착용하고 있는지 판별하는 마스크 검출 단계로 이루어진 2단계 검출 방식을 이용하였다.

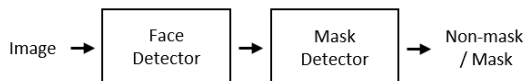


Fig. 1 The architecture of real-time mask detection system

3.1. 얼굴 검출기

얼굴 검출기는 입력된 이미지에 대해 후보 얼굴 위치를 찾아 추출하는 후보 얼굴 관심 영역 추출 작업을 수행하며, 본 논문에서는 얼굴 검출 작업을 수행하기 위해 최근 우수한 성능을 보인 MTCNN을 이용하였다.

MTCNN 모델은 단계적 모델로 구성되어 있으며, 그림2와 같이 이미지 내 관심 영역을 탐색하여 후보 얼굴 관심 영역들을 추출하는 P-Net (Proposal Network), 추출된 후보 관심 영역들 중에서 의미 있는 관심 영역들만 추출하는 R-Net(Refinement Network), 의미 있는 관심 영역들의 최종 얼굴 경계를 인식하는 O-Net(Output Network)으로 이루어져 있다.

각 네트워크의 합성곱 층은 얼굴의 특징을 찾기 위해 3x3 필터를 사용하였으며, 출력으로 얼굴인지 판별하는 얼굴 분류, 얼굴 경계와 눈, 코, 입에 대한 얼굴 특징을 나타내는 좌표 값에 대해 출력한다.

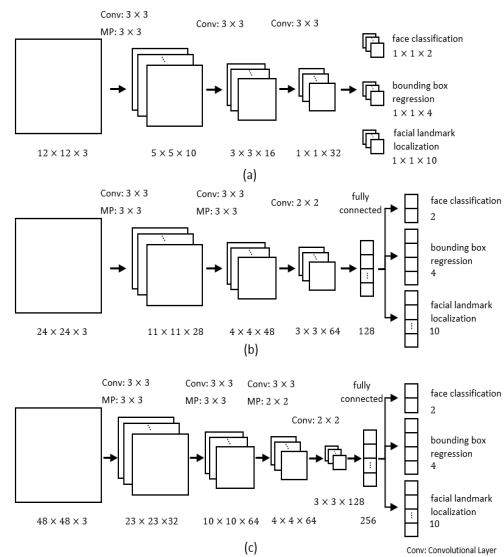


Fig. 2 The architecture of MTCNN: (a) P-Net, (b) R-Net, (c) O-Net

수식1은 MTCNN 모델을 학습하기 위한 수식이며, 학습 과정에서 얼굴 분류, 얼굴 경계 예측, 얼굴 특징 점 예측에 대한 손실 함수들의 합을 최소화하는 것을 목표로 한다.

$$\min \sum_{i=1}^N \sum_{j \in \{det, box, landmark\}} \alpha_j \beta_i^j L_i^j \quad (1)$$

수식은 네트워크 출력에 대한 손실 값을 다르게 하기 위해 작업에 대한 중요도 $\alpha_i^{x \in \{face, box, landmark\}}$ 매개 변수를 두어 학습 과정에서 기울기 값을 다르게 할 수 있으며, 3개의 손실 함수에 대한 선택 매개 변수 $\beta_i^{x \in \{face, box, landmark\}} \in \{0, 1\}$ 를 두어 각 손실 함수를 각

각 학습시킬 수 있다. L_i^j 은 i 번째 이미지에 대해 얼굴 분류, 얼굴 경계 예측, 얼굴 특징 점 예측에 대한 손실 함수이다.

먼저 얼굴 분류 손실 함수 L_i^{det} 는 신경망이 얼굴인지 아닌지 분류하기 위해 수식2와 같이 이진 크로스 엔트로피 손실 함수를 사용한다.

$$L_i^{det} = -(y_i^{det} \log(p_i) + (1 - y_i^{det}) \log(1 - p_i)) \quad (2)$$

수식에서 입력된 i 번째 이미지에 대해 p_i 는 얼굴 분류에 대한 모델의 예측 값이며, $y_i^{det} \in \{0, 1\}$ 는 정답 값을 뜻한다.

두 번째로 입력된 이미지 내의 얼굴 위치를 학습하는 얼굴 경계 손실 함수 L_i^{box} 는 수식3과 같이 유클리디안 거리 손실 함수를 사용하여 신경망이 예측한 얼굴 위치와 정답 위치간의 거리차를 이용하여 학습한다.

$$L_i^{box} = \|\hat{y}_i^{box} - y_i^{box}\|_2^2 \quad (3)$$

수식에서 \hat{y}_i^{box} 는 얼굴 경계에 대해 모델이 예측한 값이며, $y_i^{box} \in \{(x_1, y_1), (x_2, y_2)\}$ 는 얼굴 경계에 대한 정답 좌표 값을 의미한다.

마지막으로 얼굴 특징 점 예측에 대한 손실 함수 $L_i^{landmark}$ 는 얼굴 내의 눈, 코, 입에 대한 좌표를 예측하며, 얼굴 경계 추출과 동일하게 수식4와 같이 예측 값과 정답 값 간의 유클리디안 거리 손실 함수를 이용하여 학습한다.

$$L_i^{landmark} = \|\hat{y}_i^{landmark} - y_i^{landmark}\|_2^2 \quad (4)$$

3.2. 마스크 착용 여부 검출기

마스크 착용 여부 검출기는 얼굴 검출기에서 추출된 후보 얼굴 관심 영역들을 각각 분류한 이미지들에 대해 마스크 착용 여부를 판단하는 이진 분류 작업이며, 이미지 분류에 적합한 신경망 모델인 ResNet 모델을 사용하였다.

ResNet 모델은 모델의 층이 많을 때, 학습 과정에서 입력 층으로 갈수록 기울기 값이 작아지는 기울기 소실 문제(Vanishing Gradient Problem)가 발생하는 문제를 완화하기 위해 그림3과 같이 잔여 블록(Residual Block)을 사용하였으며, 입력 값과 입력 값에 대해 하나 이상

의 층들의 출력 값을 합하여 출력 값으로 사용할 수 있도록 이루어져 있다. 이러한 구조를 통해 신경망의 깊이가 깊어질수록 학습 과정에서의 기울기 소실 문제에 대한 문제를 완화할 수 있다는 장점을 갖고 있어 기존 신경망보다 높은 정확도를 얻을 수 있다.

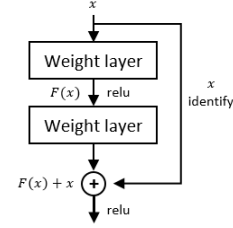


Fig. 3 Residual Block

마스크 착용 여부 학습을 위한 손실함수 L_k^{mask} 는 수식5와 같이 마스크 이미지에 대한 예측 값 p_k 과 정답 값 y_k^{mask} 에 대해 이진 크로스 엔트로피 손실 함수를 사용하였다.

$$L_k^{mask} = -(y_k^{mask} \log(p_k) + (1 - y_k^{mask}) \log(1 - p_k)) \quad (5)$$

기존 연구에서 ResNet 모델은 층의 개수(18, 34, 50, 101, 152)에 따라 정확도 및 추론 속도에 대한 성능 차이가 있었음을 확인할 수 있었고[16], 본 논문에서도 층 개수에 따른 성능 평가를 위해 정지 영상과 동영상을 이용하여 각각 정확도와 추론 속도를 실험하였다.

IV. 실험 결과 및 분석

4.1. 실험 환경 및 데이터

제안한 마스크 착용 여부 검사 시스템의 모델 학습을 위한 하드웨어는 Intel i7-7700K 4.20GHz, 64GB RAM, GeForce RTX 2080 Ti로 구성하였고, 모델 학습을 위한 딥러닝 프레임워크는 PyTorch를 사용하였다. 또한 실험에 사용된 얼굴 검출기 모델과 마스크 착용 여부 검출기 모델은 각각 사전 학습된 MTCNN 모델과 ResNet 모델을 사용하였으며, ResNet 모델은 학습률을 $1e-5$ 로 설정하여 가중치 미세 조절 기법을 이용해 학습하였다.

학습은 5회 진행하였고, 최적화 알고리즘은 Adam Optimizer를 사용하였다. 실험에 필요한 데이터 셋은 웹

크롤러를 이용해 수집하였으며, 마스크 착용한 얼굴을 포함한 이미지 17,219장을 사용하여 학습하였고, 정지 영상 정확도 및 추론 속도 측정을 위해 이미지 1,913장을 사용하였다. 실시간 동영상은 총 285장의 프레임으로 구성된 직접 촬영한 영상을 사용하여 정확도 및 fps(frame per second) 측정을 하였으며, 수작업으로 정지 영상 및 동영상에 대한 레이블링 작업을 실시하였다.

4.2. 마스크 착용 여부 검출기 정확도

표1은 정지 영상을 이용하여 마스크 착용 여부 검출기에 대해 각 ResNet 모델의 5가지 층 개수(18, 34, 50, 101, 152)에 따른 정확도 및 추론 시간을 실험한 결과를 나타낸다. 정확도는 96.39%로 ResNet101 모델이 가장 높게 나왔으며, ResNet18 모델이 96.12%로 가장 낮게 나왔다. 반면, 추론 시간은 ResNet18 모델이 평균 10.18ms의 가장 빠른 추론 속도 결과가 나왔으며, ResNet152 모델은 평균 12.70ms로 가장 늦은 추론 속도로 측정되었다.

Table. 1 The accuracy and inference time of mask detector about the number of each ResNet layers for images

# of ResNet layers	18	34	50	101	152
Accuracy(%)	96.12	96.38	96.13	96.39	96.29
Infer. Time(ms)	10.18	10.29	10.44	11.50	12.70

이를 실험을 통해 정지 영상에 대해 ResNet 모델에 대한 정확도는 평균 96%의 성능을 보였으며, 층의 개수에 따라 성능이 크게 변하지 않음을 확인할 수 있었고, 추론 속도는 모델 네트워크 층의 개수가 증가함에 따라 속도가 늦어짐을 확인하였다.

4.3. 실시간 동영상에 대한 정확도 및 fps 측정

실시간 동영상에 대해 ResNet 모델의 층 개수에 따른 정확도 및 fps 측정은 표2와 같은 실험 결과를 보였다. ResNet152 모델이 92.98%로 정확도가 가장 높게 나왔으며, ResNet18의 정확도가 90.59%로 가장 낮게 나왔다. 반면에 fps는 ResNet18 모델이 가장 빠르고 ResNet152 모델이 가장 늦었으며, 층의 개수가 많아질수록 낮은 fps를 갖게 된다는 것을 실험을 통해 확인하였다.

Table. 2 The accuracy and fps of mask detector about the number of each ResNet layers for video

# of ResNet layers	18	34	50	101	152
Accuracy(%)	90.59	92.40	91.81	91.81	92.98
fps	10.73	10.42	10.38	9.84	9.46

V. 결 론

최근 영상 처리 분야에 신경망 기술이 적극적으로 활용되고 있으며, 객체 탐지, 이미지 분류 등 다양한 영상 처리 기술에 신경망을 접목하여 높은 성능을 보이고 있으며, 현재 COVID-19로 인해 마스크 착용 여부 검사 시스템에 대해서도 신경망 기술들을 적용하는 연구가 활발히 진행되고 있다.

마스크 착용 여부 검사 시스템에 적용되는 신경망 기술은 후보 얼굴 관심 영역 검출 단계와 마스크 착용 여부 검사 단계로 이루어진 2단계 검출 방법과 마스크 착용 영역을 한 번에 검출하는 1단계 검출 방법이 있으며, 2단계 검출 방식은 단계적 검출 방법을 사용하기 때문에 상대적으로 1단계 검출 방식보다 추론 속도가 늦지만 높은 정확도를 보인다는 장점을 가지고 있다.

본 논문에서는 얼굴 인식부와 마스크 검출부로 구성된 2단계 검출 방식을 적용한 마스크 착용 여부 검사 시스템을 제안하며, 얼굴 인식부와 마스크 검출부에 각각 MTCNN모델, ResNet 모델을 사용하였다.

실험을 위해 학습 데이터는 웹 크롤러를 이용하여 수집한 17,219개의 정지 영상을 이용하였으며, 1,913개의 정지 영상과 1분 동영상 2개에 대해 각각 추론을 실시하였다. 또한 마스크 검출부에는 다양한 실 상황에서의 인식률과 추론 속도를 향상시키기 위해 5개의 ResNet 모델을 적용하여 각각 실험하였다. 시험 결과로 정지 영상인 경우 ResNet101모델이 96.39%, 동영상인 경우 ResNet152모델이 92.98%의 높은 정확도를 보였으며, 추론 속도는 ResNet18모델이 10.78fps로 가장 빠른 추론 속도를 보였음을 확인하였다.

현재 본 논문에서 제안한 마스크 착용 여부 검사 시스템은 얼굴 검출기 모델 성능에 종속적이기 때문에 향후 얼굴 검출기 성능 개선을 위한 신경망 모델과 전통적인 얼굴 검출 알고리즘을 결합한 새로운 방식에 대한 추가적인 연구가 필요하다.

ACKNOWLEDGEMENT

This paper was supported by Education and Research Promotion Program of KOREATECH.

REFERENCES

- [1] E. Blasch, S. Liu, Z. Liu, and Y. Zheng, "Deep Learning Measures of Effectiveness," in *Proceeding of the 2018 IEEE National Aerospace and Electronics Conferences*, Dayton, pp. 254-261, 2018.
- [2] J. Talukdar, S. Gupta, P. S. Rajpura, and R. S. Hegde, "Transfer Learning for Object Detection using State-of-the-Art Deep Neural Networks," in *Proceeding of the 5th International Conference on Signal Processing and Integrated Networks*, Noida, pp. 78-83, 2018.
- [3] S. S. Thomas, S. Gupta, and V. K. Subramanian, "Smart Surveillance Based On Video Summarization," in *Proceeding of the 17th IEEE Region 10 Symposium*, India, pp. 1-5, 2017.
- [4] A. K. Diop, S. Meza, M. Gordan, and A. Vlaicu, "LDA based classification of video surveillance sequences using motion information," in *Proceeding of the 20th International Conference on Advanced Communication Technology*, Korea, pp. 1-1, 2018.
- [5] S. H. Lee, H. G. Kwon, Y. J. Kim, J. S. Jeong, and H. J. Seo, "Development of CCTV for Identification of Maskless Wearers based on Deep Learning," in *Proceeding of the 28th Korea Society of Computer Information*, Korea, pp. 317-318, 2020.
- [6] W. Y. Cho, S. L. Park, H. S. Kim, and T. J. Yun, "Development of AI Systems for Counting Visitors and Check of Wearing Masks Using Deep Learning Algorithms," in *Proceeding of the 28th Korea Society of Computer Information*, Korea, pp. 285-286, 2020.
- [7] P. Soviany and R. T. Ionescu, "Optimizing the Trade-off between Single-Stage and Two-Stage Deep Object Detectors using Image Difficulty Prediction," in *Proceeding of the 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, Romania, pp. 209-214, 2018.
- [8] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499-1503, Oct. 2016.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceeding of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 770-778, 2015.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceeding of the 25th International Conference on Neural Information Processing Systems*, Nevada, pp. 94-90, 2017.
- [11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," in *Proceeding of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 1-9, 2015.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceeding of the 2015 International Conference on Learning Representations*, CA, pp. 1-14, 2015.
- [13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceeding of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Ohio, pp. 580-587, 2014.
- [14] R. Girshick, "Fast R-CNN," in *Proceeding of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 1440-1448, 2015.
- [15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceeding of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 779-788, 2015.
- [16] R. U. Khan, X. Zhang, R. Kurnar, and E. O. Aboagye, "Evaluating the Performance of ResNet Model Based on Image Recognition," in *Proceeding of the 2018 International Conference on Computing and Artificial Intelligence*, Indonesia, pp. 86-90, 2018.



남충현(Chung-Hyeon Nam)

2019년 한국기술교육대학교 컴퓨터공학부 공학사
 2019년 ~ 현재 한국기술교육대학교 컴퓨터공학과
 공학석사과정
 ※관심분야: Natural Language Processing,
 Image Processing, Incremental
 Learning, Data Mining



남은정(Eun-Jeong Nam)

2019년 부산외국어대학교 경제데이터금융학부
 학사
 2019년 ~ 현재 한국기술교육대학교 컴퓨터공학과
 공학석사과정
 ※관심분야: Image Processing,
 Image Segmentation



장경식(Kyung-Sik Jang)

1987년 고려대학교 전자공학과 공학사
 1989년 한국과학기술연구원 전기전자공학과
 공학석사
 1998년 동경공업대학 전기전자공학 공학박사
 1999년 ~ 현재 한국기술교육대학교 컴퓨터공학부
 교수
 ※관심분야: Embedded System, Natural Language
 Processing, Incremental Learning