Journal of Biomedical Engineering Research 43: 131-135 (2022)
http://dx.doi.org/10.9718/JBER.2022.43.3.131

학 술 논 문

딥러닝을 위한 마스크 착용 유형별 데이터셋 구축 및 검출 모델에 관한 연구

황호성[1] · 김동현[2] · 김호철[3]*

[1]을지대학교 의료인공지능학과, [2]시냅스이미징, [3]을지대학교 방사선학과

The Study for Type of Mask Wearing Dataset for Deep learning and Detection Model

Ho Seong Hwang[1], Dong heon Kim[2] and Ho Chul Kim[3]*

[1]Department of Medical Artificial Intelligent, Eul-Ji University, [2]Synapseimaging, [3]Department of Radiological Science, Eul-Ji University

(Manuscript received 24 May 2022 ; revised 4 June 2022 ; accepted 7 June 2022)

Abstract: Due to COVID-19, Correct method of wearing mask is important to prevent COVID-19 and the other respiratory tract infections. And the deep learning technology in the image processing has been developed. The purpose of this study is to create the type of mask wearing dataset for deep learning models and select the deep learning model to detect the wearing mask correctly. The Image dataset is the 2,296 images acquired using a web crawler. Deep learning classification models provided by tensorflow are used to validate the dataset. And Object detection deep learning model YOLOs are used to select the detection deep learning model to detect the wearing mask correctly. In this process, this paper proposes to validate the type of mask wearing datasets and YOLOv5 is the effective model to detect the type of mask wearing. The experimental results show that reliable dataset is acquired and the YOLOv5 model effectively recognize type of mask wearing.

Key words: Mask detection, Deep learning, Object classification, Object detection

I. 서 론

최근 발생한 코로나바이러스(이하 COVID-19)의 대 유행으로 인해 바이러스 감염자가 폭발적으로 증가하고 있으며 수많은 사상자가 발생하고 있다. 이에 세계보건기구 WHO (World Health Organization, 이하 WHO)는 팬데믹(Pandemic) 사태를 선언하였으며 전 세계적으로 유례가 없던 비상사태를 겪고 있다. 현재 알려진 COVID-19는 감염자의 비말, 콧물 가래 등 호흡기 분비물이 공기를 매개로 피감염자에게 전파된다. 그러므로 호흡기 분비물을 차단하는 마스크 착용이 감염 예방에 있어 최우선으로 여겨진다. 게다가 COVID-19의 다양한 변이 바이러스가 지속적으로 등장하고 있고 백신 접종을 마친 사람들에게도 감염되는 돌파감염(Breakthrough Infection)의 사례가 발생하고 있으며 관련 변형 호흡기계 질환까지 발생하고 있다. 그렇기 때문에 COVID-19는 물론 호흡기계 질환을 예방하기 위한 마스크 착용은 여전히 중요하다[1,2].

이에 따라 우리나라를 포함한 전 세계의 모든 국가들이 마스크 착용에 대한 권고 및 강력한 법률적 규제를 시행하고 있다. 관련하여 최근 영상 처리(Image Processing) 분야에 딥러닝(Deep Learning)기술이 발전됨에 따라 이를 이용하여 마스크 착용 여부를 검사하기 위한 데이터셋(Dataset)이 구축되었으며 그 데이터셋을 기반으로 한 선형연구들이 진행되어 왔다[3-5]. 딥러닝에서의 데이터셋은 기술을 적용하는 기본 척도이며 데이터셋을 기반으로 딥러닝 모델들이 발전되어 왔다. 그러므로 데이터셋은 그 구성이 잘못될 경우 학습 시 데이터

*Corresponding Author : Ho-Chul Kim, Ph.D.
Department of Radiological Science, Eul-Ji University 553, Sanseong-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do, Republic of Korea
Tel: +82-31-740-7412
E-mail: tiger1005@gmail.com

전이현상이 발생해 학습의 의미가 없어질 정도로 중요하다[6].

최근 연구[2, 7] 및 우리나라의 보건복지부[8]와 미국CDC[9] 따르면 단순 마스크 착용만이 아니라 올바른 마스크 착용 여부가 감염경로 차단에 매우 중요한 부분이다. 하지만 선행연구들은 마스크 착용여부[10]에 대한 판별에만 초점을 두어 마스크를 잘못 착용하였을 때의 유형을 검출하지 못하고 있으며, 마스크 착용 유형 데이터셋이 없어 관련 연구를 위한 기반이 마련되어 있지 않다.

따라서 본 연구의 목적은 COVID-19등 호흡기계 질환예방 연구의 기반을 마련하기위한 마스크 착용 유형 별 데이터셋 구축과 최적화된 검출모델을 선정하는 것에 있다.

II. 연구 방법

1. 데이터셋 구성

딥러닝 모델의 학습을 위해 데이터셋은 웹크롤러(Web Crawler)를 통해 얻은 오프소스(Open Source)이미지를 사용하였다. 해당 영상들은 다양한 각도로 촬영된 마스크 오착용 이미지 및 마스크 착용 이미지로 총 2,296장을 사용하였다. 영상 수집 시 더 많은 데이터셋을 구축하기 위하여 실제 마스크 착용 이미지 이외 마스크를 착용하지 않은 사람에 대하여 마스크를 합성한 영상을 활용하였다. 검출모델 최적화에 있어서 딥러닝 모델에서 제공하는 이미지 증강기술(Image Augmentation)을 활용하였고, 이는 데이터셋의 검증에 영향을 줄 수 있어 각 모델에서 기본적으로 학습 시 적용되는 기술인 CutOut, MixUp, CutMix을 사용하였다[11].

데이터셋은 그림 1과 같이 마스크를 잘못 착용하였을 때의 유형을 3가지로 구분하였다. 코 부분이 드러나 착용한 모습(그림 1a), 코와 입이 모두 드러난 모습(그림 1b), 코만 덮고 입이 드러난 모습(그림 1c)으로 세분화하였다. 그리고 마스크 착용유무 판별을 위한 라벨링(Labeling)을 진행하였다. 학습과 테스트에 사용될 영상의 비율은 통상적으로 사용되는 6:4로 구성하였다.

2. 학습 환경

모델 학습을 위한 환경으로 Intel i7-11700K 3.60GHz, 96GB RAM, GeForce RTX 3090 24GB, ubuntu 20.04LTS, nvidia driver 470, pytorch 1.11.0, CUDA 11.3.1을 사용하였다.

3. 딥러닝 모델

(1) 객체분류(Object Classification)

데이터셋 검증을 위한 딥러닝 모델로 객체분류(Object Classification)모델을 사용하였다. Pytorch 라이브러리에서 제공하는 AlexNet[12]와 Tensorflow에서 기본으로 제공하는 Dencenet201, InceptionV3, Mobilenetv2, ResNet50, ResNet101, VGG16, VGG19, Xception을 사용하였다. 분류 모델로 다양한 모델들을 사용하여 데이터셋의 신뢰도를 높였으며 모델의 학습에는 64 batch size에 80 epoch를 적용하였다.

(2) 객체검출(Object Detection)

객체검출모델로는 YOLO[13]를 버전별(YOLOv3[14], YOLOv4[11], YOLOv5[15], YOLOx[16])로 사용하였다. YOLO는 이미지를 영역으로 나누고 각 영역에 대한 경계 상자와 확률을 예측하여 빠른 시간에 할 수 있 수 있으며 동시에 실시간 감지까지 가능하여 검증된 데이터셋에 대해 마스크 착용 유형별 검출로 적합하다.

각 학습에 있어 최적화 기술로 컨볼루션 필터(Convolution Filters)는 15, 배치(Batch)는 64, 최대 훈련 배치 개수(Max

그림 1. 유형별 마스크 착용 모습은 코만 덮고 입이 드러난 모습(a), 코 부분이 드러나 착용한 모습 (b), 코와 입이 모두 드러난 모습(c)으로 구분하여 구성하였다

Fig. 1. The type of mask wearing is organized wearing a mask with the mouth exposed(a), with with only the nose exposed(b) and both the mouth and the nose exposed(c)

Batches)는 10,000, 학습률 조정 시점은 최대 훈련 배치 개수의 80%와 90%로 설정하였으며 해당수치는 YOLO에서 최적화로 권장되는 수치이다.

III. 연구 결과

딥러닝 모델에서는 정확도를 표현할 때 식 (1)과 같이 Recall, 식 (2)와 같이 Precision, 식 (3)과 같이 Precision과 Recall의 조화평균을 이용한 F1 Score를 정확도 및 식 (4)와 같이 Accuracy를 평가에 사용한다. 이때 사용되는 수치 중 TP(True positive) 실제로 참인데, 모델에서 예측이 참이라고 판단된 경우이며 TN(True negative)는 실제 거짓인데, 모델에서 예측이 거짓이라고 판단된 경우를 말한다.

FP(False positive)는 실제 거짓, 즉 없는 데이터이거나 거짓인데, 모델에서 잘못 예측한 것이며 FN(False negative)눈 실제로는 참인데, 모델에서 예측하지 못한 경우이다[17,18].

$$\text{recall} = \frac{TP}{TP+FN} \tag{1}$$

$$\text{precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{F1 score} = 2\frac{precision*recall}{precision+recall} \tag{3}$$

$$\text{Accuracy} = \frac{TP}{TP+TN+FP+FN} \tag{4}$$

객체검출 모델의 정확도는 추가적으로 Recall과 Precision의 성능비교 그래프를 나타낸 후 그래프의 아래의 영역의 합인 AP(Average Precision)로 성능을 표기하는데 그 값은 식 (4)과 같으며 AP를 유형별로 평균을 낸 mAP(mean Average Precision)를 사용한다.

$$AP = \sum_{k=0}^{k=n-1} [recall(k) - recall(k+1)] * presicion(k), \quad n = Number\ of\ thresholds \tag{5}$$

그리고 식 (5)와 같이 IoU의 영역으로 위치의 정확성을 평가하며 위 mAP에 대해 50%의 IoU인 mAP@50을 성능평가에 사용한다.

$$IoU = \frac{Area\ of\ overlap}{Area\ of\ union} \tag{6}$$

그림 2. 데이터셋 학습
Fig. 2. Training dataset

그림 3. 학습된 모델을 활용한 검출
Fig. 3. Detection using the trained model

위 평가지표에 따라 마스크 착용 유형별 데이터셋을 검증하기 위한 각 객체분류모델의 성능은 표 1과 같이 나타났다. 이는 모든 객체분류모델에서 일정수준 이상의 성능을 나타낸 것으로 해당 데이터셋의 신뢰도 수준이 유의미하다는 것을 확

표 1. 객체분류모델 성능비교
Table 1. Detection performance of the object classification model

	Detection Performance								
Model	AlexNet	Dencenet 201	Inception V3	Mobilenet v2	ResNet 50	ResNet 101	VGG16	VGG19	Xception
Accuracy(%)	80.95	97.68	96.58	97.31	84.74	89.74	95.48	96.34	98.66

표 2. 객체검출모델 성능비교
Table 2. Detection performance of the object detection model

Detection Performance				
Model	YOLOv3	YOLOv4	YOLOv5	YOLOX
Recall(%)	91	93	90.5	86.33
Precision(%)	91	89	94.4	85.41
F1-score(%)	91	91	92.40	85.87
mAP@50(%)	91.82	93.40	93.9	85.25

인 할 수 있다.

마스크 착용 유형별 검출모델 선정을 위해 그림 2과 같이 마스크 착용 유형별 학습을 진행하였다. 학습 후 실제 그림 3와 같이 검출하였으며 정확도중 mAP@50을 주요 성능평가지표로 사용하였다. 그 결과 구성한 데이터셋에 대해 YOLO의 버전별에 대해 각 성능이 표 2과 같이 나타났다. 결과적으로 그림 4와 같이 정확도가 나온 YOLOv5가 마스크 착용 유형별 데이터셋에 적합한 객체모델로 나타났다.

IV. 고찰 및 결론

본 연구는 이미지 데이터에 대해 마스크 착용 유형별로 나누어 데이터셋을 구성하였으며, 이 데이터셋의 검증을 위하여 여러 객체분류모델을 사용하여 학습하였다. 이 분류모델 중 최초의 CNN(Convolutional Neural Network)분류 모델인 AlexNet[12]은 컴퓨터 비전에서 딥러닝 혁명을 시작한 최초의 CNN 모델이다. 이후 2014년 Oxford의 VGG(Visual Geometry Group)에 의해 VGG19[19]가 제안되었다. VGG는 모델의 깊이를 16-19 레이어로 늘리고 이전에 AlexNet에서 사용되었던 5×5 및 7×7 대신 매우 작은(3×3) 컨볼루션 필터를 사용한다. 이후 ResNet(Residual Networks)은 이전에 사용된 것보다 훨씬 더 깊은 새로운 유형의 CNN을 사용해 개선을 이루었다. ResNet[20]은 레이어 입력을 참조하여 잔여 함수를 학습하는 것으로 레이어를 재구성하여 네트워크 교육을 용이하게 하는 것을 목표로 한다. 이후 EfficientNetB4[21]은 Depth, Width, Resolution을 적절히 조절하여 모델의 크기와 연산량을 줄이면서도 성능은 높인 모델이다. 그리고 본 연구에서 사용한 위 모델에 대해 마스크 착용 유형별 데이터셋이 표 1과 같이 유의미한 정확도를 보였다.

그리고 획득된 데이터셋에 대한 객체검출모델 선정을 진행하였다. 객체검출모델은 2015년 R. Girshick은 R-CNN 및 SPPNet을 더욱 개선한 Fast RCNN(Fast Regions with Convolutional Neuron Networks features)[22] 모델에서 비약적으로 성장하는데, 이 모델은 기존 선택 검색을 통해 일련의 객체 제안(객체 후보 상자)을 추출한 후 각 제안은 고정 크기 이미지로 다시 조정되고 ImageNet에서 훈련된 CNN 모델에 공급되어 특징을 추출한다. 이후 YOLO(You Only Look Once)[13]는 2015년 R. Joseph이 제안하였는데 “제안 탐지 + 검증”이라는 기존 객체검출 패러다임을 완전히 포기하고 대신 전체 이미지에 단일 신경망을 적용하는 완전히 다른 철학을 따르는 모델이다. 이 모델은 이미지를 영역으로 나누고 각 영역에 대한 경계 상자와 확률을 동시에 예측한다. 이후 성능을 향상시킨 YOLOv3[14]가 발표되었으며 거기에 YOLOv4[11]는 WRC(Weighted-Residual-Connections), CSP(Cross-Stage-Partial-Connections), CmBN(Cross mini-Batch Normalization), SAT(Self-Adversarial-training) 및 Mish-activation 등의 범용성을 가지는 기능을 사용하여 성능을 높였다. 거기에 더해 YOLOv5[15]는 실제 데이터 구별에 최적화하여 사용하였다. YOLO v5는 Mosaic기법의 데이터 증강기술, Drop block regularization을 사용하였으며 YOLOx[16]는 Anchor-free방식을 적용하였고, Decoupled Head, SimOTA와 같은 다양한 탐지 테크닉을 적용하였다. 이에 마스크 착용 유형별 검출 데이터셋에 대하여 YOLOv3, YOLOv4, YOLOv5, YOLOx를 이용하여 각 모델의 특징에 따라 최적화 하여 학습을 진행 하였고 그 결과 표 2와 같이 정확도가 측정되었으며 마스크 유형별 검출에 가장 적절한 모델이 YOLOv5라는 결론을 얻었다.

본 연구는 기존연구와 다르게 마스크 착용 유형별 데이터셋을 구축 및 검증하였고, 검증된 데이터셋에 적절한 객체검출모델을 선별하였다. 그리고 이 결론을 기반으로 마스크 착용 유형에 대한 새로운 딥러닝 모델 개발 및 이를 활용한 올바른 마스크 착용에 대한 선행연구로써 전염병 예방에 기여할 수 있다는 점과 연구결과 획득한 데이터셋을 활용하여 다양한 딥러닝을 학습에 활용이 가능하다는 점에서[23, 24] 본 연구의 의의가 있다.

하지만 본 연구에서 제안한 마스크 착용 유형별 판단 모델은 객체검출 모델에 종속적이기 때문에 객체검출 모델 개선한 딥러닝 모델구축에 대한 연구가 필요하며 관련하여 데이터셋 보완 및 관련 총괄 시스템 개발에 대한 추가연구가 필요하다.

References

[1] Esposito Susanna, Principi Nicola. To mask or not to mask children to overcome COVID-19. European journal of pediatrics, 2020;179.8:1267-1270.

[2] 장경철, et al. 쉴리렌 광학 기법을 활용한 COVID-19 대응 올바른 마스크 착용법에 대한 연구. 대한기계학회 춘추학술대회, 2020;599-602.

[3] Lee Se-Hoon, et al. Development of CCTV for Identification of Maskless Wearers based on Deep Learning. In: Proceedings of the Korean Society of Computer Information Conference. Korean Society of Computer Information, 2020;317-318.

[4] Cho Won-Young, et al. Development of AI Systems for Counting Visitors and Check of Wearing Masks Using Deep Learning Algorithms. In: Proceedings of the Korean Society of Computer Information Conference. Korean Society of Computer Information, 2020;285-286.

[5] 이승호. 출입 통제에 활용 가능한 딥러닝 기반 마스크 착용 판별. 한국산학기술학회 논문지, 2020;21.8:395-400.

[6] Quiñonero-candela Joaquin, et al. (ed.). Dataset shift in machine learning. Mit Press, 2008.

[7] Hui David S, et al. Exhaled air dispersion during coughing with and without wearing a surgical or N95 mask. PloS one, 2012;7.12:e50845.

[8] https://korean.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/about-face-coverings.html Accessed on 06 May 2022.

[9] https://www.cdc.gov/coronavirus/2019-ncov/downloads/How-to-TakeOff.pdf Accessed on 06 May 2022.

[10] Lee Se-Hoon, et al. Development of CCTV for Identification of Maskless Wearers based on Deep Learning. In: Proceedings of the Korean Society of Computer Information Conference. Korean Society of Computer Information, 2020;317-318.

[11] Bochkovskiy Alexey, Wang Chien-Yao, Liao Hong-Yuan Mark. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.

[12] Iandola Forrest N, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint arXiv:1602.07360, 2016.

[13] Redmon Joseph, et al. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016;779-788.

[14] Redmon Joseph, Farhadi Ali. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.

[15] Kuznetsova Anna, Maleva Tatiana, Soloviev Vladimir. Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images. In: International Symposium on Neural Networks. Springer, Cham, 2020;233-243.

[16] Ge Zheng, et al. Yolox: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430, 2021.

[17] Erhan Dumitru, et al. Scalable object detection using deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014;2147-2154.

[18] Girshick Ross, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014;580-587.

[19] Sengupta Abhronil et al. Going deeper in spiking neural networks: VGG and residual architectures. Frontiers in neuroscience, 2019;13:95.

[20] Targ Sasha, Almeida Diogo, Lyman Kevin. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv: 1603.08029, 2016.

[21] Zhang Pan, Yang Ling, Li Daoliang. EfficientNet-B4-Ranger: A novel method for greenhouse cucumber disease recognition under natural complex environment. Computers and Electronics in Agriculture, 2020;176:105652.

[22] Girshick Ross. Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision. 2015;1440-1448.

[23] 서희철, 박유민, 조영주. 신종 감염병 확산에 따른 마스크 착용효과 인지를 위한 VR 연동 “Wearing a Mask” SW 교육 제안. 한국컴퓨터정보학회 학술발표논문집, 2020;28.2: 625-628.

[24] 후쉬펑, 임현석, 곽정환. YOLO 모델 앙상블을 이용한 복잡한 장면에서의 Mask Detection 기법. 한국컴퓨터정보학회 학술발표논문집, 2022;30.1:97-98.