DBpia

모바일 환경을 위한 딥러닝 기반 열화상 이미지 마스크 인식 및 교차 엔트로피를 이용한 클래스 분류

Deep Learning-based Thermal Image Mask Recognition for Mobile Environments and Classification using A Cross-entropy Loss Function

저자 (Authors)	문종원, 윤호섭 Moon, Jong-Won, Yoon, Ho-Sub

출처 (Source)	전자공학회논문지 59(1), 2022.1, 118-123 (6 pages) Journal of the Institute of Electronics and Information Engineers 59(1), 2022.1, 118-123 (6 pages)
발행처 (Publisher)	대한전자공학회 The Institute of Electronics and Information Engineers
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11032931
APA Style	문종원, 윤호섭 (2022). 모바일 환경을 위한 딥러닝 기반 열화상 이미지 마스크 인식 및 교차 엔트로피를 이용한 클래스 분류. 전자공학회논문지, 59(1), 118-123.

이용정보 (Accessed)	한국전자통신연구원 129.***.164.24 2022/02/22 14:26 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s)and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s)for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

Journal of The Institute of Electronics and Information Engineers Vol.59, NO.1, January 2022 https://doi.org/10.5573/ieie.2022.59.1.118
ISSN 2287-5026(Print) / ISSN 2288-159X(Online)

논문 2022-59-1-13

모바일 환경을 위한 딥러닝 기반 열화상 이미지 마스크 인식 및 교차 엔트로피를 이용한 클래스 분류

(Deep Learning-based Thermal Image Mask Recognition for Mobile Environments and Classification using A Cross-entropy Loss Function)

문 종 원*, 윤 호 섭**

(Moon, Jong-Won and Yoon, Ho-Sub©)

요 약

전세계적 전염병인 COVID-19로 인하여 군중에서 열이 높은 일부 인원을 검출하기 위한 비접촉식 열화상 시스템의 수요가 증가하였다. 공공장소에서 마스크 착용이 의무화되었고 입구마다 관리 인원을 두어 출입을 통제하고 있는 상황속에서, 최소한의 인원으로 원활하게 통제하려면 카메라 등으로 관리를 해야 하는데 일반적으로 우리가 사용하는 RGB 이미지에서 마스크를 탐지하는 것은 열화상 이미지에서 탐지하는 것 보다 여러 가지 단점을 갖고 있다. 본 논문에서는 열화상 이미지 사용의 장점과 더불어 모바일 환경[1]에서도 충분히 사용이 가능한 경량화된 딥러닝 모델을 사용하여 이를 바탕으로 딥러닝 기반 열화상 이미지 마스크 인식에 대해서 소개한다. 또한 마스크 착용, 미착용의 이진 분류에서 더 나아가 턱에 걸치거나 코에 걸치는 애매한 경우까지 하나로 묶어 3가지의 형태를 분류하는 모델도 소개한다. 우리가 들고 다니는 일반적인 휴대폰에서도 실시간으로 사용할 수 있게 매우 경량화된 모델을 사용하였는데 마스크 인식 문제는 비교적 간단한 분류 문제이기 때문에 경량화된 네트워크로도 좋은 성능을 낼 수 있다.

Abstract

COVID-19, a global infectious disease, has increased the demand for contactless thermal imaging systems to detect some people with high heat in the crowd. In public places, mask-wearing is mandatory and access is restricted with management personnel at each entrance, to control smoothly with a minimum number of people, detecting masks from RGB images generally has more disadvantages than detecting them from thermal images. In addition to the advantages of using thermal images, this paper introduces deep learning-based thermal image mask detection based on using a lightweight deep learning model that can be sufficiently used in mobile environments. A very lightweight model is used for real-time use even in ordinary mobile phones, but the problem of mask detection is a relatively simple binary classification problem.

Keywords : Thermal image, Mask detection, Mask recognition, Mobile environments, Android

Ⅰ. 서 론

전 세계적 전염병인 COVID-19로 인해 군중 속에서 고열 환자를 찾아내기 위해 비접촉식 열화상 시스템의 수요가 증가하였다. 그와 더불어 공공장소에서는 마스크 착용이 의무화되었고 입구마다 관리 인원이 배치되어 마스크 착용 여부를 확인하고 있다. 공공장소에서 최소한의 인원으로 이러한 출입 통제를 원활하게 하기 위해서는 영상 촬영을 통한 관제가 필수적인데 현재는 RGB 카메라로 마스크 착용 여부를 판단하고 열화상 카

*학생회원, **비회원, 과학기술연합대학원대학교(UST) ETRI스쿨 (University of Science and Technology, Korea, ETRI School)
© Corresponding Author(E-mail : yoonhs@etri.re.kr)
Received ; October 22, 2021 Revised ; November 9, 2021 Accepted ; November 18, 2021

Copyright © The Institute of Electronics and Information Engineers.
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.

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)

메라로 온도를 재는 경우가 대부분이다. 보다 정확하게 온도를 측정하기 위해선 RGB 이미지에서 얼굴을 찾고 범위를 특정하여 열화상 이미지에서 그 부분의 온도를 가져오는 방식을 사용하고 있다. 이때 RGB 이미지로 마스크 착용 여부를 판단하면 몇 가지 단점이 존재한다. 첫째, 카메라의 구도나 조명이 인식 성능에 많은 영향을 미치게 된다. 카메라가 설치된 장소가 어둡다거나 역광의 환경일 경우 최적의 조건과 비교해서 구도나 조명이 다르므로 인식 성능이 떨어질 수 있다. 둘째, RGB 이미지에서 마스크에 대한 강인한 대처를 위해 마스크를 쓰지 않은 얼굴에 마스크를 합성해서 붙이는 방식[2, 3]의 학습데이터를 생성하는 방법을 취하는데, 이런 방식을 취하면 다양한 마스크나 스카프 등에 대해 대처하기 위해서는 굉장히 많은 전처리와 데이터가 필요하다. 셋째, 열화상 이미지와 RGB 이미지를 동시에 사용할 때는 두 카메라간의 화각과 해상도와 같은 기본 환경이 다르기 때문에 카메라간의 캘리브레이션이 필수적으로 필요하다.

그러므로 이러한 단점을 해결하기 위해 본 연구에서는 열화상 이미지 단일로 마스크 검출 및 온도 측정을 동시에 수행하는 것을 제안한다. 이 방법은 위의 두 개의 카메라를 이용했을 때 필요한 전처리 작업의 대부분을 줄일 수 있다. 더불어 단일 이미지만 사용하므로 그에 따른 장비에 대한 소요도 줄일 수 있다. 또한 얼굴 사진에 대한 대비가 가능하다. 출입 시스템을 사용하고 있다고 가정할 때, RGB 이미지를 사용하게 되면 사진(위장 얼굴)과 실제 사람의 구분이 쉽지 않아 추가적인 처리가 필요하지만, 열화상 이미지를 사용하면 위장 얼굴에 대한 대처가 충분히 가능하다. 만일 원형의 발열체를 사진 뒤에 놓고 통과하려고 해도 얼굴 영역 부분의 열 분포가 실제 사람과 다르기 때문에 쉽게 구분 가능한 장점이 있다.

대부분 열화상 카메라는 거치대 위에 올려놓고 쓰는 대형 카메라가 대부분인데, 본 논문에서는 모바일 환경에서 동작하는 모델을 소개하고 있으므로 모바일 기기에 꽂아서 사용할 수 있게 USB-C 타입의 형태로 나온 Flir사의 소형 열화상 카메라와 함께 갤럭시 탭 S6 Lite를 사용하였다.

모바일 환경에서 얼굴 검출 딥러닝 모델을 추론하기 위해서는 경량화된 네트워크의 사용이 필수인데 임베디드 기기에서도 작동이 가능할 수 있게 RetinaFace[4]를 이용하였다. 이때 RetinaFace의 모델 구조 일부를 수정하여 열화상 이미지에서는 필요성이 떨어지는 얼굴 랜드마크 검출 모듈을 마스크 분류 모듈로 바꾸어 사용했다.

본 논문은 다음과 같이 구성된다. 2장에서는 열화상 케메라 영상을 입력받아 얼굴 검출 및 마스크 및 안경 분류를 위한 딥러닝 모델, 학습 데이터 셋, 손실 함수를 소개하고, 3장에서는 입력 영상의 해상도 크기 조절, 손실함수를 이용한 실험 결과를 소개한다. 마지막으로 4장에서는 결론에서 본 연구에서 달성한 결과를 요약하고 향 후 연구 방향을 소개한다.

Ⅱ. 본 론

2.1 RetinaFace

RetinaFace는 CVPR 2020에 발표된 고성능의 얼굴 검출기이다. 2가지의 백본 모델을 사용할 수 있는데, 정확도가 높지만 계산량이 많은 iResNet[5] 기반과 고도로 경량화되어 속도가 매우 빠르지만 검출 성능은 약간 부족한 MobileNetV1[6] 기반 2가지이다. 본 논문에서는 실시간 작동을 위한 경량화에 초점이 맞추어져 있으므로 후자를 사용하였다. 경량화된 모델을 사용하면 얻는 이점이 두 가지가 있는데 첫째는 학습 수렴 속도가 굉장히 빠르다는 것이다. 학습 수렴 속도가 빠르면 여러번의 학습을 진행할 수 있는 만큼 시간을 아낄 수 있으며, 둘째는 모델 결과값의 수렴을 위해 필요한 학습 데이터셋이 보다적게 필요하다는 점이다. 이점은 데이터를 얻기 힘들고 부족한 열화상 이미지에 적합한 환경이다.

RetinaFace는 얼굴 검출은 물론 5개의 랜드마크 포인트, 검출된 영역의 이진 분류(얼굴, 배경)와 3D 얼굴 생성까지 4가지를 한 번에 수행하는 모델이다. 공개된 코드에서는 3D 얼굴 생성은 빠져있으며 본 논문에서는 나머지 3가지 중에서 5개의 랜드마크 포인트 검출하는 모듈을 마스크 인식 모듈로 바꾸어 사용했다.

열화상 이미지가 모델에 입력으로 들어가면 정규화를 진행하고 백본 네트워크에 의해 추출된 정보가 영상 내의 다양한 얼굴 크기를 대비하기 위한 FPN[7]과 얼굴을 검출하기 위한 특징을 추출하는 SSH[8]를 거쳐 얼굴 검출, 얼굴 이진 분류, 마스크 이진 분류 모듈을 통과하게 된다. 최종적으로 도출된 값에 정규화했던 값들을 다시 원래대로 복구시켜 비최대억제를 수행하여 가장 신뢰도 높은 값만을 취해 얼굴의 좌표 및 마스크 착용 여부가 결과값으로 나오게 구성하였다.

2.2 데이터셋

열화상 이미지의 데이터셋은 공개되어있는 것도 매우 드물지만 공개되어있는 데이터셋[9]으로 학습을 시켜

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)

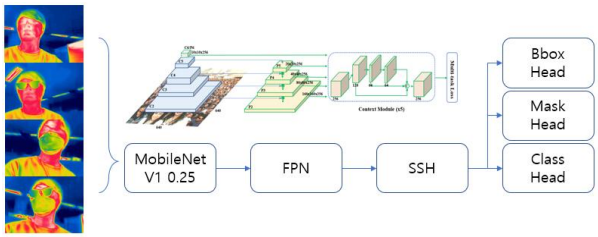

그림 1. 기존의 랜드마크 모듈이 마스크로 수정된 모델의 전체 구조도

Fig. 1. The overall structure of the existing landmark module modified with mask module.

도 해결하고자 하는 문제에 직접 적용하기가 매우 힘들다. 카메라마다 센서도 다르고 그 값을 이미지로 시각화하는 과정마저 다르기 때문이다. 따라서 열화상 이미지의 마스크 인식을 위해 데이터 셋을 직접 구축하여 사용했다. 특히 얼굴을 검출함에 있어 고려했던 사항은 안경과 마스크였다. 열화상 이미지 특성상 안경과 마스크는 피부온도의 정보를 가리기 때문에 RGB 이미지에서보다 민감한 부분이며 각각에 대해 착용, 미착용 2가지가 있으므로 총 4개의 클래스가 생성된다. 두가지 모두 미착용의 경우 2243장, 안경만 착용했을 경우 2164장, 마스크만 착용했을 경우 1438장, 둘 다 착용했을 경우 2118장, 총 7963장의 학습셋을 구축하였으며 테스트셋은 마스크 미착용과 착용의 비율을 2:3으로 나누어 구축하여 마스크 미착용 147장, 마스크 착용 221장, 총 368장으로 구축했다. 학습셋에서 마스크 미착용의 경우 걸쳐서 쓰는 애매한 형태의 데이터는 들어있지 않다. 테스트셋의 경우 보다 다른 구도를 위해 학습셋에 포함되지 않은 데이터인 턱 밑쪽으로 걸쳐서 쓴 마스크(미착용 분류)와 얼굴 근처에 발열체를 포함해 학습 데이터셋 보다 어렵게 구성하였다.

2.3 손실함수 및 마스크 모듈

본 논문에서 사용하고 있는 수정된 RetinaFace의 3

표 1. 열화상 이미지 데이터셋에 대한 정보(테스트 셋의 구분은 마스크의 여부만 고려, 단위 장)

Table 1. Information on the thermal image data set(the classification of the test set is only considering whether the mask is present or not).

	학습용	테스트용
마스크 X 안경 X	2243	147
마스크 X 안경 O	2164	
마스크 O 안경 X	1438	221
마스크 O 안경 O	2118	
합계	7963	368

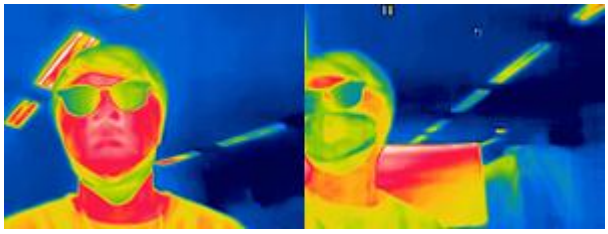

그림 2. 학습용 데이터셋에는 포함되어 있지 않은 턱 끝에 걸쳐서 쓴 것과 발열체가 얼굴 근처에 있는 테스트용 데이터셋의 일부

Fig. 2. Writing over the tip of the chin, which is not included in the training dataset, and part of the test dataset where the heating element is near the face.

가지 기능(얼굴 검출, 얼굴/배경 이진 분류, 마스크 이진 분류)에 대하여 손실함수의 합을 최종 손실함수로 사용하고 있는데 이를 정의하면 다음과 같다.

$$L = \lambda_{bbox} p_i^* L_{bbox}(t_i, t_i^*) + \lambda_{cls} L_{cls}(p_i, p_i^*) + \lambda_{mask} L_{mask}(m_i, m_i^*) \quad (1)$$

L_{bbox}는 smooth-L1 손실함수이며 t_i, t_i^*는 각각 anchor와 정답 데이터를 의미한다. smooth-L1 함수는 일반적으로 회귀에 많이 사용되며 특이치에 강인한 특징을 가지므로 좌표 검출을 하는 모듈에 적합한 손실함수이다. anchor는 사전에 미리 얼굴의 비율대로(가로 1: 세로 2) 생성해 놓은 정답 데이터 후보들이다. 추론의 결과와 미리 생성해 놓은 정답 데이터 후보가 일정 임계값을 넘어 겹쳐있다면 얼굴이 있다고 판정한다.

L_{cls}는 교차 엔트로피를 사용하며 앞에서 나온 추론 결과 anchor i 대해 대상이 배경인지 얼굴인지 검증하는 모듈이다. 교차 엔트로피 함수는 비교 대상이 되는 집단 간의 차이를 나타내는 손실 함수이며 분류 문제에 많이 사용된다. 교차 엔트로피를 시각화하면 그림 3과 같은데 교집합에 해당하는 부분을 새로운 분류로 정의

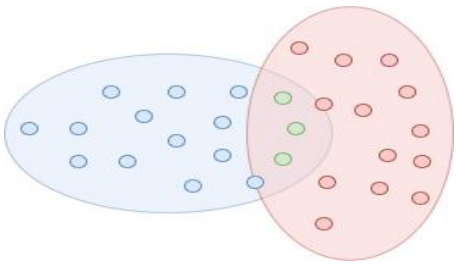

그림 3. 교차 엔트로피를 시각화 한 모습. 두 집합의 교집합 부분을 새로운 클래스로 정의할 수 있다

Fig. 3. Visualizing cross entropy. The intersection of two sets can be defined as a new class.

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)

해서 사용할 수 있다.

p_i는 anchor i의 추론 결과의 정확도이며 p_i^*는 anchor가 얼굴이면 1, 배경이면 -1이다. 추론에서 사용할 때는 A, B 값이 출력되는데 anchor 중에서 B 값을 임곗값 0.02보다 큰 값으로 1차로 후보군을 제거한 뒤에 비 최대 억제 임곗값으로 IoU(Intersection over Union) 0.4를 이용하여 겹쳐져 있는 정답 후보 중에서 최종 값을 걸러낸다.

L_{mask}도 L_{cls}와 같이 교차 엔트로피를 사용하며 m_i는 마스크 착용 시 1, 미착용 시 0이다. 기존의 모듈에서는 랜드마크 포인트 5개를 추론하는 모듈이 존재하지만 열화상 이미지에서는 랜드마크 검출이 쉽지 않으며 데이터 확보도 어렵고 마스크 착용에 대한 정보가 더 의미 있으므로 교체하여 사용하였다. 앞의 모듈에서 이미 비최대억제를 수행하여 후보들이 다 걸러졌으므로 따로 후처리 연산은 없다.

$\lambda_{bbox}, \lambda_{cls}, \lambda_{mask}$는 각각 1, 2, 1의 가중치 값을 가진다. 열화상 이미지 특성상 얼굴과 비슷한 다른 타원형의 발열체와 구분하기 위해서 세부 특징 학습을 위해 얼굴/배경 이진 분류 손실 함수에 비교적 큰 가중치를 주었다.

얼굴/배경 이진 분류 모듈과 마스크 모듈은 완전히 동일한 구조를 가지는데 이는 동일하게 이진 분류를 수행하기 때문이다. FPN 모듈을 통해 여러 스케일로 이미지의 크기를 조절한 뒤 SSH 모듈을 통과하여 각 스케일 레벨마다 특징을 추출하게 된다. 여기서 추출된 값이 손실 함수로 전달되어 학습이 진행된다. 마스크 모듈의 경우 출력 값이 0이나 1처럼 정해진 값 하나가 출력되는 것이 아니라 A, B 형식처럼 두 개의 값이 한 쌍으로 출력되는데 A의 값이 크면 미착용, B의 값이 크면 착용으로 판정한다. 얼굴/배경 이진 분류 모듈도 출력값은 2개이나 마스크 모듈과는 다른 점은 이진 분류 모델은 생성된 anchor에서 비최대억제로 데이터를 선택할 때에 쓰인다. 마스크 모듈의 값은 그 후에 최종 마스크 판정으로 쓰이기 때문에 둘의 성격이 약간 다르다고 할 수 있다.

Ⅲ. 실 험

3.1 크기 조절

모바일 환경에서는 CPU의 컴퓨팅 파워가 크지 않기 때문에 입력 이미지의 원본 크기 그대로 이용하는 것은 효율적이지 않다. 즉, 큰 해상도의 이미지를 입력으로 사용하게 되면 연산량이 많아져 실시간 사용에 어려움을 겪기 때문에 사용 환경에 따라 최적의 해상도에 맞춰서 사용해야 한다. 본 연구에서 사용한 Flir 열화상 카메라에서 출력되어 나오는 이미지의 해상도는 VGA 해상도이며 RetinaFace는 입력으로 들어오는 이미지를 정방의 형태로 맞추어 사용하기 때문에 해상도 변환없이 그대로 사용하게 되면 크기 변환이 1번 발생한다. 저사양 CPU 환경에서 큰 해상도의 이미지를 연산하는 시간과 큰 해상도의 이미지를 작은 해상도로 크기를 조절해서 연산하는 시간을 비교해보면 후자가 연산 시간에서 이득을 더 많이 볼 수 있다. 다만 해상도가 작아질수록 정확도도 낮아질 수 있으므로 모바일 환경을 고려해서 적절한 해상도를 선택해야 한다.

표 2. 열화상 이미지 해상도에 따른 수행시간과 정확도
Table 2. Performance time and accuracy according to the resolution of the thermal image.

크기	수행시간(ms)	정확도(%)
160 * 120	22.51	98.55
320 * 240	57.85	99.13
480 * 360	119.4	99.71
640 * 480	206.86	99.42

표 2의 수행시간은 영상 입력부터 다음 영상 입력까지의 시간을 의미한다. 전처리로 이미지 회전 및 크기 조절을 수행하며 그 뒤에 모델 추론이 이어진다. 열화상 카메라의 이미지 최대 출력이 8fps(125ms)임을 고려하여 이미지 크기를 조절하여 테스트 해 보았다. 애매한 착용의 경우는 학습용 데이터에는 존재하지 않으므로 정확도는 마스크의 이진 분류 정확도를 의미한다. 본 논문에서는 열화상 영상의 프레임간의 간격이 125ms임을 고려하여 정확도와 크기를 둘다 확보할 수 있는 480 * 360의 해상도를 채택하였다.

해상도와 관계없이 공통적으로 오류가 나는 이미지는 RGB 이미지에서도 번지는 이미지가 성능에 영향을 미치듯이 열화상 이미지에서도 똑같이 번진 이미지가 오류의 원인이었다. 그림 4의 왼쪽 그림은 640 * 480의 해상도에서만 마스크 클래스 분류에 오류가 있었으며 오른쪽 그림은 4가지 해상도 모두 공통된 오류를 보여주었다. 특이점은 마스크를 쓰지 않은 이미지에 대해서만 오류가 발생했다는 점이며, 얼굴 밑 부분의 일부가 잘려있거나 이미지가 번져서 마스크를 착용한 것처럼 온도분포가 변해서 오류가 생긴 것으로 추정이 된다. 그

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)

림 4의 경우는 학습용 데이터셋에는 포함 되어있지 않으며 오직 테스트용 데이터셋에만 포함 되어있다. 출입관리 등에 쓰이는 등 신뢰도에 굉장히 민감하기 때문에 성능 보고용 결과에서 그치는 것이 아닌 실사용의 경우를 고려하여 일반적인 테스트에서는 전처리를 통해 빠지게 되는 노이즈성 이미지를 테스트용으로 추가하였다.

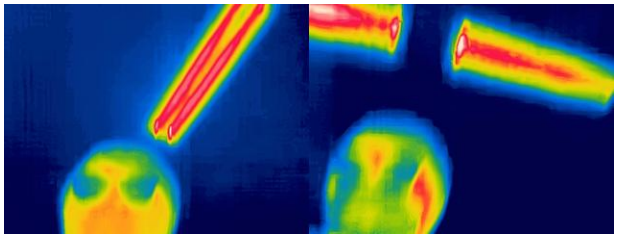

그림 4. 이미지의 해상도와 관계없이 공통적으로 일어나는 오류 영상

Fig. 4. Error images that occur in common regardless of the resolution of the image.

3.2 손실함수를 이용한 새로운 분류

턱이나 코에 살짝 걸쳐서 있는 경우는 일반적으로 미착용으로 분류되지만, 이러한 특수한 경우에는 A와 B의 값의 차이의 절댓값이 착용, 미착용의 경우보다 매우 작다. A와 B의 값의 차이의 절댓값이 보통의 경우는 3~6정도라면 특수한 경우는 1 미만으로 줄어들기 때문에 1 미만으로 차이가 나는 경우에 대해서 임계값으로 설정하여 새로운 분류를 정의할 수 있다. 마스크 모듈의 입장에서 두 값의 차이가 작다는 것의 의미는 교차 엔트로피에서 두 집합간의 차이가 작아 헷갈린다는 의미로 받아들일 수 있다. 학습은 착용, 미착용 두 분류에 대해 수행했지만 결과값을 이용하여 임계값을 조절하면 세가지 분류(착용, 미착용, 애매한 착용)를 결정할 수 있다. 임계값의 조절은 그림 3의 교집합 부분을 조절하는 역할을 한다.

그림 5. 턱에 걸쳐서 착용을 하고 있는 위쪽의 사진과 마스크를 착용/미착용한 밑의 사진. 병기된 숫자는 위쪽부터 마스크 판정의 정확도, 마스크 모듈에서 반환된 A, B값, 두 값의 차이를 나타냄. 보이는 보라색 점은 주어진 얼굴 영역에서의 최고 온도점, 파란점은 최저 온도점을 나타냄

Fig. 5. Photoes of the top show wearing over the chin and photoes of the bottom show wearing/not wearing a mask. The combined number from top represents the difference between the two values, the accuracy of mask determination and the A, B values returned from the mask module. The visible purple dot represents the highest temperature point and the blue dot represents the lowest temperature point in a given face area.

그림 5는 모델의 추론 결과를 보여주고 있다. 출력된 값을 보게 되면 마스크의 착용 여부가 확실한 경우에 A와 B 두 값의 차이의 절댓값이 커지는 경우를 볼 수 있다. 마스크 판정의 정확도는 반환된 A, B 두 개의 값의 정확도를 나타내는 지표이며 일반적으로 생각하는 마스크 착용에 대한 확률은 두 값의 차이의 절댓값에 적절한 숫자(20~25)를 곱해서 사용할 수 있다. 턱에 걸쳐서 착용하는 것이 코 밑부분 착용하는 것보다 더 낮은 값을 나타내는 것이 일반적인데 코 밑부분의 경우 올바르게 착용한 경우와 값의 차이가 크지 않기 때문에 애매한 착용의 새로운 분류를 위한 적절한 임계값 찾기가 상당히 어렵다. 임계값을 더 자세하게 조절한다고 해서 코 밑부분에 걸쳐서 착용한 마스크와 턱에 걸쳐서 착용한 마스크까지 구분할 수는 없었다. 임계값을 조절하는 것의 의미는 애매한 착용의 범위를 조절하는 것과 같은 효과를 가지고 있다. 임계값을 크게 잡을수록 영역이 넓어져 일반적인 마스크 착용도 애매한 착용으로 오인할 수 있다.

Ⅳ. 결 론

본 연구에서는 마스크를 인식할 때 RGB 이미지 대신 열화상 이미지를 사용해야 하는 이유와 모바일 환경에서도 딥러닝 추론이 가능한 얼굴 검출기와 결합된 경량화된 추론 모델을 이용하는 법을 제시하였다. 실험을 위해 공개된 열화상 이미지 데이터셋 중에 얼굴과 마스크를 집중적으로 다루는 데이터 셋이 없어 직접 데이터셋을 제작하여 사용하였는데 학습용 데이터셋에 비해 테스트용 데이터셋의 양이 적은 것은 학습용 데이터셋에는 존재하지 않는 발열체와 함께 촬영하는 등의 다른 구도의 데이터 획득이 어렵기 때문이다.

열화상 영상의 전처리 과정중 최소 1번은 크기 조절이 필요하여 안드로이드 기기의 저사양 CPU 환경에서

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)

딥러닝 추론을 할 때는 가장 알맞은 해상도를 선택해서 사용하는 것이 중요하다. 즉, 이미지에 추가적인 전처리 과정이 필요하다면 활용 목적에 알맞은 해상도를 선택하여 최대한의 프레임을 사용할 수 있는 해상도를 사용하는 것이 좋다. 본 연구에서는 해상도가 커진다고 해서 무조건 정확도가 올라가는 모습은 보여주지 않았는데 모델이 작기 때문에 특징 추출단계에서 생기는 문제일 것으로 추정을 할 수 있다.

본 연구에서는 마스크 인식을 함에 있어 교차 엔트로피 함수를 이용해서 학습 과정에서는 두 가지 분류로 학습을 했지만, 추론 과정에 있어서는 임계값을 조절해서 세 가지 분류로 나누는 것도 제시하였다. 마스크를 코와 입술 사이에 걸쳐서 착용하는 경우는 보통 2~3의 값을 나타내지만 완전히 착용했을 때는 2~6의 값을 나타내므로 명확히 임계값을 설정할 수 없기 때문에 턱에 걸쳐서 착용하는 부분은 임계값 조절로 충분히 검출해 낼 수 있지만, 코와 입술사이에 걸쳐서 착용하는 경우는 임계값 조절을 수행해도 검출해 낼 수 없었다.

향후 모델 개선을 통해 고해상도 이미지를 사용할 때 속도 개선과, 턱에 걸친 경우에 더해 코와 입술 사이에 걸쳐서 착용하는 경우도 검출해내는 것도 기대할 수 있다.

Acknowledgement

이 성과는 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2021-0-00172, 감시카메라 기반 마스크 착용자 얼굴인식 및 재인식(Re-ID) 기술 개발)

REFERENCES

[1] Lee FF, Chen F, Liu J. Infrared thermal imaging system on a mobile phone. Sensors. 2015 May; 15(5): 10166-79.
[2] CHEN, Ying, et al. Mask-based data augmentation for semi-supervised semantic segmentation. arXiv preprint arXiv:2101.10156, 2021.
[3] CHEN, Pengguang, et al. Gridmask data augmentation. arXiv preprint arXiv:2001.04086, 2020.
[4] DENG, Jiankang, et al. Retinaface: Single-shot multi-level face localisation in the wild. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. p. 5203-5212.
[5] LIANG, Zhengfa, et al. Learning for disparity estimation through feature constancy. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. p. 2811-2820.
[6] KRIŠTO, M.; IVAŠIĆ-KOS, Marina. Thermal imaging dataset for person detection. In: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2019. p. 1126-1131.
[7] HOWARD, Andrew G., et al. Mobilenets: Efficient convolutional neural networks for mobile vision application. arXiv preprint arXiv:1704.04861, 2017.
[8] LIN, Tsung-Yi, et al. Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 2117-2125
[9] NAJIBI, Mahyar, et al. Ssh: Single stage headless face detector. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 4875-4884.

저 자 소 개

문 종 원(학생회원)
2020년 한성대학교 컴퓨터공학 학사 졸업.
2020년 UST ETRI 스쿨 컴퓨터 소프트웨어 석박사통합 과정 입학.
<주관심분야 : 컴퓨터, 영상처리, 딥러닝, 얼굴인식, 안드로이드>

윤 호 섭(비회원)
1989년 숭실대학교 전산학 학사 졸업.
1991년 숭실대학교 전산학 석사 졸업.
2004년 한국과학기술원 전산학 박사 졸업.
<주관심분야 : 영상처리, 얼굴인식, 딥러닝, 연령인식, 성별인식, 얼굴 3D 모델링>

한국전자통신연구원 | IP:129.***.164.24 | Accessed 2022/02/22 14:26(KST)