

YOLO 모델 앙상블을 이용한 복잡한 장면에서의 Mask Detection 기법

후취펑^o, 임현석*, 곽정환(교신저자)*

^o한국교통대학교 소프트웨어학과,

*한국교통대학교 소프트웨어학과

e-mail: hxufeng66@gmail.com^o, {hyunseoki, jgwak}@ut.ac.kr*

Mask detection in complex scenes using an ensemble of YOLO models

Xufeng Hu^o, Hyunseok Lim*, Jeonghwan Gwak(Corresponding Author)*

^oDept. of Software, Korea National University of Transportation,

*Dept. of Software, Korea National University of Transportation

● 요약 ●

코로나바이러스-19 팬데믹 이후 매일 수만 명의 환자가 발생하고 있다. 보건당국은 사람들의 생활 안전을 보호하기 위해 공항, 정류장 등 공공장소에서는 반드시 마스크를 착용하라고 지시하고 있다. 마스크를 착용하는 목적은 감염으로부터 신체를 보호하고 바이러스 전파와 확산을 막기 위한 것이다. 공공장소에서는 많은 인원에게 대한 일괄적인 마스크 착용 검사를 하기 어렵고, 육안으로 확인하는 마스크 착용 검사 방법은 인파가 몰리는 장소에서 검사 효율이 떨어지며 누락되는 경우도 많이 발생한다. 본 연구에서는 입력 이미지에 존재하는 얼굴 영역을 YOLOv4와 YOLOv5 모델을 통해 예측하여 마스크의 착용 여부를 판단하되, 앙상블 기법을 적용하여 보다 효과적인 BB(Bounding Box) 추출 및 마스크 착용 탐지 기법을 적용한다. 따라서 공공장소의 마스크 착용실태를 효과적으로 모니터링 할 수 있는 방법을 제안한다.

키워드: 코로나바이러스-19 (COVID-19), 인공지능(Artificial intelligence), 딥러닝 (Deep learning), YOLO, Mask Detection

I. Introduction

코로나바이러스-19[1]의 세계적 확산으로 인해 어느 나라나 공공장소에서 마스크를 착용해야 한다. 하지만 일부 사람들은 공공장소에서 마스크를 쓰지 않거나 제대로 쓰지 않는 경우가 많다. 인력을 투입해 마스크 착용 상태를 확인하게 되면 엄청난 인력이 낭비되고 모든 공공장소를 확인하는 것도 비효율적이다. 공공장소에서의 검사는 검사 시간과 효율성이라는 두 가지 문제를 극복해야 한다.

본 논문은 YOLO[2] 알고리즘을 이용하여 복잡한 환경에서 마스크 착용 여부를 확인하는 새로운 방식을 제시했다. 보통 한 사람 또는 두 사람일 경우 식별이 쉽지만 사람이 붐비는 복잡한 환경에서 이미지 전체를 대상으로 하는 검사는 쉽지 않다. 본 연구에서 제시된 검사 알고리즘은 장면의 복잡도가 높은 경우에 대해 성능이 뛰어난 것을 확인하였다.

II. Related works

최근 몇 년 동안 목표 검사 알고리즘은 매우 큰 진전을 이루었다. Region Proposal 기반 R-CNN 알고리즘(R-CNN[3], Fast R-CNN[4], Faster R-CNN[5])은 2-스테이지 (Two-stage) 방식으로 작동된다. 먼저 Selective search[7], CNN Network, RPN(Region Proposal Network)[8]을 사용하여 Region Proposal을 만든 후 분류 및 회귀한다. 또 다른 부류로는 YOLO: You Only Look Once[2]가 있으며 단 한 번의 CNN 연산으로 End-to-End 예측을 제공하며 더 빠른 실시간 성능을 보여준다. 하나의 개별 End-to-End 네트워크에 기반하여 원본 이미지의 입력과 객체 위치, 클래스 출력은 CNN 네트워크를 사용하여 서로 다른 목표의 카테고리 위치를 직접 예측한다. Faster R-CNN는 정확도는 높고 속도는 느리지만, YOLO는 속도는 빠르지만 정확성은 떨어지는 특성이 있다. YOLO는 첫 출시 이후 지금까지 계속해서 버전이 향상되어왔으며 현재는 객체 감지 속도 면에서 월등한 성능을 보이는 YOLOv5[6]가 발표되었다.

III. The Proposed Scheme

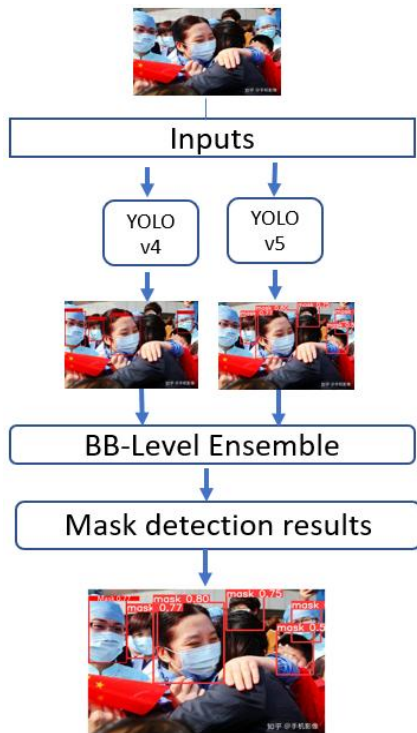


Fig. 1. System Architecture

본 논문은 복잡한 장면에서의 YOLO 모델들의 앙상블을 이용한 마스크 검출 기법을 제시한다. Fig 1은 YOLOv4[9] 모델과 YOLOv5 모델을 이용하여 출력한 BBox(Bounding box) 결과를 병합하여 검출하는 BB-level Ensemble 기법이 도시되어 있다. YOLOv4와 YOLOv5의 예측 결과를 결합하여 만약 같은 개체가 두 알고리즘에서 모두 검출된다면 두 모델이 출력한 BBox를 비교하여 BBox 면적이 적은 것을 최종 예측 결과로 선정한다. 만약 같은 개체가 두 알고리즘 중 하나의 모델에서만 감지되고 다른 모델에서 감지되지 않는다면 감지된 모델이 출력한 BBox를 출력 결과로 선정한다. Fig. 2는 복잡한 환경에서의 마스크 검출 실험 결과를 보여주며, BB-level 모델 Ensemble 기법이 기존 대비 성능이 우수함을 보여준다.

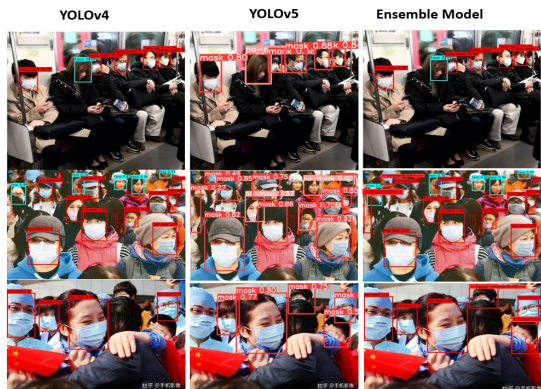


Fig. 2. Results of the Ensemble Model

IV. Conclusions

본 논문에서는 YOLOv4와 YOLOv5에 기반한 앙상블 마스크 탐지 알고리즘 장면 복잡도에 따른 적응적 얼굴 감지 시스템을 제안하였다. 실험 결과 제안한 BB-level 모델 Ensemble 방법은 사용하지 않은 개별 YOLOv4 및 YOLOv5 모델보다 높은 성능을 나타냄을 보여주었다. 제안한 시스템을 통해 여러 공공장소에서 사람들이 마스크를 제대로 잘 착용했는지를 높은 정확도로 실시간으로 감지할 수 있을 것으로 기대한다.

ACKNOWLEDGEMENT

This results was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE) (2021RIS-001).

REFERENCES

- [1] W. H. Organization et al. "Coronavirus disease 2019 (COVID-19)weekly epidemiological update - 29 december 2020," 2020.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, 2016.
- [3] R. Girshick J. Donahue T. Darrell J. Malik "Rich feature hierarchies for accurate object detection and semantic segmentation tech",arXiv:1311.2524v5 [cs.CV] 22 Oct 2014.
- [4] R. Girshick, "Fast R-CNN," In Proceedings of the IEEE International Conference on Computer Vision, pp. 1440-1448, 2015.
- [5] R. Shaoqing, K. Kaiming, R. Girshick, and S. Jian, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", arXiv:1506.01497v3 [cs.CV] 6 Jan 2016.
- [6] G. Jocher," YOLOv5" ,oct ,2020.
- [7] J.R.R Uojings, T. Gevers,"Selective Search for Object Recognition",2012.
- [8] B. Li, Y. Junjie, W.Wei, Z. Zheng, H.Xiaolin, "High Performance Visual Tracking with Siamese Region Proposal Network"2015.
- [9] A. Bochkovskiy, W.ChienYao, and H.Yuan M.Liao, "Yolov4: Optimal speed and accuracy of object detection", arXiv:2004.10934,2020.
- [10] Google Brain, Toronto, "When Does Label Smoothing Help?",arXiv:1906.02629v3 [cs.LG] 10 Jun 2020.