논문 2023-1-7 http://dx.doi.org/10.29056/jsav.2023.3.07

임베디드 시스템에 인공지능을 적용한 안전모 인식

김현아*, 이규대*†

Automatic safety helmet detection with Machine Learning

Hyun-A Kim*, Kyu-Tea Lee*†

요 약

산업별 재해상황 통계로 나타나는 작업현장에서의 안전사고는 추락, 넘어짐 등의 원인이 50% 이상이며, 이에 따른 두부 손상을 방지하기 위한 안전모 착용이 요구되고 있다. 그러나 작업현장의 특성상 안전모의 미착용 및 작업의 편의성에 따라 작업모의 불완전 착용 등으로 안전사고에 노출되고 있다. 현재 작업 감독자가 현장을 순찰하면서, 작업모의 착용 여부 및 안전상황을 점검하는 제도가 시행되고 있지만, 계속적인 작업 감독의 어려움이 있다. 이러한 환경에서 안전사고를 예방하기 위해서는 작업현장에서 개인보호구 착용을 시스템에 의한 자동 관리 감독시스템을 구축하여, 감시와 경고 신호를 발생하는 시스템이 요구된다. 본 연구에서는 라즈베리파이에 인공지능 모델을 포팅한 안전모 착용 여부 감지 시스템을 개발하였다. 객체 검출(Object Detection) 컴퓨터 비전 기술을 이용하여 안전모 착용 여부를 탐지하는 모델을 구성하고, 실시간 카메라 영상 분석을 통해 데이터를 입력받은 후, 안전모 미착용이 탐지되는 경우 경고음이 발생하는 시스템을 구축하였다. 이동체에 탑재된 실시간 영상을 통한 검출 정확도는 87%를 달성하였으며 TensorFlow lite 파일로 변환하는 과정을 통해 기존의 TensorFlow 모델보다 20% 향상된 fps 성능 결과를 보였다.

Abstract

More than 50% of safety accidents at work sites, by the statistics of accident situations in the industry, have been caused by falling off and falling down. Accordingly, it is required to wear a safety helmet to prevent head injury. Currently, a system in which work supervisors patrol the site and check whether the helmet is worn and the safety situation is being implemented. However, it is difficult to continuously supervise work. In order to prevent safety accidents, it is necessary to establish an automatic management and supervision system by system for wearing personal protective equipment at the work site. Also a system is required to generate warning signals. In this study, a safety helmet-wearing detection system was developed by porting an artificial intelligence model to Raspberry Pi. A model was constructed to detect whether or not a helmet was worn using object detection computer vision technology. The system generates a warning sound when not wearing a helmet is detected. The system was mounted on a mobile body and analyzed images. The detection accuracy was achieved 87%, and through the process of converting to TensorFlow lite file, fps performance improved by 20% compared to the existing TensorFlow model.

한글키워드 : 안전사고, 안전모, 객체검출, 모델학습, 텐서플로우

keywords : safety accident, safety helmet, object detection, model train, tensorflow

* 공주대학교 스마트정보기술공학과
† 교신저자: 이규대(email: ktlee@kongju.ac.kr)

접수일자: 2023.03.06. 심사완료: 2023.03.16.
게재확정: 2023.03.20.

1. 서 론

건설 현장에서 추락, 두부 손상 등의 재해사고를 예방하기 위해 안전모 착용 관리가 필수적이다. 하지만 최근 고용노동부에 따르면 작업 현장 3곳 중 1곳에서 안전 보호구 착용이 미흡한 것으로 나타났다[1]. 현실적으로 안전 관리자가 현장에서 작업자를 상시 감독, 관리하는 것은 어려우며, 특히 중소규모의 사업장의 경우 현장점검 안전관리 인력과 비용 부담으로 안전모 착용과 감시가 이루어지지 않고 있다[2].

본 연구에서는 건설, 제조 등과 같은 작업 현장의 안전모 착용 점검 인력 부족 문제를 해결하기 위해 머신러닝 기반 자동 안전모 착용 탐지 시스템을 구성하였다. 안전모 이미지 데이터를 수집하여 SSD(Single Shot multi-box Detector) 모델을 전이 학습(Transfer Learning)하였다. 전이 학습된 모델은 52% mAP(mean Average Precision)의 성능 결과를 보였다. 훈련된 모델을 라즈베리파이에 임베디드 장치에 포팅한 후, 카메라를 통한 입력 영상에서 안전모 착용을 탐지 성능을 실험하여, 87%의 검출 정확도를 확인하였다. 또한 안전모 미착용이 감지될 경우 임베디드 보드 시스템에서 경고음과 LED로 경고 알림을 출력하는 실시간 시스템을 구성하였다.

2. 모델 학습

2.1 객체 검출

객체 검출(Object Detection)은 컴퓨터 비전 분야에서 중요한 기술로 이미지에서 객체의 위치를 찾고, 찾은 객체에 레이블을 지정하는 알고리즘이다[3]. Object detection(객체 검출)은 입력 영상을 대상으로, 영상 내에 존재하는 모든 영역에 대해서 분류(classification)와 지역화(localization)를 수행하는 작업이다. 여기서 지역화(Localization)는 bounding box를 찾는 회귀(regression)이고, 분류(classification)는 bounding box 내 어떤 물체가 존재하는지 분류하는 작업이다.

object detection = classification + localization

따라서 처리하려는 입력 영상에 따라 존재하는 물체의 개수가 다르고, 물체의 종류가 다양하기 때문에 객체 검출은 상황에 따른 변수가 많은 작업이다. 객체검출과 유사한 기법은 semantic segmentation 이 있다. semantic segmentation은 각 픽셀 마다 클래스를 정해주는 것으로, 각 픽셀마다 클래스 예측이 가능한 이유로 dense prediction이라고도 한다.

여러 객체 검출 알고리즘 중 SSD(Single Shot multi-box Detector)는 속도가 가장 빠른 방법이다. 객체 분류와 경계 박스를 동시에 예측하는 1-stage 방식으로 컨볼루션 레이어마다 경계 박스를 생성하여 작은 객체까지 검출할 수 있다. SSD의 기본 구조는 그림 1과 같다[4].

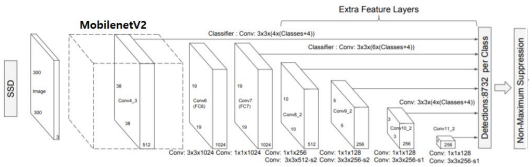

그림 1. SSD_MobilenetV2 structure
Fig. 1. SSD_MobilenetV2 structure

SSD 는 2016년 11월 말에 발표되었으며 물체 감지 작업에 대한 성능 및 정밀도 측면에서 74% 이상의 mAP(평균 정밀도) 실험 결과를 보이고 있다. 여기서 Single Shot 은 개체 현지화 및 분류 작업이 네트워크의 단일 순방향 패스 에서 수행되는 것이고, MultiBox는 바운딩 박스 회귀 기술을 의미한다. 또한 감지기는 네트워크

에서 감지된 객체를 분류하는 기능이다. 그림의 다이어그램과 같이 SSD의 아키텍처는 VGG-16 아키텍처를 기반하며, 기본 네트워크로 사용되어, 고품질 이미지 분류 작업에서 우수한 성능과 전이 학습 결과 개선에 활용된다. 기존의 VGG 연결 레이어 대신 보조 컨볼루션 레이어 세트 (conv6)가 추가 되어 여러 스케일에서 추출된 정보를 후속 레이어의 입력 크기가 점차적으로 감소되는 효과를 나타낸다. SSD의 바운딩 박스 회귀 기술은 바운딩 박스 좌표추출 방법인 MultiBox 작업으로, 클래스에 제한 없는 기술이다. MultiBox에는 Inception 스타일의 컨볼루션 네트워크가 적용되었다. 즉, 1x1 컨볼루션은 "너비"와 "높이"는 동일하게 유지하면서 차원 감소 효과를 준다.

본 연구에서는 모바일 환경에서 향상된 속도와 성능을 위해 SSD_MobilenetV2 모델을 사용하여 학습을 진행하였다. 이는 SSD의 CNN 기본 네트워크를 MobilenetV2로 사용한 모델로, 기본 네트워크로 VGG-16을 사용한 SSD보다 필요한 연산과 파라미터의 수를 줄인 알고리즘이다. 모바일 디바이스에 최적화되어 하드웨어의 성능 제한을 극복한 모델로 fps 와 성능이 향상된 구조이다.

2.2 학습

안전모 착용 이미지 분석을 위한 데이터는 구글에서 제공되는 사전 훈련 모델을 사용하는 전이 학습(Transfer Learning)을 진행하였다. 전이 학습은 사전에 학습된 가중치를 초기 가중치로 이용하여 원하는 모델의 가중치를 얻도록 학습하는 것으로, 전이 학습에 필요한 이미지와 어노테이션은 구글 kaggle에서 수집하였다[5]. 안전모의 전, 후, 좌, 우를 모두 고려하고 다양한 색상의 안전모 이미지를 데이터 세트로 구성하였다. 약 5,000여 장의 데이터를 train set 80%와 test set 20%로 나누어 학습 훈련에 사용하였다.

2.3 학습평가

학습 평가의 정확도를 나타내는 지표인 mAP(mean Average Precision)는 Faster R-CNN, SSD와 같은 object detector의 정확도를 측정하는 평가지표로 사용된다. mAP는 다음과 같이 precision(정밀도), recall(재현율), IoU(intersection of union)의 의미를 포함한다. 즉 정밀도(Precision)는 모델이 True라고 예측한 결과에서 정답도 True인 것과의 비를 나타낸다. 재현율(recall)은 실제 값이 True인 것 중에서 모델이 True라고 예측한 것과의 비를 나타낸다. 정밀도와 재현율의 계산은 다음과 같다. 수식(1)과 같다[6].

$$preision = \frac{TP}{TP+FP} \ , \ recall = \frac{TP}{TP+FN} \quad -- \quad (1)$$

TP: True positivve
TN: True negative
FP: False positive
FN: False negative

정확도와 재현율을 고려한 종합적인 평가 지표가 AP이다. 각 클래스에 대해 AP를 계산한 후 그 평균을 구하는 것이 mAP(mean Average Precision) 값이다. 실험에서 학습데이터로 사용된 이미지, 각 객체 클래스에 대한 AP는 그림 2와 같으며 학습 결과 모델은 52% mAP를 가진다. 안전모 미착용으로 분류되는 person과 head의 클래스는 helmet 클래스에 비해 낮은 AP 성능을 보였다.

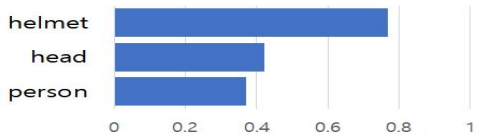

그림 2. 클래스별 AP 평가 결과
Fig. 2 Precision AP by Class

3. 임베디드시스템 적용

영국의 라즈베리 파이(Raspberry Pi) 재단에서 2012년 3월 출시된 저가의 컴퓨터 기능으로 제작된 라즈베리파이 보드는 GPIO 40개로, 다양한 센서와 카메라 모듈을 연결을 할 수 있는 하드웨어 오픈소스 플랫폼이다. 1.5GHz의 쿼드코어 64-bit ARM Cortex-A72 CPU가 적용된 브로드컴 BCM2711 SoC를 사용하고, 라즈베리파이 4B 모델 경우 USB3.0 포트 2개, USB2.0 포트 2개가 있어, USB2.0보다 10배 정도 고속이다.

임베디드 시스템으로 사용되는 라즈베리파이 보드는 자체적인 운영체제를 설치가능하고, 내부에서 프로그래밍 가능한 특징을 갖는다. 기존에 많이 사용되던 아두이노와 달리 모니터, 키보드, 마우스 등이 연결되어 개인용 컴퓨터와 유사하게 활용이 가능하며, 통신 인터페이스로 이더넷, HDMI, USB를 지원하는 싱글보드 시스템이다.

전이 학습으로 재구성된 모델을 라즈베리파이에 포팅하여 사용하도록 한다. TensorFlow lite 모델의 라즈베리파이 포팅과 입출력 구성은 그림 3와 같다. 카메라를 통해 실시간으로 영상을 입력받아 OpenCV를 이용해 이미지가 처리된 후 모델의 가중치 연산이 수행된다. 영상 이미지의 연산 처리 결과 미착용으로 예측이 된 경우 1초 간격으로 경고음이 발생하고, 빨간색 LED가 GPIO핀을 통해 출력된다.

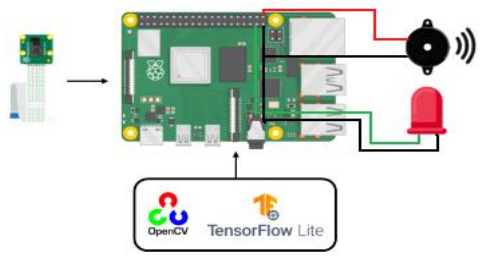

그림 3. Raspberry pi architecture
Fig. 3 Raspberry pi architecture

TensorFlow Lite란 TensorFlow 모델의 경량화 모델로 실수형 변수 단위를 정수형으로 바꾸어 연산의 속도를 개선한다. Tensorflow API를 이용하여 50,000번의 전이 학습을 진행하였다[4]. 그림 4는 훈련 스텝 마다 총 손실을 나타내는 그래프이다. 손실이 감소하며 훈련이 잘 진행됨을 보여준다.

훈련을 진행 한 후 생성된 TensorFlow 모델을 TensorFlow Lite 모델로 변환하였다. TensorFlow와 Tensorflow Lite의 성능을 비교하였을 때 20%의 fps 성능향상을 보였다.

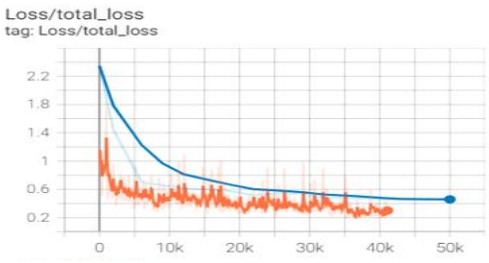

그림 4. 총 손실
Fig.4 Total loss

그림 5는 안전모 인식 결과를 box로 표현한 시스템의 화면을 나타낸다. (a)는 단일 객체 인식 (b)는 다수의 객체가 검출 되는 것을 확인 할 수 있다. 실험 결과 안전모를 착용하였을 경우 bounding box와 helmet 라벨이 화면에 출력된다.

영상인식의 활용실험은 그림 6과 같은 이동장치에 실장하여 측정되었다. 이동체는 실제로 사용하게 되는 실내·외 주행 환경과 바닥과의 마찰을 고려하여 오프 로드용 바퀴를 사용하였다. 초기 자신의 위치를 설정하고 목적지를 입력하게 되면 장애물들의 정보와 Navigation 패키지 내에 있는 Local planner, Global planner, move_base 등의 설정한 Parameter에 따라서 설정한 목적지로 이동하도록 하였고, 이동과정에서 안전모 착용작업자와 미착용 작업자의 상황을 설정하여

실시간 카메라로 입력 처리하였다.

(a) 단일 객체 (a) single object

(b) 다수 객체 (b) Multi objects

그림 5. 실험 결과 화면
Fig. 5 Test result photo

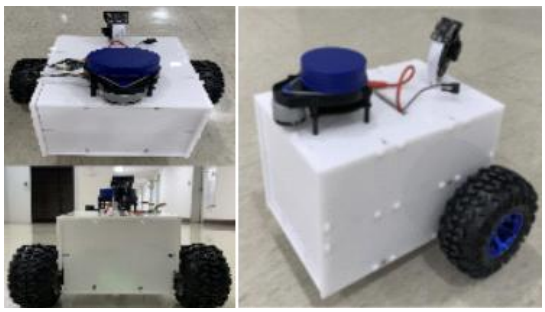

그림 6. 이동체의 전·후면 및 좌측
Fig. 6 Picture of Mobile car

실험은 실시간 카메라를 통해 100번의 얼굴을 입력하였고 결과는 표 1과 같다. 라벨과 박스 위치 모두 맞은 경우(TP) 87회, 오검출(FP) 11회, 미검출(FN) 12회의 결과로 검출 정확도는 87%이다.

표 1. 검출 정확도 평가
Table 1. Evaluation of precision

모델 평가	횟수
TP	87
FP	11
FN	12

4. 결 론

본 연구에서는 딥러닝 객체 검출 컴퓨터 비전을 이용한 자동 안전모 착용 탐지 시스템을 제안하였다. 모바일 기기에 적합한 SSD_MobilenetV2 모델을 전이 학습하고 가중치를 라즈베리파이에 포팅하였다. 딥러닝 기반으로 안전모 착용을 감지하고 라즈베리 파이의 GPIO를 이용한 알림음 발생으로 적극적인 감시 및 경고를 가능하게 하였다. 시스템의 정확도는 87%이며 모바일 기기에 적합한 TensorFlow Lite 모델을 통해 기존의 TensorFlow 모델 보다 20% 향상된 fps를 보였다. 추후 데이터 증강 등의 기법을 통해 낮은 AP 성능의 클래스의 검출 성능 향상을 위한 연구가 계속될 예정이다. 본 연구를 통해 작업 현장에서 안전모 미착용을 적극적으로 감시하고 착용을 권장시켜 안전 재해 사고를 줄이는 데 활용 가능할 것으로 기대된다.

참 고 문 헌

[1] https://www.joongang.co.kr/article/24108646#home “Construction site safety hat”, 2021.07
[2] https://m.kmib.co.kr/view.asp?arcid=0924201448, “Construction site safety hat”, 2022.01
[3] Rajalingappaa Shanmugamani, “Deep Learning for Compuer Vision : Expert

techniques to train advanced neural networks using Tensorflow and Keras", Packt Publishing, pp. 147-175, August 2018.[ISBN-10 : 1788295625]
[4] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy,Scott Reed, Cheng-Yang Fu, Alexander C. Berg, "SSD: Single Shot MultiBox Detector", 2016. [doi.org/10.1007/978-3-319-46448-0_2]
[5] Zheng Zeyu, Gu Siyu, Woo jin Jang, Google Deep Learning Framework: TensorFlow Practice, 2018 [ISBN: 9781491980453]
[6] Rowel Atienza, "Advanced Deep Learning with TensorFlow2 and Keras - Second Edition", Packt Publishing, 2020.[ISBN-10 : 1838821651]
[7] Mark Sandler Andrew Howard Menglong Zhu Andrey Zhmoginov Liang-Chieh Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", pp. 1-2, March 2019. [https://doi.org/10.48550/arXiv.1801.04381]
[8] Gu, J.; Lan, C.; Chen, W.; Han, H. Joint Pedestrian and Body Part Detection via Semantic Relationship Learning. Appl. Sci. 2019, 9, 752.[doi.org/10.3390/app9040752]
[9] Gao, H.; Chen, S.; Zhang, Z. Parts Semantic Segmentation Aware Representation Learning for Person Re-Identification. Appl. Sci. 2019, 9, 1239.[doi.org/10.3390/app9061239]
[10] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234 - 241. [doi.org/10.48550/arXiv.1505.04597]
[11] Bouindour, S.; Snoussi, H.; Hittawe, M.M.; Tazi, N.; Wang, T. An On-Line and Adaptive Method for Detecting Abnormal Events in Videos Using Spatio-Temporal ConvNet. Appl. Sci. 2019, 9, 757. [doi.org/10.3390/app9040757]
[12] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 2015, 28, 91 - 99. [doi.org/10.48550/arXiv.1506.01497]
[13] Radac, M.B.; Precup, R.E. Data-Driven Model-Free Tracking Reinforcement Learning Control with VRFT-based Adaptive Actor-Critic. Appl. Sci. 2019, 9, 1807. [doi.org/10.3390/app9091807]
[14] Wei, C.; Ni, F.; Chen, X. Obtaining Human Experience for Intelligent Dredger Control: A Reinforcement Learning Approach. Appl. Sci. 2019, 9, 1769. [doi.org/10.3390/app9091769]
[15] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends perspectives and prospects", Science, vol. 349, no. 6245, pp. 255-260, 2015. [DOI: 10.1126/science.aaa8415]
[16] A. Ng, "Machine learning yearning: Technical strategy for ai engineers in the era of deep learning", 2019.
[17] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision", Proc. IEEE Int. Symp. Circuits Syst., pp. 253-256, May/Jun. 2010. [DOI: 10.1109/ISCAS.2010.5537907]

저 자 소 개

김현아(Hyun-A Kim)

2019.3 - 공주대학교 스마트정보기술공학과 학부생
<주관심분야> 인공지능, 영상인식

이규대(Kyu-Tae Lee)

1991 고려대 전자공학과 박사
1992~국립공주대학교 정보통신공학부 교수
<주관심분야> 신호처리, VLC, 저작권보호, 임베디드 시스템, 상황인식 및 학습