

제조업 근로자 안전관리를 위한 데이터셋 구축과 모델 학습

이태준¹ · 김윤정² · 정희경^{3*}

Dataset Construction and Model Learning for Manufacturing Worker Safety Management

Taejun Lee¹ · Yunjeong Kim² · Hoekyung Jung^{3*}

¹Graduate Student, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea

²Principal Researcher, KISTI, 245 Daehak-ro, Yuseong-gu, Daejeon, 34141 Korea

^{3*}Professor, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea

요약

최근 “중대재해 등에 관한 법률”이 제정되고 안전사고에 관한 제도적, 사회적 관심이 높아지고 있다. 본 논문에서는 제조업 현장에서 발생한 안전사고에 대해 정부 기관에서 발간한 통계 자료를 분석하고, 안전사고 발생을 줄이기 위해 위험 상황을 판정하는 모델을 구축하기 위한 딥러닝 기반에 다양한 객체 탐지 모델을 학습시켜 비교 분석하였다. 제조업 현장에 있는 CCTV에서 영상을 수집하여 직접 데이터셋을 구축하였으며, YOLO-v4, SSD, CenterNet 모델에 훈련 데이터와 검증 데이터로 이를 활용하고 학습을 진행하였다. 그 결과 YOLO-v4 모델이 mAP 81%의 수치를 얻었다. 산업 현장에서 클래스를 선정하고 데이터셋을 직접 구축하여 모델 학습을 하는 데 의의가 있으며 이를 통해 위험 상황을 판정하고 이를 추론하는 시스템의 초기 연구자료로 활용할 수 있을 것으로 사료된다.

ABSTRACT

Recently, the “Act of Serious Disasters, etc” was enacted and institutional and social interest in safety accidents is increasing. In this paper, we analyze statistical data published by government agency on safety accidents that occur in manufacturing sites, and compare various object detection models based on deep learning to build a model to determine dangerous situations to reduce the occurrence of safety accidents. The data-set was directly constructed by collecting images from CCTVs at the manufacturing site, and the YOLO-v4, SSD, CenterNet models were used as training data and evaluation data for learning. As a result, the YOLO-v4 model obtained a value of 81% of mAP. It is meaningful to select a class in an industrial field and directly build a dataset to learn a model, and it is thought that it can be used as an initial research data for a system that determines a risk situation and infers it.

키워드 : CenterNet, SSD, YOLO, 객체 탐지, 산업 안전

Keywords : CenterNet, Industrial safety, Object detection, SSD, YOLO

Received 3 June 2021, Revised 4 June 2021, Accepted 22 June 2021

* Corresponding Author Hoekyung Jung(E-mail:hkjung@pcu.ac.kr, Tel:+82-42-520-5640)

Professor, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea

Open Access <http://doi.org/10.6109/jkiice.2021.25.7.890>

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

산업재해란 작업 근로자가 작업이나 업무를 하던 중 이와 관계된 사건에 의해 사망, 부상을 겪거나 질환을 얻는 것을 말한다. 2021년 정부에서는 “중대재해 처벌 등에 관한 법률”을 제정하여 산업재해를 방지하기 위한 제도적 장치를 강화하는 추세이다. 제도적 장치는 이미 일부 사업장에서 안전관리 책임자를 두도록 법으로 강제하였으나 법의 사각지대에 있는 소규모 사업장에서 높은 비율로 사고가 발생하고 있다.

한편 대부분 사업장에는 도난 방지나 사고 방지를 위해 CCTV를 설치하여 운영하고 있으며, 이 데이터를 이용해 딥러닝 모델 학습에 활용하여 안전사고를 방지하는 기술이 활발히 연구되고 있다[1, 2]. 이러한 딥러닝 모델은 먼저 작업자, 안전모, 방진 마스크, 작업자가 사용하는 장비가 검출하는 모델이어야 하며 이러한 사물을 검출하는 분야를 객체 탐지(object detection)라고 한다. 학습 데이터셋으로는 대표적으로 Microsoft 사의 COCO(Common Object in Context)나 Cityscapes 등이 쓰인다. 이는 데이터의 수가 많고 클래스가 다양하지만 제공하려고 하는 서비스와 관련된 학습 데이터셋이 쓰여야 하므로 제조업 현장에서 안전관리를 위한 모델을 만들기 위해서는 이를 활용하기 어렵다. 따라서, 학습 데이터를 직접 수집한 영상을 이용하여 안전관리 모델 학습에 이를 이용해야 한다.

본 논문에서는 제조업 현장에서 직접 수집한 학습 데이터를 직접 구축하였다. 또한 기존에는 1가지의 모델만을 선정하여 진행하였으므로 본 논문에서는 3가지 모델을 선정하여 비교하였다. 이를 통해 위험 상황 추론 모델에 관한 선행 연구로 사용하고자 한다.

II. 산업재해 현황 및 관련 연구

본 장에서는 고용노동부에서 발간한 산업재해 통계 자료를 통해 업종별, 규모별로 나눠보고 이에 대한 원인에 대해 알아본다[3].

2.1. 제조업 세부 업종별 산업재해 현황

2020년 산업재해 발생 현황을 살펴보면 전체 산업재해 사건 수가 108,379건 발생했으며 이 중 제조업이

28,840건으로 전체의 26.6%를 차지하였으며 제조업 세부 업종별로는 표 1과 같다. 제조업에서는 금속을 다루는 업종이 가장 많이 재해가 발생하는 것으로 나타났으며 그다음으로 화학이나 고무 공장, 식품품 공장, 조선 공장에서 나타나는 것으로 나타났다.

Table. 1 Status of detailed industries where many accidents occur by manufacturing in Korea

detailed industry	2020	2019	Increase/Decrease
Machanic/Metal/Nonmetal	14,121	14,439	-318
Chemistry / Rubber	3,080	3,172	-92
Groceries	2,920	2,905	+15
Ship building & repair	2,492	2,306	+186

2.2. 제조업 산업재해 주요 원인

산업재해 현황을 살펴보면 건설/제조업의 비중이 최근 5년간 74.1%를 차지하고 있고 제조업 분야에서 추락 및 끼임 사고가 전체의 48.8%를 차지한 것으로 나타났다. 이들 사고는 안전사고 예방 조치만을 취해도 예방할 수 있는 사고임에도 높은 수치이며 정부에서는 5만여 개의 100인 미만 제조 사업장에 대해 위에서 언급한 사고에 대해 집중 관리를 시행한다고 밝혔다.

2.3. 산업재해 방지를 위한 연구

산업재해 방지를 위한 관련 연구를 살펴보면, 건설 현장에서 작업자가 안전모를 착용하였는지를 판단하기 위해 Faster R-CNN을 활용한 연구가 있다[4]. 또한, 건설 장비를 다루는 작업자와 주변 작업자 간의 거리를 측정하여 위험 상황을 인지하기 위해 YOLO-v3 기반으로 구현한 연구도 존재한다[5]. 두 연구 모두 객체 탐지 모델을 기반으로 했다는 점에서 의의가 있다.

현재 건설 현장 분야에서 이러한 기술들이 활발히 연구되는 사업을 보면 국토교통부에서 추진하고 있는 건설 생산성 혁신과 안전성 강화를 위한 스마트 건설 로드맵 사업과 관련되어있다[6]. 이는 안전사고 예방을 위해 4차 산업혁명의 핵심 기술인 드론(drone), 3D 스캐너, 웨어러블(wearable) 기기, VR(Virtual Reality), AR(Augmented Reality) 등의 기술을 적극적으로 활용하도록 하고 있다.

III. 객체 탐지 모델 선정과 학습

본 장에서는 선정 모델인 YOLO-v4, SSD, CenterNet에 대해 설명하고 학습 데이터셋과 모델 하이퍼파라미터에 대해 기술한다. 이에 대한 전체적인 구조는 그림 1과 같다.

3.1. YOLO-v4

YOLO-v4는 YOLO-v3 후속으로 나온 객체 탐지 모델로 YOLO-v3를 기반으로 하였으며, PANet(Path-aggregation Network)을 사용한 모델이다[7]. PANet은 FPN(Feature Pyramid Network)을 보완한 모델로 그림 2와 같다[8].

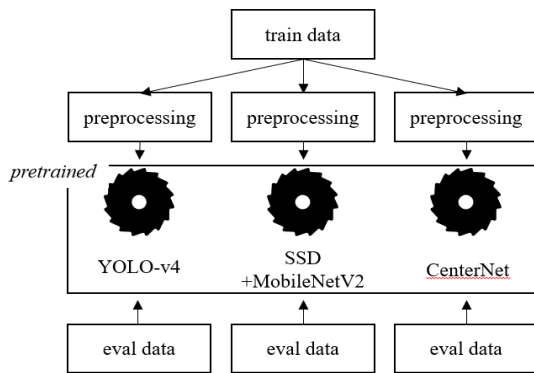


Fig. 1 Overall structure

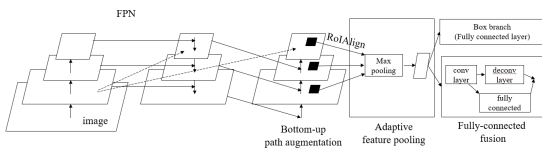


Fig. 2 Schematic of path-aggregation network

PANet은 FPN에 상향식 경로 증강(bottom-up path aggregation)을 추가하여 기존 FPN의 반대의 네트워크를 추가한 것이다. 이는 가장 밑에 있는 계층에서 나온 특징이 마지막 계층까지 될 반영되기 때문인데 이에 shortcut을 추가하여 상위 계층까지 전달하기 위함이다. 또한, SPP(Spatial Pyramid Network)를 추가하여 수용장(receptive field)을 넓혔다.

3.2. SSD와 MobileNetV2

SSD는 YOLO 모델과 마찬가지로 1-stage 방식이며 YOLO-v1의 단점인 작은 객체의 정확도가 떨어지는 단점을 극복한 모델이다[9]. 여러 개의 특징 맵을 이용하고 완전 컨볼루션 네트워크(fully convolutional network)가 사용된다.

본 논문에서 사용한 모델은 기존 VGG-16이 아닌 MobileNetV2를 백본(backbone) 네트워크로 하였으며 MobileNetV2는 역잔차(inverted residual)기법을 적용한 네트워크이다[10]. 역잔차 기법은 깊이별 분리 컨볼루션(depthwise separable convolution)과 선형 병목(linear bottleneck)을 결합한 모듈로 기존의 제안된 잔차 블록에 있는 병목을 역으로 하여 선형 변환(linear transformation) 정보를 상위 계층까지 잘 전달하도록 하기 위함이다. 이를 통해 모바일 환경에서도 사용할 수 있도록 MobileNetV1보다 훈련 매개변수를 획기적으로 줄이고 정확도도 높은 모델이라고 할 수 있다. 표 2는 Keras에서 제공하는 VGG-16 모델과 MobileNetV1, MobileNetV2의 매개변수를 비교한 것이다.

Table. 2 Comparison of parameters of VGG-16 and MobileNet provided by Keras

	VGG-16	MobileNet	MobileNetV2
Total	14,714,688	3,228,864	2,257,984
Trainable	14,714,688	3,206,976	2,223,872
Non-trainable	0	21,888	34,112

3.3. CenterNet

CenterNet도 1-stage 방식이며 키포인트 추정(key-point estimation) 모델이다[11]. 키포인트 추정은 자세 추정(pose estimation) 분야에서 사람의 머리, 목, 가슴, 배 등 특정 부위를 기준으로 활용하기 위한 기술이다. SSD와 YOLO와 다른 점은 1개의 앵커 박스(anchor box)만을 이용하는데 이는 중심점이란 위치를 이용하기 때문에 가능한 것이다. 따라서, SSD와 YOLO와 다르게 박스가 중복되는 문제를 해결하기 위해 제안된 NMS (Non-Maximum Suppression) 기법을 사용하지 않아도 된다.

3.4. 학습 데이터 수집과 데이터 증강 기법

본 논문에서 사용한 학습 데이터는 CCTV가 설치되

어 있는 정밀가공 제조 사업장 1개를 선정하였으며 2020년 11월 16일부터 20일까지 5일간 촬영한 영상이다. 모델을 비교하기 위한 데이터로 사용하기 위해 일자별로 6등분하였으며, 이 중 120개의 프레임만을 이용하였다. 그림 3은 데이터를 나누었던 과정을 보여준다.

어노테이션 툴은 CVAT(Computer Vision Annotation Tool)를 이용하였으며 CVAT는 Django 프레임워크 기반의 오픈소스로 제공되는 툴로 객체 탐지, 인스턴스 분할(instance segmentation) 등의 기능을 제공하며 MS COCO, PASCAL VOC, TFRecord 등 다양한 포맷을 지원한다. 그림 4는 그림 3을 CVAT로 작업 분할한 화면이다.

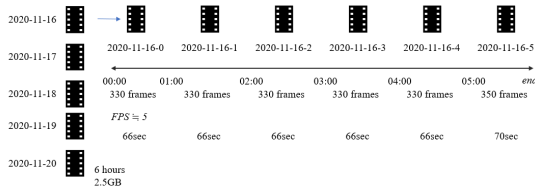


Fig. 3 Process the collected video by dividing

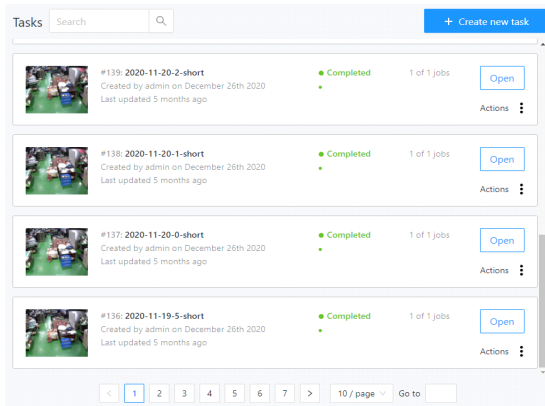


Fig. 4 Task of dividing video using CVAT

객체 클래스는 제조업 현장을 반영하여 작업자, 자재, 카트, 지게차 등을 어노테이션 하였으며, 위험 지역은 절삭 기계나 밀링 등 작업자가 상해를 입을 수 있는 기계를 대상으로 진행하였다. 그림 5는 CVAT로 어노테이션한 일부 예이다.

각 모델에서 제공하는 데이터 증강 기법도 일부 활용하였다. YOLO-v4는 두 이미지를 비율에 따라 동시에 표현하는 Mosaic 방식을 이용하였고, SSD에서는 데이

터 증강 기법을 이용하지 않았으며, CenterNet의 경우 키 포인트를 이용한 증강 기법이 제공되는데 키포인트를 대상으로 어노테이션을 진행하지 않았기 때문에 이를 제거했고 그 외에 이미지를 무작위로 자르는 기법 (randomized cropping)과 패딩(padding) 기법을 이용하였다.



Fig. 5 Part of image annotated

3.5. 하이퍼파라미터 설정

세 모델 모두 반복 횟수(epoch)는 동일하게 6,000번으로 했다. YOLO-v4는 학습률을 0.0013으로 설정하였고 burn in을 시작하는 반복을 1,000번째부터 했으며, SSD는 학습률을 0.8로 설정하고 CenterNet은 5e-3으로 설정하였다. SSD와 CenterNet은 warm up을 시작하는 반복을 2,000번째로 동일하게 설정했다.

배치(batch) 크기는 YOLO-v4와 SSD의 경우 16으로 설정하였고, CenterNet은 8로 하였다. CenterNet의 배치 크기를 줄인 이유는 GPU 메모리를 YOLO-v4나 SSD보다 더 많이 쓰기 때문에 줄인 것이다.

3가지 모델 모두 같은 사양의 환경에서 진행했으며 표 3이 모델 학습을 진행한 환경이다.

Table. 3 Hardware specifications

CPU	i7-10700 (2.9 Ghz, 16 Core)
RAM	16 GB
GPU	Nvidia Geforce GTX 1660
Storage	SSD 250 GB / HDD 4TB
OS	Ubuntu 20.04.1 LTS
Virtual Environment	Anaconda 3

3.6. 성능 지표

세 모델 모두 COCO detection metric을 이용하였다.

COCO detection metric에서 mAP(mean Average Precision)는 IoU(Intersection over Union) 값이 0.5부터 0.95까지의 임계값 범위에 대한 AP의 평균이다. AP는 PR-curve의 아래 면적을 의미하는데 PR-curve는 정밀도와 재현율을 그래프로 표현한 것이며 정밀도와 재현율에 사용되는 TP(True Positive)와 FP(False Positive)를 나누는 기준이 IoU이다. IoU는 정답(ground truth) 박스와 모델이 예측한 경계 상자의 합집합 분의 교집합이 되는 값이다. 이 값을 모델 학습을 하기 전에 미리 값을 정하고 이 값 이상이면 객체를 검출했다 판단하는 것이다.

3.7. 학습 결과와 고찰

학습 결과 YOLO-v4, SSD, CenterNet 순으로 mAP를 기록하였다. 표 4는 학습 결과이며 그림 6은 SSD(주황)와 CenterNet(빨강)의 최종 손실 값 그래프이다.

Table. 4 Result of the mAP

	YOLO-v4	SSD + MobileNetV2	CenterNet
mAP	0.81	0.4702	0.3034

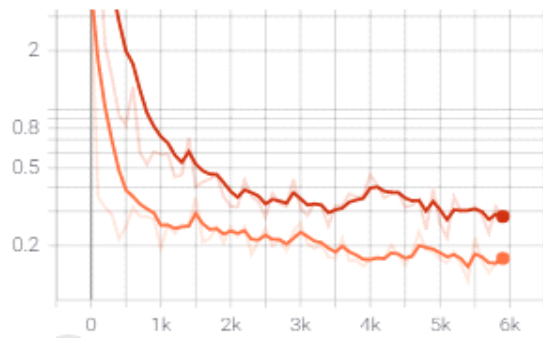


Fig. 6 Result of training with SSD and CenterNet

YOLO-v4의 학습 결과는 그림 7과 같으며 파란색 선이 손실 값을 의미하고 빨간색 선이 mAP를 나타낸다. x축은 반복 횟수이고, y축은 손실 값이다.

본 논문에서 사용한 세 가지 모델의 경우 객체를 검출하는 기법, 손실 함수 등이 조금씩 다르므로 동일선상에 놓고 비교하기에는 무리가 있지만, 모델이 제공하는 기본적인 설정을 유지하고 모델을 학습했을 때의 결과로 YOLO-v4가 가장 좋은 성능을 보였다. CenterNet의 경우에는 3.4절에서 설명했듯이 키 포인트를 이용한 데이터 증강 기법을 제외했기 때문에, 가장 낮은 성능을 보

인 것으로 판단된다. 그림 8은 모델 평가 시 출력된 영상 추론의 예이다.

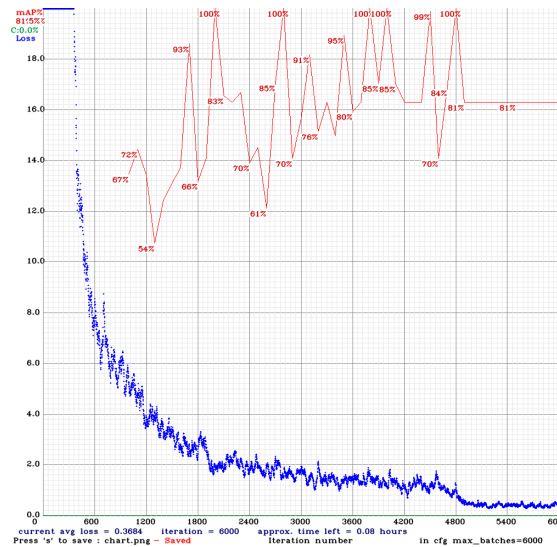


Fig. 7 Result of training with YOLO-v4

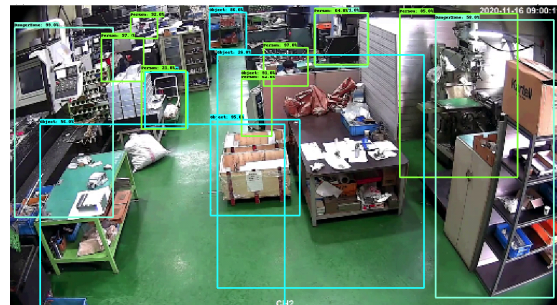


Fig. 8 Some examples of image inference

IV. 결론

본 논문에서는 제조업 근로자의 안전관리를 위한 현황을 분석해보고 학습 데이터를 직접 구축하였으며 세 가지 객체 모델을 선정하여 학습을 진행하였다.

정부와 산업계에서는 작업 현장에서 발생하는 사고 예방을 위한 요구가 커지고 있으며 이를 딥러닝 기반의 객체 탐지 모델을 활용하는 관련 연구가 존재한다. 이러한 모델을 활용하기 위해서는 작업자, 보호장비, 굴착기와 같은 작업 환경에서 위험한 장비들을 선정하고 충분

한 데이터셋이 존재해야만 한다. 이를 통해 객체를 인식하여 객체 간의 거리를 측정하거나 보호장비를 착용했는지 여부 등을 판단한 다음 위험 상황을 판정해야 한다.

객체를 탐지하기 위해 YOLO-v4, SSD, CenterNet을 비교하여 YOLO-v4가 가장 높은 mAP 0.81을 달성하였다. 이 세 가지 모델의 아키텍처, 전처리와 후처리 기법, 최적화(optimizer) 기술, 데이터 증강 기법 등이 다르므로 동일선상으로 비교하기에는 무리가 있지만, 제조업 현장에서 발생할 수 있는 위험 상황에 대한 보호장구, 작업자가 사용하기에 위험한 장비 등을 선정해 직접 데이터셋을 구축하여 이를 모델 훈련에 활용하였다는 점에서 향후 연구로 위험 상황 판정 모델의 선행 연구자료로 사용될 수 있을 것으로 사료된다.

ACKNOWLEDGEMENT

This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: P0015365)

REFERENCES

[1] M. Jeong, C. Lee, and W. Cho, "Design of Integrated Safety System for Sealed Places," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 23, no. 1, pp. 97-102, Jan. 2019.

[2] Z. Fang and P. Kim, "Comparison of Deep-Learning Algorithms for the Detection of Railroad Pedestrians," *Journal of Information and Communication Convergence Engineering*, vol. 18, no. 1, pp. 28-32, Mar. 2020.

[3] Ministry of Employment and Labor. Status of industrial accidents at the end of December 2020 [Internet]. Available: https://www.moel.go.kr/policy/policydata/view.do?bbs_seq=20210401122.

[4] D. Kim, J. Kong, J. Lim, and B. Sho, "A Study on Data Collection and Object Detection using Faster R-CNN for Application to Construction Site Safety," *Journal of Korean Society of Hazard Mitigation*, vol. 20, no. 1, pp. 119-126, Feb. 2020.

[5] S. Shim and S. Choi, "Development on Identification Algorithm of Risk Situation around Construction Vehicle using YOLO-v3," *Journal of the Korea Academia-Industrial cooperation Society*, vol. 20, no. 7, pp. 622-629, Jul. 2019.

[6] Ministry of Land, Infrastructure and Transport. Realization of commercialization of smart construction core technology by 2025 [Internet]. Available: https://www.molit.go.kr/USR/NEWS/m_71/dtl.jsp?lcmepage=1&id=95083436.

[7] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

[8] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path Aggregation Network for Instance Segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8759-8768, 2018.

[9] W. Liu, D. Auguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," in *European Conference on Computer Vision (ECCV)*, Springer, Cham, pp. 21-37, 2016.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottleneck," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510-4520, 2018.

[11] X. Zhou, D. Wang, and P. Krahenbuhl, "Object as Points," *arXiv preprint, arXiv:1904.07850*, 2019.



이태준(Taejun Lee)

2020년 배재대학교 컴퓨터공학(공학사)
2021년~현재, 배재대학교 컴퓨터공학(석사과정)
※ 관심분야 : Deep Learning, Object Detection, Big data



김윤정(Yunjeong Kim)

2005년~현재 한국과학기술정보연구원 재직
1998년 홍익대학교 전자전산공학과(공학사)
2005년 홍익대학교 전자전산공학과(공학석사)
2019년~현재 배재대학교 컴퓨터공학과(박사 과정)
※ 관심분야 : 정보분석, 빅데이터, 머신러닝



정희경(Hoekyung Jung)

1985년 광운대학교 컴퓨터공학과(공학사)
1987년 광운대학교 컴퓨터공학과(공학석사)
1993년 광운대학교 컴퓨터공학과(공학박사)
1994년~현재 배재대학교 컴퓨터공학과 교수
※ 관심분야 : Machine learning, Big data, Embedded system, U-Healthcare, IoT