

## 스마트건설 현장에서 개인 보호장비 검출의 개선 방법

박용석\*, 이송연\*, 이경택\*

Improving Personal Protective Equipment Detection  
in Smart Construction

Yong-Suk Park\*, Song-Yeon Lee\*, Kyung-Taek Lee\*

요약

건설현장에서 산업재해 및 사망사고를 줄이기 위해 작업자의 PPE(개인 보호장비) 착용여부를 확인할 필요가 있다. 현장 관리자가 육안으로 작업자의 PPE 착용여부를 확인하는 것은 많은 시간적, 인적 자원이 소요된다. 스마트 건설 기술의 도입으로 객체검출 기술을 사용하여 작업자의 PPE 착용여부 모니터링을 자동화할 수 있게 되었다. 여러 객체검출 알고리즘 중 YOLO는 실시간 객체검출이 가능하여 PPE 검출에 활용되고 있다. 건설현장에서 넓은 시야각 확보를 위해 모니터링 카메라를 원거리에 설치할 경우 출력영상에서의 PPE 객체 크기는 매우 작아진다. YOLO 알고리즘의 고질적인 문제로 작은 객체 검출에 취약한 부분이 있다. 본 논문에서는 이를 보완하기 위해 모션 모자이크 이미지 전처리 기법을 제안한다. 이 기법은 이미지에서 모션 감지를 통해 관심 PPE 객체의 존재가능 영역들을 추출하고, 해당 영역들을 모자이크 형태로 조합하여 YOLO로 입력되는 이미지를 재구성한다. 이로 인해 검출 영역의 해상도를 최대한 보존할 수 있어 YOLO를 이용한 작은 크기의 PPE 검출율을 향상시킬 수 있다.

키워드 : 스마트건설, 개인 보호장비, 객체검출, 모션 감지, 모자이크 이미지

Key Words : Smart construction, Personal protective equipment, Object &amp; motion detection, Mosaic imaging

## ABSTRACT

In order to reduce accidents and fatalities in construction sites, it is important to ensure that the workers are wearing appropriate personal protective equipments (PPEs). Manual inspection and monitoring of PPE usage is time consuming and requires significant human resources. The introduction of smart construction enabled object detection techniques to be used for automatic detection of PPE usage. Among various object detection algorithms, YOLO enables PPE detection in real-time and is widely used. If the monitoring camera is installed afar in order to secure a wide view of the construction site, the size of PPE objects in the output image becomes very small. YOLO has weaknesses in detecting small objects, so in this case, it will not be able to perform effective PPE detection. In this paper, a motion mosaic image pre-processing method is proposed to improve small PPE detection by YOLO. The method uses motion detection in order to extract areas of probable PPE locations. A mosaic image is created from the extracted areas and fed as input to YOLO. This method preserves the detection area resolution as best as possible, which increases small sized PPE detection rate by YOLO.

※ 본 연구는 산업통상자원부 및 산업기술평가관리원 산업용임베디드시스템기술개발사업-임베디드인공지능시스템기술개발 연구비 지원으로 수행되었습니다(20007872, 사용자 행동예측을 위한 임베디드 인공지능 엔진 및 시스템 기술개발).

•\* First and Corresponding Author : Korea Electronics Technology Institute, yspark@keti.re.kr, 책임연구원, 정회원

\* Korea Electronics Technology Institute, sylee8466@keti.re.kr, 연구원; ktechlee@keti.re.kr, 수석연구원, 정회원

논문번호 : 202011-271-0-SE, Received October 29, 2020; Revised November 20, 2020; Accepted November 23, 2020

## 1. 서 론

정보통신 기술의 발전과 함께 오늘날 우리가 살고 있는 도시는 스마트시티로 진화하고 있다. 건설업은 스마트시티를 구성하는 다양한 스마트 인프라를 구축하기 위한 공급의 주체이다. 안타깝게도 건설업은 산업분야 중 사망사고가 가장 많이 발생하는 분야이기도 하다. 한국산업안전보건공단의 산업재해통계 자료에 따르면 2020년 6월 기준 전체 산업재해 사고 사망자 중 54%가 건설업에서 발생했다. 건설현장에서 산업재해 사망사고를 줄이기 위한 안전성 확보가 시급한 이유다.

건설현장에서 안전성과 생산성을 높이기 위해 사물인터넷(IoT, Internet of Things)과 인공지능(AI, Artificial Intelligence)을 주축으로 한 다양한 ICT 기술이 융합된 스마트건설 기술이 시범 도입되고 있다. CCTV 카메라, 센서, 드론 등 다양한 출처로부터 현장 정보를 수집, 공유, 분석하여 안전한 작업환경 구축을 추구한다. 예를 들어, IoT 가스센서를 이용하여 유해 가스를 감지하고 허용농도 초과 시 환기시스템과 연계하여 공기를 순환시켜, 주변 작업자들에게 스마트폰 등을 통해 통보한다. 또한, AI가 탑재된 스마트 CCTV 카메라를 이용하여 현장의 고위험 상황, 불안정한 작업자의 행동, 제한구역 침입 등을 실시간으로 감지하고 대응하기도 한다.

스마트건설 관련 기술 중 개인 보호장비(PPE, Personal Protective Equipment) 검출은 컴퓨터 비전과 AI 기술을 이용하여 현장 근로자의 보호 장비 착용 여부를 확인할 수 있도록 한다. 안전모, 안전조끼, 안전벨트, 안전화, 장갑 등이 PPE에 해당된다. 우리나라는 산업안전보건기준에 관한 규칙과 산업안전보건법에 따라 사업주의 보호 장비 지급과 근로자의 보호 장비 착용이 의무화 되어있다. 현장 근로자는 담당하고 걸리적거리는 보호 장비를 착용하지 않고 작업하는 경우가 빈번하고, 현장 관리자는 모든 근로자들을 감지하기에는 한계가 있다. 따라서 PPE 검출 기술을 통해 근로자의 보호 장비 착용여부를 자동으로 확인할 수 있도록 하면 근로자들의 안전을 지킴과 동시에 관리감독의 인적, 시간적 노력을 줄이고 업무효율을 높일 수 있다.

PPE 검출에 적용되는 핵심 기술은 객체검출(object detection) 기술이다. 객체검출은 이미지에서 관심 객체를 배경과 구분해 식별하는 자동화 기법으로, 이미지 내에서 관심 객체의 존재 여부와 위치를 경계박스(bounding box) 설정을 통해 파악할 수 있게 해준다.

객체검출을 위해 딥러닝(deep learning) 알고리즘이 많이 활용되며 대표적으로 YOLO(You Only Look Once) 알고리즘이 사용된다. YOLO는 다른 딥러닝 기반 객체검출 알고리즘과 비교하여 빠른 처리속도와 상대적으로 높은 검출 정확도를 보인다<sup>1)</sup>. 이로 인해 YOLO는 PPE 검출을 포함하여 다양한 실시간 객체검출 응용분야에 활용되고 있다. YOLO의 고질적인 문제로 작은 객체 검출에 취약한 부분이 있는데, 검출 성능을 높이기 위해 Bag of Freebies, Bag of Specials 등 다양한 기법이 적용된 YOLOv4가 최근 발표되었다<sup>2)</sup>.

건설현장 모니터링 및 PPE 검출을 위해 사용되는 CCTV 카메라는 다양한 현장위치에 설치된다. 실외 환경에서 넓은 시야각 확보를 위해 원거리에 설치되는 카메라의 영상출력에서 근로자는 상대적으로 매우 작게 보인다. 이런 위치에 설치된 카메라의 이미지에 대해 YOLOv4 기반 PPE 검출을 수행할 경우, 작게 표출되는 보호 장비는 검출이 안 되는 경우가 많다.

본 연구에서는 원거리 촬영 동영상에서 PPE 검출 시, 크기가 작은 PPE 객체에 대한 검출 성능을 향상시킬 수 있는 이미지 전처리 기법을 제안한다. YOLOv4에서 객체검출을 위해 입력되는 이미지의 크기는 학습 시 설정된 입력영상의 크기와 동일하게 조절(resize)된다. 따라서 고해상도 영상이 입력으로 제공되어도, 원본 입력 이미지 크기가 축소 조절되는 현상이 발생하며, 이는 객체 검출에 영향을 미치게 된다. 전처리 과정을 통해 원본 영상의 품질을 유지한 입력 영상이 제공된다면 검출 성능이 향상될 수 있다. 제안된 기법은 영상 프레임에서 객체검출 시 의미 있는 부분만을 추출하여 입력으로 제공한다. 즉, 영상 프레임에서 모션(motion, 움직임)이 발생하는 부분만을 추출하여 모자이크 방식으로 결합, 재구성하여 YOLOv4 객체검출 모델의 입력으로 제공한다. 관심 영역 또는 의미 있는 부분만을 추출함으로써 입력 영상의 크기를 줄이고 영상에서 검출 대상이 되는 영역의 원래 해상도를 최대한 유지할 수 있다. 이를 통해 작은 객체에 대한 검출 성능을 개선할 수 있다.

본 논문은 다음과 같이 구성된다. II장에서는 PPE 검출과 관련하여 진행된 여러 연구 성과를 소개하고, III장에서는 PPE 검출 성능을 개선하기 위해 새롭게 제안하는 모션 모자이크 이미지 전처리 기법을 설명한다. IV장에서는 제안하는 전처리 기법을 YOLOv4에 적용한 실험 결과를 필드 영상에 대한 객체 검출 성능 비교를 통해 분석하고, V장에서는 본 연구에 대한 결론과 향후 연구 진행 방향에 대해서 논의한다.

## II. 관련 연구

근로자의 안전과 직결된 PPE 검출기술의 개발을 위해 다양한 기법과 알고리즘이 연구되었다. 초기에는 cascade classifier를 이용한 컴퓨터 비전 기반의 PPE 검출 방법이 연구되었다<sup>3)</sup>. HOG(Histogram of Oriented Gradients), Haar-like, LBP(Local Binary Pattern) 기반의 다양한 다단계 분류자(cascade classifier)를 사용하여 근로자의 안전모 착용여부를 판단하였다. 하지만 안전모에서 추출되는 유효 특징점(feature)의 개수가 부족하여 고유 문양의 스티커를 부착하여 추가적인 마커(marker)를 생성해야 하는 등의 불편함이 있었다.

대부분의 PPE 검출과 관련된 연구는 근거리, 실내 환경에서 많이 진행되었으며, 실제 건설현장을 반영한 원거리 실외환경에서의 PPE 검출 연구도 진행되었다<sup>4)</sup>. 계절과 날씨, 카메라와 작업자 간 거리(visual range), 현장에서 발생하는 occlusion 현상 등이 안전모 착용여부 판단에 미치는 영향이 연구되었으며, 해당 연구에서는 Faster R-CNN<sup>5)</sup> 기법이 높은 정밀도(precision)를 보였다. R-CNN(Regions with Convolutional Neural Network)은 2개의 단계(stage)로 구성되는 검출 알고리즘으로 1단계에서는 region proposal을 이용하여 객체가 존재할만한 영역을 추출하고 2단계에서는 추출된 영역에서 객체검출을 수행한다. Faster R-CNN은 RPN(Region Proposal Network)을 이용하여 1단계의 region proposal을 모델 내부로 통합시켜 기존 방식 대비 처리속도를 향상시킨다.

Faster R-CNN이 실시간으로 거리와 조도에 비례하여 안전모, 안전조끼, 마스크, 장갑, 안전화 등 복수의 PPE 검출을 효과적으로 할 수 있는지에 대한 연구도 진행되었으며 장갑과 안전화 등 작은 크기의 객체에 대해서는 검출의 한계가 확인되었다<sup>6)</sup>.

에지 컴퓨팅(edge computing) 환경에서의 AI가 조명받기 시작하면서 실시간 운용과 연산 경량화가 용이한 YOLO 알고리즘이 주목을 받기 시작하였다. YOLO는 이미지를 격자 그리드(grid)로 나누어 한 번에 객체의 위치와 클래스를 판단하고 최종 객체를 구분하기 때문에 실시간으로 동작이 가능하다. YOLO는 버전 업그레이드를 통해 연산속도를 높이기 위한 개선방안을 꾸준히 도입하였으며 YOLOv4에서는 컴퓨팅 파워가 낮은 환경에서도 성능 저하를 최소화한 객체검출이 가능하게 한다. [7]에서는 YOLO를 사용하여 실내환경에서 안전모, 안전조끼, 장갑, 벨트, 바

지, 안전화의 6가지 PPE에 대한 실시간 검출 성능을 검토하였다. [8]에서는 YOLOv2<sup>9)</sup>를 사용하여 실내환경에서 안전모에 대한 실시간 검출 성능을 검토하고 알람연동 기능을 연구하였다. YOLOv3<sup>10)</sup>를 실시간 PPE 검출에 적용하는 [11]에서는 PPE 구분(classification)을 다양한 방식으로 시도하고 학습하여 성능을 측정하였다. 해당 논문에서는 근로자, 안전모, 안전조끼를 각각의 객체로 구분하여 학습하는 방식과 근로자-안전모, 근로자-안전조끼, 근로자-안전모-안전조끼와 같이 PPE 객체 그룹으로 구분하여 학습하는 방식을 진행하였다. 개별 객체로 학습한 모델은 검출 속도가 빨랐고, 객체 그룹으로 학습한 모델은 검출 성능(mAP, mean Average Precision)이 높았다. 이외에도, YOLOv4를 PPE 검출에 적용하였을 때 mAP가 81.81%까지 나오는 결과를 보여준 연구도 있다<sup>12)</sup>. 해당 연구에서는 PPE 클래스를 작업자, 안전모 착용, 안전모 미착용, 안전조끼 착용, 안전벨트 착용의 5개로 구성하였다.

## III. 모션 모자이크 이미지 전처리 기법

YOLO에서는 학습과 추론 시 입력 이미지를 32 픽셀의 배수로 조절(resize)하며 기본값은 416x416 픽셀이다. 예를 들어 크기가 1920x1080 픽셀인 이미지를 입력으로 제공할 경우, YOLO는 입력 이미지를 416x234 픽셀로 축소하고 나머지 부분은 패딩(padding) 처리하여 416x416 픽셀 크기의 이미지를 생성한다. 원본 입력 이미지에 검출해야 할 작은 크기의 객체가 존재할 경우, 이미지 크기 조정 과정에서 픽셀 수 감소로 인하여 유효 특징이 줄어들거나 사라질 수 있다. 입력 이미지의 해상도를 최대한 보존할 수 있게 된다면 작은 크기의 객체 검출 성능도 향상될 수 있다.

PPE 검출 시 우리는 작업자와 작업자가 착용하고 있는 보호 장비에 관심이 있다. 작업자가 존재하지 않는 이미지 영역은 처리할 필요가 없다. 이미지에서 작업자가 존재할만한 영역만 선별적으로 추출하여 이미지를 재구성한다면 전체적인 입력 이미지 크기를 줄이면서 작업자의 해상도를 최대한 보존할 수 있게 된다. 현장에서 작업자는 움직이고 이동하기 때문에 이미지에서 모션이 발생하는 부분을 감지하면 작업자가 존재할만한 영역을 유추할 수 있다.

그림 1은 제안하는 모션 전처리 기법이 적용된 PPE 검출 프로세스 과정을 보여준다. Motion Region Detection 단계에서는 영상 프레임에서 모션이 발생하

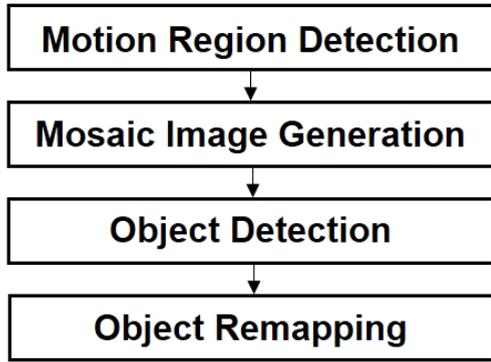


그림 1. 모션 전처리가 적용된 PPE 검출 프로세스  
Fig. 1. PPE detection with motion pre-processing

는 부분을 감지 및 추출한다. Mosaic Image Generation 단계에서는 추출된 이미지를 모자이크 방식으로 결합하고 재구성한다. Object Detection 단계에서는 YOLOv4 객체검출 모델을 사용한다. Object Remapping 단계에서는 Mosaic Image에 대하여 검출된 객체 위치정보를 원본 이미지에 대응시켜 적용한다. Object Remapping 단계는 검출된 PPE 객체를 원본 이미지에 표시하지 않아도 될 경우 생략 가능하다.

그림 2는 모션 영역 감지 프로세스를 보다 자세히 보여준다. 모션 영역 감지를 위해 두 개의 연속적인 이미지 프레임이 요구된다. 프레임들을 gray scale로 변환 후, 현재 이미지 프레임(current frame)과 이전 이미지 프레임(previous frame) 간 픽셀 차이를 계산하여 픽셀 간 변화를 찾아낸다. Frame differencing 과정을 통해 이진화된 threshold intensity 이미지를 얻게 되면 움직임이 발생한 영역을 찾을 수 있다. 영상 소스(CCTV)가 움직이지 않고 고정된다면 정적인 배경(background)과 동적인 전경(foreground)을 분리하는 background subtraction 기법을 사용하여 모션 영역을 찾을 수도 있다.

이미지 픽셀 영역을 키우는 morphology 처리 과정을 거쳐 객체 검출영역을 확보한다. 전경이 되는 이미지의 경계 부분을 침식시켜 배경 이미지로 전환하는 erosion을 통해 영상의 노이즈를 제거한 다음 이미지를 팽창시키는 dilation을 적용하여 움직임이 발생한 영역을 확장하고 blob 형태로 구성한다.

Morphology 처리를 통해 생성된 모션 blob들에 convex hull을 적용하여 윤곽(contour)을 찾아낸다. Convex hull은 이미지 blob 객체의 convexity defect (오목한 부분)를 보정하여 blob 객체를 감싸는 윤곽을 찾는다. 계산된 윤곽에 외접하는 사각형 영역을 계산

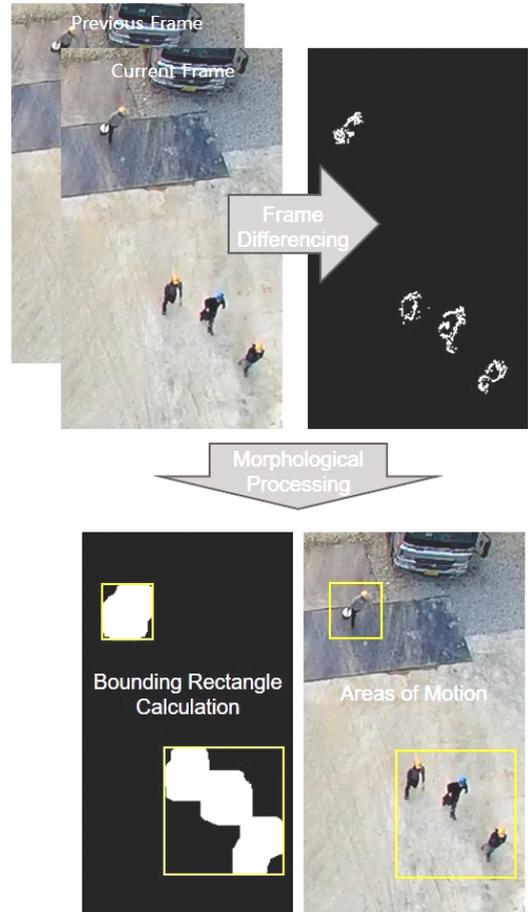


그림 2. 모션 영역 감지 방법  
Fig. 2. Detecting image regions with motion

하여 경계박스로 설정하고 이를 움직임이 발생하는 영역(areas of motion)으로 지정한다.

그림 3은 감지된 모션 영역들을 하나의 이미지로 재구성하는 모자이크 이미지 생성 절차를 보여준다.

감지된 모션 영역들에 대해 면적(area)을 구하고 면적이 큰 순서로 정렬한다. 모션 경계박스 중 가장 큰 면적을 가진 것을 기준으로 차순위 면적 크기의 모션 경계박스들을 인접시켜 새로운 이미지를 구성한다. 모자이크 이미지를 구성할 때 가로, 세로 비율이 가급적 1:1이 되도록 하며, 가급적 YOLO의 입력 이미지 크기(예. 416x416 픽셀)를 벗어나지 않는 크기로 구성한다. 모자이크 이미지에서 모션 영역으로 채워지지 않는 부분은 padding 처리하여 사각형 형태의 이미지를 구성한다. 각 모션 경계박스에 대해 원본 이미지에서의 좌표와 모자이크 이미지에서의 좌표를 저장하여, 필요 시 그림 4에서와 같이 모자이크 이미지에서 검

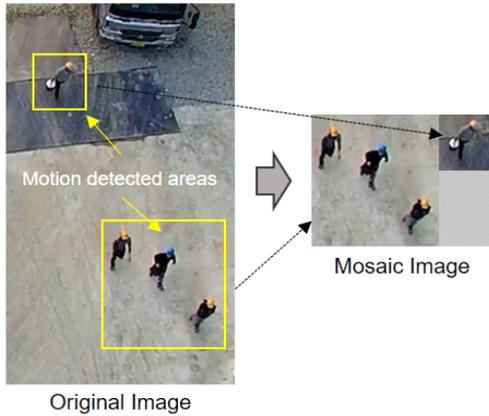


그림 3. 모션 모자이크 이미지 생성  
Fig. 3. Motion mosaic image generation

출되는 PPE가 원본 이미지로 표출될 수 있도록 한다.

#### IV. 실험방법 및 결과

##### 4.1 YOLOv4 모델 생성

YOLOv4 모델 학습은 Ubuntu 16.04.06 LTS 서버 환경에서 진행하였다. 서버는 NVIDIA TITAN Xp GPU, Intel Xeon CPU와 125GB 메모리로 구성된다. 좀 더 빠르고 정확한 학습 진행을 위해 pre-trained 모델인 yolov4.conv.137을 사용하였다. 학습에 사용된 YOLOv4 모델의 backbone은 CSPDarknet53이다. CSPDarknet53은 53개 계층으로 구성된 DarkNet-53을 사용하는 CNN으로 연산량을 완화하는 CSPNet(Cross Stage Partial Network) 모델을 사용한 다<sup>[13]</sup>. 해당 모델에는 네트워크 연산 속도의 감소 없이 주요 특징을 추출할 수 있도록 SPP(Spatial Pyramid Pooling) 블록이 추가되었다. SPP 블록을 추가함으로써 네트워크의 수용영역을 증가시키고 이전 버전인 YOLOv3 대비 성능을 향상시켰다.

GPU 메모리 사용을 고려하여 입력 이미지는 416x416 픽셀 크기로 고정하였으며, 인터넷을 통해 수집된 다양한 크기로 구성된 1,740장의 건설현장 이미지를 8,000번 반복 학습시켰다. 학습 시 데이터 증강(data augmentation) 기법 중 하나인 mosaic 옵션을 활성화하여 정확도를 향상시켰다.

PPE 검출을 위해 정의된 클래스는 worker(작업자), hard hat(안전모), vest(안전조끼) 세 가지이며, 각 클래스가 학습을 진행한 ground truth 개수는 표 1과 같다.

학습된 모델의 정확도를 평가하기 위하여 ground truth 영역과 검출 영역의 IoU(Intersection over



그림 4. 원본 이미지로 검출정보 표출  
Fig. 4. Object remapping from mosaic image

표 1. 클래스 당 학습을 진행한 ground truth 수  
Table 1. Ground truth samples trained per class

	Worker	Hard Hat	Vest	Total
Ground Truth	3073	2299	848	6219

표 2. 클래스 별 AP  
Table 2. AP per class

	True Positive	False Positive	AP(%)
Worker	752	245	85.01
Hard Hat	577	193	82.53
Vest	200	81	67.08

Union)를 이용한 정확도 AP@.50은 표 2와 같다. IoU > 0.5인 경우 True Positive로 결정되고, 아닐 경우 False Positive로 결정된다. 정확도 평가에는 AP가 가장 높은 5,000번째 학습 시 저장된 가중치(weights) 모델을 사용하였다.

##### 4.2 전처리 적용 테스트

모션 모자이크 이미지 전처리를 적용했을 때 성능 비교분석을 위해 Full HD(1920x1080 픽셀) 해상도의 건설현장 CCTV 카메라 영상이 사용되었다. 영상에서 PPE를 착용한 작업자들이 보이는 다양한 장면 50개를 선별하고, 전후 프레임 시퀀스(sequence)를 캡처하여 모션 모자이크 이미지를 생성하였다. 장면에는 각종 PPE를 착용한 작업자가 최소 6명에서 최대 12명까지 나타난다.

50장의 원본 이미지와 이에 해당되는 50장의 모션 모자이크 이미지에 대해 YOLOv4 모델을 이용하여 PPE 검출을 진행하였다. PPE 검출은 학습을 진행한

컴퓨터와 동일한 사양의 NVIDIA TITAN Xp, Intel Xeon CPU, 125GB 메모리가 장착된 Ubuntu 16.04.06 LTS 환경에서 진행했다. 한 개의 테스트 이미지를 처리하여 PPE를 검출하는데 평균적으로 17.5ms의 시간이 소요되었다.

그림 5는 전처리하지 않은 원본 이미지 일부에 대한 PPE 검출 결과(Detection using YOLOv4)와 모션 모자이크 이미지 전처리를 적용했을 때의 PPE 검출 결과(Detection using Motion Mosaic + YOLOv4) 예시를 보여준다. 전처리 전에는 감지되지 않던 안전모가 전처리 후에는 검출됨을 볼 수 있다. Worker에 대한 경계박스도 보다 정교하게 생성됨을 볼 수 있다.

모션 모자이크 이미지 전처리를 적용했을 때와 하지 않았을 때의 성능 비교를 위해 검출율(recall)과 정확도(precision)를 측정하였다. Recall은 검출 ground truth로 설정된 PPE를 누락시키지 않고 얼마나 잘 검출하는지를 나타낸다. 즉, 실제 True인 것 중 모델이 True라고 예측한 것의 비율이다. Precision은 검출 결과 중 실제 검출 대상이 얼마나 포함되어 있는지를 나타낸다. 즉, 모델이 True라고 분류한 것 중, 실제 True인 것의 비율이다. 표 3은 PPE 검출 테스트 결과를 보여준다.

표 3에서 Size Range는 테스트 이미지에 포함된 PPE경계박스 픽셀 크기의 최대 및 최소 범위를 나타낸다. Average PPE Count는 각 테스트 이미지에 포함된 평균 PPE 개수를 나타낸다. 검출율은 Worker, Vest, Hat 순으로 객체의 실제 크기가 클수록 높음을 볼 수 있다. Worker의 검출율은 전처리를 적용하였을

표 3. PPE 검출 테스트 결과  
Table 3. PPE detection test results

Class	Size Range	Hard Hat	Recall	
			YOLOv4 only	Proposed method
Worker	150x250 30x100	5.9	0.61	0.84
Hat	40x50 40x25	6.25	0.52	0.74
Vest	70x90 45x50	6	0.58	0.85

경우 0.23 높으며, 전처리를 하지 않았을 경우 대비 37% 상승하였다. Hat의 경우 전처리 적용 시 검출율은 전처리를 하지 않았을 때와 비교하여 0.22 높으며 이는 42% 상승에 해당된다. Vest의 검출율 또한 전처리를 적용한 경우가 YOLOv4만 적용했을 때 보다 0.27 높았다(46% 상승). 검출율은 평균적으로 모션 모자이크 이미지 전처리를 적용한 경우(Proposed method)가 전처리를 하지 않은 경우(YOLOv4 only) 보다 약 42% 정도 높음을 볼 수 있다. 정확도 값은 모든 경우에서 1로 나와 표 3에 포함시키지 않았다.

표 4는 원본 이미지에서는 검출이 안 되고 모션 모자이크 이미지 전처리가 적용된 경우에만 검출된 PPE 클래스의 정보를 보여준다. Worker의 경우 20x40 픽셀, Hat의 경우 20x15 픽셀 크기의 객체를 검출할 수 있었다. 따라서 제안된 기법으로 YOLOv4를 단독으로 사용했을 때는 검출되지 않는 작은 PPE 객체도 검출할 수 있다.

표 4. 전처리 적용 시에만 검출된 PPE 정보  
Table 4. PPE detected only by using proposed method

Class	Size Range	Avg. PPE count	Recall (Proposed method)
Worker	30x60 20x40	3.5	0.58
Hard Hat	30x25 20x15	6.7	0.67

## V. 결 론

본 논문에서는 스마트건설 산업현장 영상에서 작은 크기의 PPE 검출이 가능하도록 하는 모션 모자이크 이미지 전처리 기법을 제안하였다. 제안된 기법으로 입력 이미지를 전처리할 경우 YOLOv4와 같은 객체

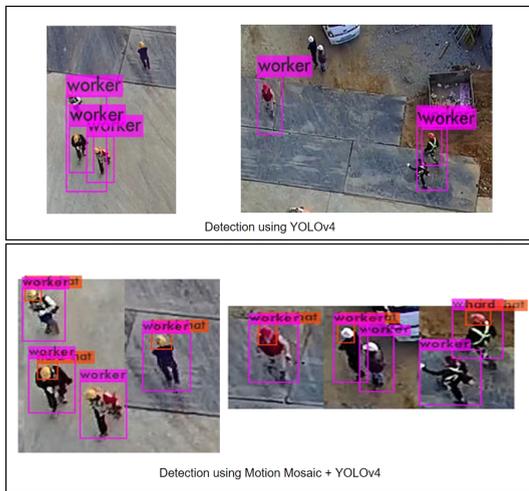


그림 5. PPE 검출 결과 예시  
Fig. 5. Sample PPE detection results

검출 알고리즘의 검출율을 높일 수 있다. 실험결과를 통해 모션 모자이크 이미지 전처리를 적용한 경우 YOLOv4를 이용한 PPE 검출율이 약 42% 향상됨을 확인하였다. 제안된 기법은 기능적으로는 객체가 존재 할만한 영역을 추출하는 region proposal과 유사한 역할을 한다. 제안된 기법에서는 전후 프레임 비교를 통해 모션을 감지하고 관심 객체의 존재가능 영역을 추출하기 때문에 사진과 같은 정지영상에는 적용할 수 없으며 동영상 기반의 PPE 검출에만 적용이 가능하다. 향후 안전벨트, 마스크 등 PPE 종류를 추가하여 PPE 검출 테스트를 진행하고, 보다 많은 학습 데이터 수집을 통해 전체적인 모델의 객체검출 성능을 향상시킬 예정이다.

## References

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. 2016 IEEE CVPR*, pp. 779-788, Las Vegas, NV, Jun. 2016.
- [2] A. Bochkovskiy, C. Wang, and H. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv:2004.10934v1* [cs.CV], Apr. 2020.
- [3] B. E. Mneymneh, M. Abbas, and H. Khoury, "Automated hardhat detection for construction safety applications," *Procedia Eng.*, vol. 196, pp. 895-902, Jun. 2017.
- [4] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T. M. Rose, and W. An, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Automat. in Construction*, vol. 85, pp. 1-9, Jan. 2018.
- [5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 39, no. 6, pp. 1137-1149, Jun. 2017.
- [6] F. Zhafran, E. S. Ningrum, M. N. Tamara, and E. Kusumawati, "Computer vision system based for personal protective equipment detection, by using convolutional neural network," in *Proc. 2019 IES*, pp. 516-521, Surabaya, Indonesia, Sep. 2019.
- [7] H. M. Hung, L. Lan, and H. S. Hong, "A deep learning-based method for real-time personal protective equipment detection," *J. Sci. and Technique*, no. 199, pp. 23-34, Jun. 2019.
- [8] C. B. Souto Maior, et al., "Personal protective equipment detection in industrial facilities using camera video streaming," *Safety and Reliability - Safe Societies in a Changing World*, 1st Ed., Taylor & Francis Group, pp. 2863-2868, 2018.
- [9] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proc. 2017 IEEE CVPR*, pp. 6517-6525, Honolulu, HI, Jul. 2017.
- [10] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv:1804.02767v1* [cs.CV], Apr. 2018.
- [11] N. D. Natha, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automat. in Construction*, vol. 112, pp. 1-20, Apr. 2020.
- [12] J. Park, S. Jeon, J. Jeon, and J. Kim, "A study on deep learning based personal protective equipment detection," in *Proc. 2020 The Korean Inst. of Broadcast and Media Eng. Summer Conf.*, pp. 650-651, online, Jul. 2020.
- [13] C. Y. Wang, H. Y. Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, and I. H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proc. 2020 IEEE/CVF CVPRW*, pp. 779-788, Seattle, WA, Jun. 2020.

**박 용 석 (Yong-Suk Park)**



1998년 5월 : 카네기멜론대학교  
전기컴퓨터공학과 석사 졸업  
2018년 2월 : 연세대학교 전기  
전자공학과 박사 졸업  
2003년 12월~현재 : 한국전자기  
술연구원 콘텐츠응용연구센  
터 책임연구원

<관심분야> 이미지 프로세싱, 인공지능, 생체인식  
[ORCID:0000-0002-6694-5125]

**이 경 택 (Kyung-Taek Lee)**



2008년 8월 : 연세대학교 전기전  
자공학과 박사 졸업  
1996년~1998년 : (주)인켈 연구원  
1998년~2001년 : (주)아이앤씨테  
크놀로지 팀장  
2015년~2016년 : Carnegie Mellon  
University, HCII Visiting  
Researcher

2006년~현재 : 한국전자기술연구원 콘텐츠응용연구센  
터 센터장  
<관심분야> XR(eXtended Reality), 디지털트윈, 딥러  
닝

[ORCID:0000-0002-5048-9268]

**이 송 연 (Song-Yeon Lee)**



2017년 2월 : 서울여자대학교  
멀티미디어학과 졸업  
2019년 2월 : 서울여자대학교  
컴퓨터학과 석사 졸업  
2019년 9월~현재 : 한국전자기  
술연구원 콘텐츠응용연구센  
터 연구원

<관심분야> 딥러닝, AI, 이미지 프로세싱  
[ORCID:0000-0002-3119-6099]