

<http://dx.doi.org/10.17703/JCCT.2023.9.5.605>

JCCT 2023–9–75

## 다중 카메라와 객체 탐지를 활용한 건설 현장 사고 감지 시스템

### Accident Detection System for Construction Sites Using Multiple Cameras and Object Detection

김민형\*, 감민성\*\*, 류호성\*\*, 박준혁\*\*, 전민수\*\*, 최형우\*\*, 민준기\*\*\*

Min hyung Kim\*, Min sung Kam\*\*, Ho sung Ryu\*\*, Jun hyeok Park\*\*

Min soo Jeon\*\*, Hyeong woo Choi\*\*, Jun-Ki Min\*\*\*

**요약** 건설 현장의 사고는 중증외상환자가 발생하기 쉬운 특성 탓에 사망으로 이어지는 비율이 매우 높다. 중증외상환자의 사망률을 줄이기 위해서는 빠른 대처가 필요하며, 빠른 사고 대처를 위해 인공지능 기술과 카메라를 이용하여 사고를 감지하는 시스템들이 개발되었다. 그러나 기존 사고 감지 시스템들은 단일 카메라만을 사용하여, 사각지대로 인해 건설 현장의 모든 사고를 감지하기에 한계가 있다. 따라서, 본 논문에서는 다수의 카메라를 사용하여 감지 사각지대를 최소화하는 시스템을 구현하였다. 구현된 시스템은 다수의 카메라의 영상에서 YOLO-pose 라이브러리로 특징점을 추출하고, 추출된 특징점을 장단기 메모리(Long Short Term Memory) 기반 순환신경망에 입력하여 사고를 감지하였다. 실험 결과, 우리는 제안하는 시스템이 복수의 카메라 사용으로 감지 사각지대를 최소화하면서도 높은 정확도를 가지는 것을 확인하였다.

**주요어** : 객체 탐지, 인공지능, 사고 감지

**Abstract** Accidents at construction sites have a very high rate of fatalities due to the nature of being prone to severe injury patients. In order to reduce the mortality rate of severely injury patients, quick response is required, and some systems that detect accidents using AI technology and cameras have been devised to respond quickly to accidents. However, since existing accident detection systems use only a single camera, there are blind spots. Thus, they cannot detect all accidents at a construction site. Therefore, in this paper, we present the system that minimizes the detection blind spot by using multiple cameras. Our implemented system extracts feature points from the images of multiple cameras with the YOLO-pose library, and inputs the extracted feature points to a Long Short Term Memory-based recurrent neural network in order to detect accidents. In our experimental result, we confirm that the proposed system shows high accuracy while minimizing detection blind spots by using multiple cameras.

**Key words** : object detection, AI, accident detection

\*준회원, 한국기술교육대학교 컴퓨터공학과 석사과정 (제1저자) Received: July 30, 2023 / Revised: August 27, 2023

\*\*준회원, 한국기술교육대학교 컴퓨터공학부 학사과정 (참여지) Accepted: September 5, 2023

자)

\*\*\*Corresponding Author: jkmin@koreatech.ac.kr

\*\*\*정회원, 한국기술교육대학교 컴퓨터공학부 교수 (교신저자) Dept. of Computer Science and Engineering, Korea University of Technology and Education, Korea

접수일: 2023년 7월 30일, 수정완료일: 2023년 8월 27일

University of Technology and Education, Korea

게재확정일: 2023년 9월 5일

## I. 서 론

건설업의 발전과 더불어 건설 현장에서의 사망사고 사례 또한 꾸준히 보고되고 있다. 고용노동부의 산업 재해 현황 통계 자료[1]에 따르면 2022년 한해 건설업에서의 사고재해자 수는 27,432명으로 전체 사고재해자 수의 25.6%에 달하며, 사고사망재해의 경우 402건의 사망으로 전체 사고사망재해의 46%에 달하는 것으로 나타났다. 이는 비슷한 사고재해자 수를 가지는 제조업에 비해 두 배에 달하는 사망자 수로, 사고로 인해 중증의 상환자가 발생하기 쉬운 건설 현장의 특성으로 인한 것으로 보인다. [2]에 따르면 이러한 중증외상환자의 처치에 걸리는 시간은 사고 이후 예후에 영향을 미치는 중요 요소로, 외상 발생 후 첫 60분의 ‘골든 타임’이 생존 확률을 결정한다고 한다. 따라서, 사망사고로 이어지는 최악의 사태를 줄이기 위해 빠르게 사고를 인지하고 대처하는 것이 중요하다. 그러나 실제 건설 현장에서는 부족한 관리 감독과 큰 소음 등으로 인해 빠르게 사고를 인지하는 것이 쉽지 않다.

이와 같은 문제 해결을 위해 카메라를 통해 영상을 획득하고 인공지능을 통해 사고 상황을 인식하는 영상 기반 감지 시스템이 개발되었다[3]. 영상 기반 감지 기법은 건설 현장의 카메라를 이용하면 추가적인 장비가 필요하지 않고, 작업자에게 주는 부담이 없어 실제 적용에 유리하다. 영상 감지 기법 중, 장단기 메모리(Long Short Term Memory: LSTM) 기반 순환 신경망(Recurrent Neural Network: RNN)을 사용한 영상 기반 감지 기법[3]의 경우 높은 사고 감지 정확도를 보인다. 그러나 기존 영상 기반 감지 기법은 단일 카메라를 사용하여, 많은 사각지대가 존재하는 건설 현장의 모든 사고를 감지하기에는 한계가 있다.

본 논문에서는 기존 영상 기반 감지 기법이 가지는 사각지대 문제를 해결하기 위해 다수의 카메라를 사용하여 사고 상황을 감지하는 시스템을 제안한다. 제안하는 시스템은 4대의 카메라를 이용하여 영상을 획득하고, 합성곱 신경망 기반 객체 탐지 라이브러리인 YOLO(You Only Look Once)[4]를 통해 객체인 작업자를 탐지한다. 객체 탐지를 통해 탐지된 작업자로부터 특징점(keypoints)과 특징점의 좌표를 추출하고, 추출된 특징점 좌표를 통해 장단기 메모리 기반 순환 신경망을 사용한 인공지능을 사용하여 사고 상태(fall)와 안전 상

태(safe)의 두 가지 상태로 분류한다. 분류 과정에서 사고 상태로 분류되면 메신저 앱의 알림 API를 통해 관리자에게 사고 상태임을 알린다. 알림을 통해 관리자는 작업자의 사고 상태를 신속히 파악하고 대처하여 건설 현장에서의 사상자의 피해를 줄일 수 있을 것으로 기대된다.

## II. 배경 지식

### 1. 인공 신경망

인공 신경망은 사람의 뇌 속 뉴런의 동작 원리를 모방한 학습 알고리즘으로, 여러 계층과 각 계층에 속하는 신경망 셀들로 이루어져 있다. 인공 신경망의 계층은 데이터를 입력받는 입력층(input layer), 입력층과 출력층 사이에 위치하여 데이터가 걸으므로 드러나지 않는 은닉층(hidden layer), 학습된 데이터의 결과값을 출력하는 출력층(output layer)의 3가지 계층으로 구성된다. 이때, 한 계층에 속하는 모든 신경망 셀들이 다음 계층에 속하는 모든 신경망 셀과 연결되면 이를 완전 연결 계층(fully connected layer)이라 한다. 각 신경망 셀은 효과적인 출력값을 만들기 위해 활성화 함수(activation function)를 사용하며[5], 본 논문에서는 ReLU 함수를 사용하였다. ReLU 함수는 값이 양수면 입력값을 출력하는 함수로, 식 (1)과 같이 나타낼 수 있으며 복잡한 연산이 필요하지 않아 빠른 계산속도를 가진다.

$$ReLU(x) = \max(0, x) \dots\dots\dots(1)$$

### 2. 합성곱 신경망

합성곱 신경망이란 특징을 추출 및 분류하는 인공 신경망으로, 객체 탐지나 이미지 분류 등 이미지 분석에 사용되고 있다[6]. 그림 1은 합성곱 신경망의 구조를 간략하게 나타낸 것으로, 합성곱 신경망은 입력된 이미지(input image)로부터 합성곱(convolution) 단계와 풀

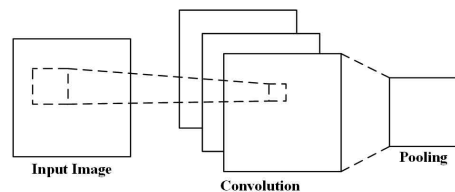


그림 1. 합성곱 신경망 구조도  
Figure 1. Convolution Neural Network structure diagram.

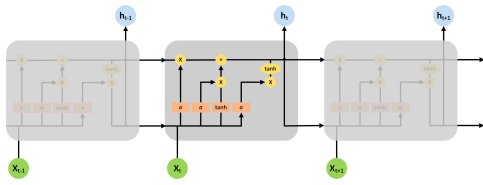


그림 2. 장단기 메모리 (LSTM) 구조도  
 Figure 2. Long Short Term Memory (LSTM) structure diagram

링(pooling) 단계를 반복하며 이미지의 특징을 추출한다. 합성곱 단계는 데이터에서 필요한 특징들을 추출하는 단계로, 데이터와 그보다 크기가 작은 필터를 합성곱 연산하여 특징 맵(feature map)을 추출한다. 풀링 단계는 합성곱 단계를 거친 특징 맵의 크기를 줄이고 특징을 강화하는 과정으로 최대 풀링, 평균 풀링 등의 연산을 사용하여 특징 맵의 요약 통계를 추출한다[7].

## 2. 장단기 메모리(LSTM)

그림 2 는 LSTM 기반 순환 신경망의 구조도이다. LSTM[8]이란 장기간 메모리 역할인 셀 상태와 연결의 강도를 조절하는 입력, 망각, 출력 게이트로 구성된 인공신경망이다. 기존 순환 신경망의 경우 장시간 학습 진행 시, 학습 초기의 결과가 미치는 영향이 작아져 학습 능력이 저하되는 기울기 소실 문제[9]가 발생한다. LSTM는 망각 게이트를 통해 이전 신경망이 현재의 신경망에 미치는 영향을 결정하도록 하여, 기존의 순환 신경망이 가지는 기울기 소실 문제를 해결하였다.

## III. 관련 연구

카메라에서 획득한 영상을 통해 사람의 자세를 예측하는 기법들과, 예측된 자세를 통해 사고 상태로 판단하는 영상 기반 사고 감지 기법들이 연구되었다.

[10]에서는 객체 탐지 라이브러리인 YOLO와 LSTM을 사용하여 사람의 자세를 감지하고, YOLO만 단독으로 사용하였을 때와 LSTM을 같이 사용하였을 때의 정확도를 비교하였다. 비교 결과 YOLO와 장단기 메모리 기반 순환 신경망을 모두 사용하였을 때 더 높은 분류 정확도를 보여주었다.

[11]는 YOLO와 OpenPose라이브러리를 사용하여 건설현장 근로자의 자세를 탐지하는 연구를 수행하였으

며 [12]에서는 적외선 열상 카메라를 사용하여 수집한 비디오 데이터셋을 합성곱 신경망과 장단기 메모리 기반 순환 신경망을 사용하여 사람의 자세를 분류하고, 사고를 감지하였다.

[13]에서는 카메라를 통해 취득한 영상 데이터를 장단기 메모리를 사용하여 5가지 자세로 사람의 자세를 분류하고, 사고 상태를 감지하여 알림을 전송하는 시스템을 개발하였다.

[12,13] 등의 기존 영상 기반 사고 감지 시스템들은 모두 단일 카메라를 이용한 사고 감지 시스템으로, 다양한 물체로 인한 사각지대가 다수 존재하는 건설 현장에서는 감지 범위에 한계가 있다. 또한 여러 사람의 탐지를 고려하지 않아 단일 대상에 대해서만 탐지가 가능한 한계가 있다.

이러한 문제점을 해결하기 위하여 본 논문에서는 다수의 카메라를 이용하여 사각지대를 최소화하고, 객체 탐지를 통해 다수의 부상자가 동시에 발생해도 판별할 수 있는 건설현장 사고 감지 시스템을 제안한다.

## IV. 사고 감지 시스템의 구현

그림 3은 본 논문에서 구현한 시스템의 전체 시스템 구조도를 나타낸 것이다. 본 논문에서 구현한 영상 기반 사고 감지 시스템은 영상 획득, 사고 탐지, 서비스 제공의 3가지 단계로 이루어진다.

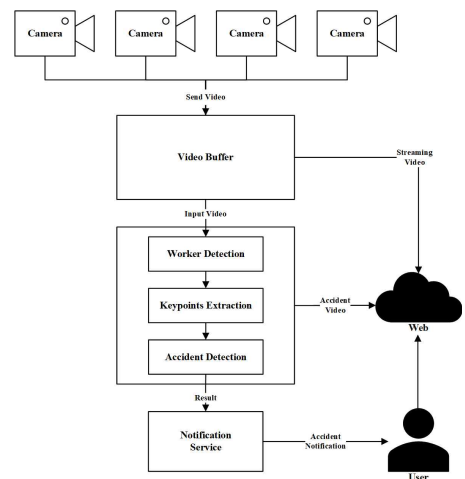


그림 3. 전체 시스템 구조도  
 Figure 3. Overall system structure diagram

1. 영상 획득

영상 획득 단계는 그림 3에서 카메라로부터 영상 이미지가 촬영되어 비디오 버퍼(video buffer)로 전송되는 단계로 사고 감지를 위해 영상을 획득하는 단계이다. 본 논문에서 구현한 시스템에서는 2D카메라를 사용하여 시스템을 구현하였다. 구현한 시스템에서는 획득하는 영상의 사각지대를 최소화하기 위해, 4대의 카메라를 설치하여 사고 감지를 위해 사용하였다. 이렇게 획득된 영상 데이터는 사고 탐지 단계의 입력으로 사용되며 동시에 관리자용 웹 페이지로 전송되어 실시간으로 사용자가 작업자들의 모습을 관찰할 수 있도록 하였다.

2. 사고 탐지

본 단계에서는 영상 획득 단계에서 넘어온 영상 데이터로부터 사고를 탐지하는 단계로 작업자 인식, 특징점 추출, 사고 감지 단계로 구성된다.

1) 작업자 인식

작업자 인식은 주어진 영상 내에서 작업자를 인식하는 단계로, 제안하는 시스템에서의 작업자 인식을 위해 YOLO(You Only Look Once) v5를 사용하였다. YOLO는 합성곱 신경망을 기반으로 한 객체 추출 알고리즘으로, R-CNN[14] 등의 경쟁 알고리즘에 비해 빠른 탐지 속도와 가벼운 계산량을 요구하는 것이 특징이다. 초기 버전의 경우 정확도가 다소 떨어지는 문제가 있었으나, 버전업을 통해 가벼운 계산량을 유지하면서도 준수한 성능을 보이고 있다[15]. YOLO v5는 인공지능 모델의 크기별로 v5n, v5m, v5l, v5x의 라이브러리를 제공하며, 자체 실험을 통해 실시간 감지가 가능하면서도 가장 뛰어난 정확도를 보여준 YOLO v5l를 사용하였다.

2) 특징점 추출

이 단계는 작업자의 자세를 예측하기 위해 입력된 영상으로부터 특징점을 추출하고, 추출된 특징점의 좌표값을 획득하는 단계이다.

특징점은 YOLO v5 기반 라이브러리인 YOLO-pose [16] 라이브러리를 사용하여 추출하였으며, 각 영상으로부터 17개의 특징점을 추출하였다. 추출된 특징점은 x 좌표와 y좌표 34개의 좌표값을 획득하여 인공지능 분류를 위해 사용하게 된다. 획득한 좌표값을 그대로 인공지능 모델에 사용할 경우, 객체의 크기에 따라 좌표값의 크기 차이가 발생하게 된다. 크기 차이가 발생하

는 경우 인공지능 모델의 가중치 설정에 영향을 미치게 되며, 이는 분류 정확도에 악영향을 미치게 된다. 따라서 본 시스템에서는 가중치 반영 정도를 균일하게 만들기 위해 정규화를 수행하였다. 정규화 기법은 식 (2)의 공식을 사용하여 0 ~ 1 사이의 값을 가지도록 설정하는 기법인 최소-최대 정규화를 사용하였다.

$$x = \frac{x - x_{min}}{x_{max} - x_{min}} \dots\dots\dots(2)$$

3) 사고 감지

특징점 추출 단계를 거친 특징점들의 좌표를 인공지능 모델에 입력하고, 작업자의 자세를 분류하여 사고를 감지하는 단계이다.

본 논문에서 구현한 모델에서는 4개 2D 카메라 영상에서 탐지된 객체들의 특징점 좌표를 입력받아 분류를 수행하고 결과를 출력하게 된다. 그림 4는 본 논문에서 사고 감지를 위한 인공지능 모델을 나타낸 것으로, 시계열 데이터인 영상 데이터에 유리한 장단기 메모리 기반 순환 신경망을 사용하여 구성하였다.

인공신경망은 입력층, 은닉층, 출력층의 3단계로 구성되어 있다. 입력층(input layer)은 좌표값의 개수인 34의 크기(hidden\_size)를 가지며, 감시 카메라가 15프레임 정도의 주사율을 가지는 것을 고려하여, 1초에 해당하는 프레임 개수인 15의 길이(seq\_length)를 입력받도

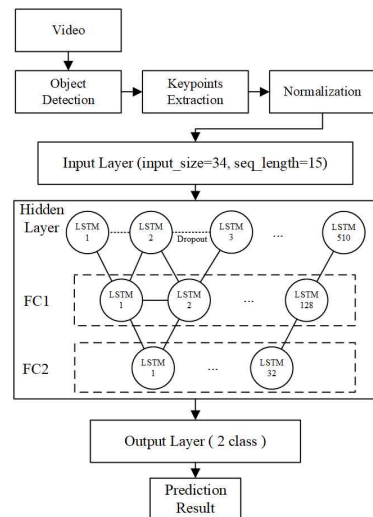


그림 4. 구현 모델 구조도  
Figure 4. Implementation model structure diagram



그림 5. 관리자 페이지 캡처  
 Figure 5. Administrator page capture

록 구성하였다. 은닉층은 ReLU를 활성화 함수로 사용하는 장단기 메모리 기반 순환 신경망(LSTM) 510개가 완전 연결(fully connected: FC)된 장단기 메모리 셀 계층과, 128개의 출력값을 가지는 완전 연결(FC1) 계층, 32개의 출력값을 가지는 완전 연결(FC2) 계층으로 이루어져 있다. 출력층은 ReLU를 활성화 함수로 사용하며, 은닉층의 출력값을 입력받아 사고 상태(fall)와 정상 상태(safe)의 두 가지로 분류하여 결과를 출력한다.

출력된 결과는 물체로 인한 사각지대나 화각으로 인해 작업자를 탐지하지 못하는 경우를 고려하여 4개의 결과 중 하나만이라도 사고로 탐지할 경우 실제 사고 상태로 판단하여 서비스 제공을 위한 단계로 진행하도록 하였다.

### 3. 알림 서비스

사고가 감지되면 이를 관리자에게 알리고, 현장 상황을 확인할 수 있도록 서비스를 제공하는 단계이다 (그림 3의 No. 시스템에서 사고가 감지되었을 경우 메시지의 알림 API를 통해 메시지를 관리자에게 전송하게 된다. 모델의 순간적 오류로 잘못 판단하는 경우를 고려하여, 일정 시간 이상 사고 상태로 감지될 시에만 메시지를 보내도록 설정할 수 있다. 본 논문의 실험에 사용한 모델에서는 7초 이상 사고로 감지하면 실제 사고 상태로 판단하고 사고 알림을 보내도록 설정하였다. 또한 다수의 카메라에서 사고가 감지되면 이를 반영하여 알림을 전송하도록 구현하였다.

그림 5는 본 논문에서 구현한 관리자 페이지를 캡처한 것이다. 관리자 페이지를 통해 실시간 현장 모니터링 및 분류 결과를 확인할 수 있으며, 하위 메뉴를 통해 사고 당시의 영상 또한 확인할 수 있다.

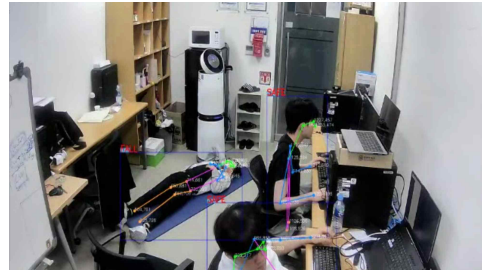


그림 6. 사고 상황 캡처  
 Figure 6. Accident situation capture

그림 6은 사고 발생 상황의 페이지 화면을 캡처한 것으로 사고 발생 시 사고를 감지한 화면을 확대하여 표시되도록 구성되어 있다. 그림 6에서 보는 바와 같이 여러 작업자들을 탐지하고 그 중 일부 작업자가 쓰러져 있을 경우에도 이를 탐지할 수 있다.

## V. 실험 및 결과

### 1. 실험 환경

본 논문에서 구현한 시스템의 운영체제는 Windows 11, CPU는 Intel(R) Core(TM) i9-12900K, GPU는 Nvidia Geforce RTX 3060, 램은 DDR4 32GB 램을 사용하였다. 사용 언어는 Python 3.10.9를 사용하였으며 Yolo v5기반 라이브러리인 Yolo-pose v5와 PyTorch7.0을 사용하였다. 사용 인공지능은 장단기 메모리 기반 순환 신경망을 사용하였으며 활성화 함수로 ReLU, 최적화 알고리즘으로 Adam 알고리즘을 사용하였다.

표 1. 인공신경망 매개변수  
 Table 1. Artificial Neural Network Hyperparameters

Parameter	Value
epoch	1000
input_size	34
seq_length	15
hidden_size	510
dropout	0.3

표 1은 실험을 위해 사용한 인공신경망의 매개변수를 나타낸 것이다. 학습 횟수(epoch)는 1000이다. 입력 크기(input\_size)는 34, 입력 길이(seq\_length)는 15이며 입력 크기와 길이를 고려하여 hidden\_size는 510으로 설정하였다. 추가로 학습 중 임의로 일부 특징을 제거하여 과적합 현상을 최소화하는 기법인 드롭아웃(dropout)

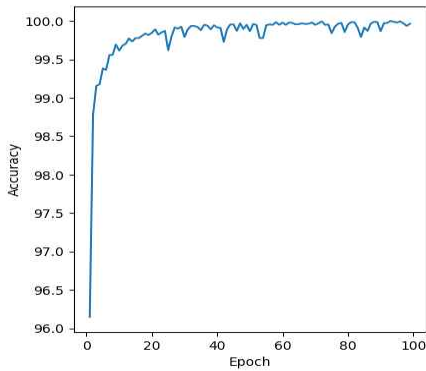


그림 6. 학습횟수에 따른 정확도  
Figure 6. Accuracy with increasing Epoch

ut)[17]을 적용하였으며, 이때 0.3을 매개변수로 사용하였다.

### 2. 사용 데이터셋

데이터셋은 자체 제작한 내부 데이터와 AI hub[18]에서 제공하는 외부 데이터로 구성하였다. 데이터셋의 구성은 사고 상태 129,480프레임, 정상 상태 254,460프레임으로 총 383,940프레임의 데이터를 사용하였다.

본 논문에서 사용하는 데이터셋의 경우 사고 상태 데이터와 정상 상태 데이터의 데이터 수에 차이가 존재하여 데이터 불균형 문제[19]가 발생하게 된다, 이는 분류 정확도에 영향을 미치게 된다. 이를 해소하기 위해 적은 데이터를 가진 클래스의 데이터 수를 많은 데이터를 가진 데이터의 데이터 수에 맞게 늘려주는 기법인 오버샘플링을 적용하였다. 오버샘플링 기법으로는 SMOTE[19] 오버샘플링을 적용하였다. SMOTE 오버샘플링은 적은 데이터를 가진 클래스와, 많은 데이터를 가진 클래스 사이에 소수 클래스에 해당하는 임의의 새로운 데이터를 생성하는 기법이다. 오버샘플링 결과 클래스별 254,460프레임, 총 508,920프레임의 데이터가 되었으며, 8:2 비율로 학습 데이터와 검증 데이터로 나누어 실험을 진행하였다.

### 3. 실험 및 평가

그림 6은 제안하는 모델의 학습 횟수(epoch)에 따른 정확도를 나타낸 그래프이다. epoch 값이 증가에 따라 정확도가 안정적으로 증가하여, 구현한 인공지능 모델의 학습이 정상적으로 이루어지는 것을 확인하였다.

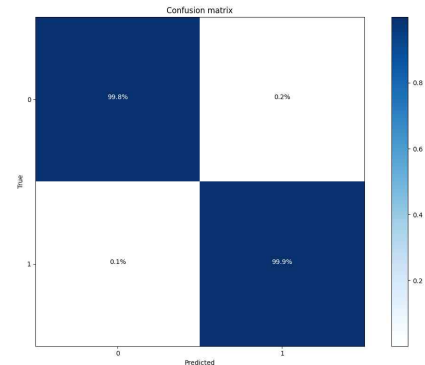


그림 7. 학습 정확도 혼동 행렬  
Figure 7. Confusion matrix

그림 7은 제안하는 모델의 테스트 결과를 혼동 행렬로 나타낸 것으로 세로축은 정답, 가로축은 예측 결과를 나타낸다. 테스트 결과, 정확도는 약 99.8% ~ 99.9%로 나타났다. 실험을 통해, 다수의 카메라를 사용하고, 다수 인원의 사고를 감지하면서도 높은 정확도를 보이는 것을 확인할 수 있다.

## VI. 결 론

본 연구에서는 객체 탐지 라이브러리인 YOLO와 장단기 메모리 기반 순환 신경망을 사용하여, 다수의 카메라로부터 획득한 영상을 통해 건설현장 작업자의 사고를 감지하는 시스템을 제안하였다. 실험을 통해 제안하는 시스템은 다수의 카메라를 사용하고, 다수 인원의 사고를 감지하면서도 높은 정확도를 보여주는 것을 확인하였다. 본 연구를 통해, 기존 영상 기반 사고 감지 기법이 사각지대로 인해 감지하지 못하던 사고를 추가로 감지하여 더욱 많은 사망사고 예방이 가능할 것으로 기대한다.

## References

[1] Ministry of Employment and Labor “Additional Statistics on Industrial Accidents in 2022” 2022. [https://www.moel.go.kr/policy/policydata/view.do?bbs\\_seq=20230100992](https://www.moel.go.kr/policy/policydata/view.do?bbs_seq=20230100992)

[2] J.C. Yang and J.D. Moon, “The effects of prehospital care on on-scene time in patients with major trauma” The Korean Journal of Emergency Me

- dical Services, Vol. 24, No. 1, pp. 67 - 76, 2020. doi: 10.14408/KJEMS.2020.24.1.067n
- [3] S. Jeong, S. Kang, and I. Chun, "Human-skeleton based fall-detection method using LSTM for manufacturing industries" 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), pp. 1-4, 2019 doi: 10.1109/ITC-CSCC.2019.8793342
- [4] J. Redmon, S. Divvala, R. Grishick and A. Farhadi, "You only look once: Unified, real-time object detection" Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016. doi: 10.1109/CVPR.2016.91
- [5] A. K. Jain, J. Mao and K. M. Mohiuddin, "Artificial neural networks: A tutorial" Computer, Vol. 29, No. 3, pp. 31-44, 1996. doi: 10.1109/2.485891
- [6] B.S. Hwang, J.H. Kim, Y.R. Lee, C.U. Kyeong, J.H. Seon, Y.G. Sun, J.Y. Kim. "Performance of Exercise Posture Correction System Based on Deep Learning" The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 22, No. 5, pp. 177-183, 2022 doi: doi.org/10.7236/IIBC.2022.22.5.177
- [7] I.S. Oh, "Machine Learning" Hanbit Academy, 2017.
- [8] S. Alex, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network" Physica D: Nonlinear Phenomena, Vol. 404, pp. 132306, 2020. doi: 10.1016/j.physd.2019.132306
- [9] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions" International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 6, No. 3, pp. 107-116, 1998. doi: 10.1142/S0218488598000094
- [10] J. Lu, M. Nguyen and W.Q. Yan, "Deep learning methods for human behavior recognition" 2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ). IEEE, pp 1-6, 2020. doi: 10.1109/IVCNZ51579.2020.9290640
- [11] Y. Yoon and T. Oh, "A Study on the Improvement of Construction Site Worker Detection Performance Using YOLOv5 and OpenPose" The journal of the convergence on culture technology, Vol. 8, No. 5, pp. 735 - 740, 2022. doi: 10.17703/JCCT.2022.8.5.735.
- [12] Y.W. Lee, J.H. Park, S.Y. Sin, "Implementation of Fall Detection Based on CNN-LSTM" Vol. 47, No. 2, pp. 340-347, 2022. doi:10.7840/kics.2022.47.2.340
- [13] W.H. Choi, C.D. Kwon, B.S. Yoo, M.H. Kim and J.K. Min. "Accident Detection System Based on RNN Exploiting Keypoints and LSTM" KIISE Transactions on Computing Practices, Vol. 29, No. 7, pp. 309-315, 2023. doi: 10.5626/KTCP.2023.29.7.309
- [14] R. Girshick, J. Donahue, T. Darre and J. Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation" Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587, 2014. doi: 10.1109/CVPR.2014.81
- [15] N.J. Kwak, D.J. Kim. "A study on Detecting the Safety helmet wearing using YOLOv5-S model and transfer learning" International Journal of Advanced Culture Technology, Vol.10, No.1, pp. 302-309, 2022 doi: 10.17703/IJACT.2022.10.1.302
- [16] D. Maji, S. Nagori, M. Mathew and D. Poddar. "Yolo-pose: Enhancing yolo for multi person pose estimation using object keypoint similarity loss" Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2637-2646, 2022. doi: 10.1109/CVPRW56347.2022.00297
- [17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. "Dropout: a simple way to prevent neural networks from overfitting" The journal of machine learning research, Vol. 15, No. 1, pp. 1929-1958, 2014.
- [18] AIHub, "Senior Abnormal Behavior Video" 2022. <https://aihub.or.kr/>
- [19] N.V. Chawla, K.W. Bowyer, L.O. Hall and W.P. Kegelmeyer. "SMOTE: synthetic minority over-sampling technique" Journal of artificial intelligence research, Vol. 16, pp. 321-357, 2002. doi: 10.48550/arXiv.1106.1813

※ 이 논문은 2023년도 한국기술교육대학교 교수 교육연구진흥과제 지원에 의하여 연구되었음.