

Swoon Monitoring System based on YOLOv4-CSP Object Detection Algorithm

YOLOv4-CSP 기반 실신 감시 시스템

Jung-woo Chae · Tae-kyeong Kim · Hyun-chong Cho

채정우* · 김태경** · 조현종†

Abstract

When swoon occurs, it is important to detect it immediately. If immediate and quick detection is not possible, additional serious accidents may occur. Therefore, this study proposed swoon monitoring system that can detect swoon from real-time images of CCTV based on the increasing number and installation of CCTV. For the detection of swoon, deep learning based object detection algorithm applied with the Cross Stage Partial connections network model was used. Through this, swoon detection is performed within a single frame, and the amount of computation required for swoon detection is reduced and can have a fast processing speed. In addition, it is possible to have high detection performance through various BoF and BoS techniques with CSP network model. The result of swoon detection and classification performance through object detection algorithm showed 94.2% F1-Score, 92.6% mAP and 95.4% Accuracy.

Key Words

CCTV, Cross Stage Partial Connections, Deep Learning, Object Detection, Swoon

1. 서론

현대 사회의 경제가 발전해나감에 따라 영양 및 위생 상태와 같은 기초적인 보건 수준은 높아졌다. 의료 기술 역시 빠른 발전을 이루었으며, 앞선 변화와 함께 사회적 기대 수명도 빠른 속도로 증가하기 시작했다. 하지만 기대 수명의 증가와 함께 저출산 문제가 더해지면서 전체적인 고령화가 진행되고 있다. UN은 만 65세 이상 노인 인구가 전체 인구의 7%를 넘으면 고령화사회, 14%를 넘으면 고령사회, 20% 이상이면 초고령사회로 정의한다. 통계청에 의하면 현재 대한민국은 2017년부터 노인인구가 14% 이상인 고령사회에 진입했으며, 2025년에는 노인인구가 20% 이상인 초고령사회로의 진입이 예상된다. 이처럼 빠르게 고령화가 진행되면서 고령 인구가 늘어남에 따라 고령 질환으로 인한 입원 및 사망 비율 역시 증가하고 있다. 순환계통 질환은 고령 질환 사망률에 큰 영향을 미치고 있으며, 뇌혈관 질환은 순환계통 질환 중 가장 높은 비율을 차지하고 있다. 뇌혈관 질환은 뇌의 혈관이 막혀서 혈액과 산소가 공급되지 않아 뇌가 일시적으로 정지하는 뇌경색 증 증세를 포함하며, 실신 중 가장 높은 원인 비율을 가지는

미주신경성 실신의 원인이 된다. 이외에도 실신은 심장 질환, 심폐 질환, 신경계 질환 및 기립성저혈압 등 다양한 원인을 통해 발생할 수 있다.

실신의 경우 사건이 발생한 후 즉각적인 확인과 빠른 치료가 중요하다. 치명적인 질환으로 인한 실신의 경우 즉각적인 확인을 통한 원인 질환의 빠른 치료가 필요하며, 다른 경우라도 발견이 늦어진다면 추가적인 사고로 이어질 수 있다. 따라서 실신이 발생했을 때 확인 가능 여부는 매우 중요한 요소라고 할 수 있다. 실신이 발생했을 때 주변에 확인 가능한 사람이 없는 경우 CCTV는 실신을 확인하는 대안이 될 수 있다. 현재 공공기관의 CCTV 설치량 및 증가량은 그림 1과 같이 증가하는 추세를 보인다. 또한, 9,000개의 기업, 4,500명의 개인 사업자에 대해 통계청에서 조사한 CCTV 보유 현황의 경우 표 1과 같이 CCTV가 없는 사업장은 평균 9.1%로 매우 적은 것을 확인할 수 있다. 이처럼 점점 증가하는 CCTV 설치량과 함께 높은 CCTV의 보급률은 사람이 없는 상황에서 실신의 확인에 도움이 될 수 있다. 하지만 모든 CCTV를 사람이 24시간 관찰하는 것은 현실적으로 어렵다고 볼 수 있다. CCTV의 확인을 위해서는 노동력 및 인건비가 소모될 것이며,

† Corresponding Author: Dept. of Electronics Engineering and Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Korea.
E-mail: hyuncho@kangwon.ac.kr
<https://orcid.org/0000-0003-2122-468X>

* Dept. of Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Korea.
<https://orcid.org/0000-0003-0760-9735>

** Dept. of Electronics Engineering, Kangwon National University, Korea.
<https://orcid.org/0000-0001-8896-416X>

Received : Nov. 12, 2021 Revised : Dec. 12, 2021 Accepted : Dec. 23, 2021

Copyright © The Korean Institute of Electrical Engineers

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

CCTV를 확인하는 인원의 피로와 집중력 저하 등 다양한 원인에 따라 상황을 파악하지 못할 가능성도 있다.

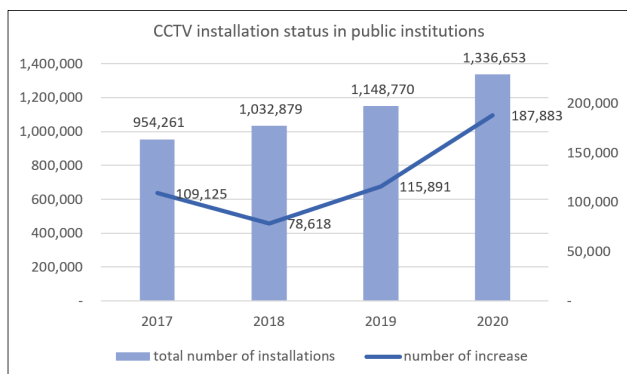


그림 1 공공기관 CCTV 설치 현황 (출처 : 통계청)

Fig. 1 CCTV installation status in public institutions

표 1 기업 및 개인 CCTV 보유 현황 (출처 : 통계청)

Table 1 Percentage of CCTV not installed in industries

Industry	CCTV not installed		
Agriculture and fishery industry	20.2	Information and communication industry	9.8
Manufacturing industry	13.6	Finance and insurance industry	4.7
Construction industry	9.4	Real estate industry	2.3
Wholesale and retail industry	6.4	Technical service industry	13.3
Transportation and warehousing industry	4.9	Facility management and rental service industry	11.6
Accommodation and restaurant industry	12.3	Personal service industry	4.3
Other industry	5.4	Average	9.1

앞선 문제점의 대안이 될 수 있는 실시간 영상을 통해 실신을 감지하는 연구가 진행되기도 하였다[1]. 카메라를 통해 이미지를 입력받고 전처리를 진행한 후, 사람의 높이와 너비를 통계모델의 기준값과 비교해서 실신을 감지하였다. 하지만 일반 카메라가 아닌 높은 비용을 가지는 열화상 카메라를 사용했으며, 열화상 카메라의 특성상 사람의 객체가 여럿이거나 열을 발생하는 다른 객체가 있을 때 실신의 관찰이 어려울 것이 예상된다. 피사체와의 깊이 정보를 측정할 수 있는 키넥트(Kinect) 카메라 기반의 실신 감지 연구도 진행되었다[2]. 해당 연구의 경우 사람의 깊이 영상 정보를 받아온 후 손끝, 팔꿈치, 어깨, 가슴 및 허리 등 사람의 특정한 관절 부위를 검출해서 관절 부위의 위치 정보에 따라 실신을 감지하였다. 하지만 키넥트 카메라의 기술 특성상 사람과의 거리가 멀어지거나, 세로로 서서 관절 부위가 검출될 수 없는 경우, 사람의 일부가 물체에 가려진 경우는 검출이 어려울 것이다. 배경차(Background Substraction) 영상 처리 기법을 통해 영상에서 배경과 사람을 분리한 후, 실신을 검출하는 연구도 진행되었다[3]. 해당 연구의 경우 배경차 기법을 통해 분리된 사람 객체의 길이를 측정했으며, 기하학적 거리에 따라 실신에 대한 길

이 기준을 설정하고 실신의 검출을 진행하였다. 이를 통해 전체 실신 데이터에 대해 평균 약 93%의 Precision 값으로 검출에 성공했으나, Recall 값은 72%의 높지 못한 수치를 보여주었다.

본 연구에서는 앞선 문제점들을 해결하기 위해 CCTV의 실시간 영상을 기반으로 실신의 검출이 가능한 기법을 제안하였다. 연구는 실시간 영상에서 활용되고 실신의 즉각적인 확인이 필요한 특성상 실신 검출에 객체 인식 알고리즘을 사용하였다. 객체 인식 알고리즘은 실신이라는 동작, 즉 행동 인식(Activity Recognition)을 위해 실신 구간의 프레임 연관성을 확인할 필요 없이 실신을 하나의 객체로 정의하여 단일 프레임 내에서 검출이 진행되며, 대다수의 영상 기반 행동 인식 알고리즘과 비교했을 때 상대적으로 낮은 연산량을 통한 빠른 처리 속도를 보여주었다. 또한, 다양한 BoF(Bag of Freebies)와 BoS(Bag of Specials) 기법의 사용과 CSP(Cross Stage Partial connections) 네트워크 구조가 적용된 객체 인식 알고리즘의 사용을 통해 높은 인식 정확도의 실신 검출이 이루어졌다[4].

2. 본 론

본 연구는 AI Hub에서 구축한 검증된 개방 데이터인 실신(Swoon) 영상 데이터를 기반으로 진행됐다[5]. 실신은 사람이 순간적 및 일시적으로 의식을 잃은 후, 자세를 유지하지 못하면서 갑자기 쓰러지는 증상을 나타낸다. 실신은 수원시의 CCTV 관제부서에서 정의한 실생활 속 12종의 이상행동 중 1종으로 정의되어 있으며, 실신 영상 데이터는 앞선 기준들을 토대로 AI Hub에서 구축한 이상행동 CCTV 영상 데이터셋에 포함되어 있다. 또한, 실신 영상 데이터는 장소, 시간, 계절, 날씨, 인원, 성별 및 연령대 등을 고려하여 다양한 상황 및 환경을 가질 수 있는 현실과 최대한 근접한 영상의 제작을 위해 전문 촬영 업체와 협업을 진행했으며, 이를 통해 이상행동을 재현한 CCTV 영상 형태로 구성되었다.

2.1 실신 감시 모델 데이터셋 제작

실신 감시 시스템의 개발을 위한 데이터셋은 앞서 언급된 AI Hub의 실신 영상 데이터를 기반으로 진행되었다. 전체 480개로 구성된 실신 영상 데이터는 실외에서 1인 이상의 사람을 대상으로 쓰러짐과 비틀거림 이상행동 동작을 포함하는 영상 데이터로 구성되었으며, 480개의 실신 영상 데이터는 합계 약 40시간의 영상 길이를 가지고 있다. 또한, 3840(w) x 2160(h)의 해상도와 30(fps)의 프레임 속도를 가지며, 가장 널리 쓰이는 표준 비디오 코덱 중 하나인 MPEG-4 형태로 저장되었다.

전체적인 실신 감시 데이터셋 제작은 그림 2와 같은 순서로 진행되었다. 우선적으로 480개의 실신 영상 데이터는 학습 데이터셋(Train Set) 및 검증 데이터셋(Test Set)으로 분류되었다. 실신 영상 데이터의 분류는 모델 학습을 진행할 때 보편적으

로 사용하는 8 : 2 비율에 근접하도록 학습 데이터셋 380개 및 검증 데이터셋 100개로 구성하였다. 또한, 각각의 데이터셋에 포함된 실신 영상 데이터는 선정의 공정성을 위해 난수를 통해 무작위로 선정하고 분배하였다.

학습 데이터셋 및 검증 데이터셋으로 분배된 실신 영상 데이터는 실신 이상행동이 존재하는 구간에 대한 추출이 진행되었다. 480개의 실신 영상 데이터는 전체 영상 길이 모두가 실신으로 구성된 건 아니며, 각각의 실신 영상 데이터는 일정 구간에서 실신 이상행동이 존재한다. 따라서 AI Hub에서 지정한 실신 이상행동 구간을 기준으로 사용하여 추출을 진행하였다. 추출된 실신 이상행동 구간은 실신 전 비틀거림이 포함된 평균 50초 정도의 길이를 가지고 있다.

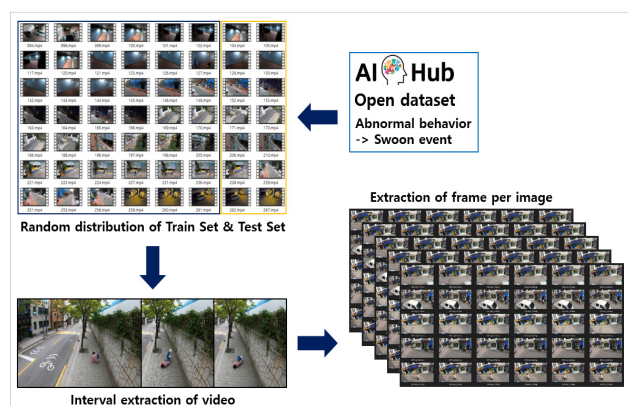


그림 2 실신 감시 데이터셋 제작 과정

Fig. 2 Process of making swoon monitoring dataset

실신 이상행동 구간별로 추출이 끝난 실신 영상 데이터는 실신 감시 모델 학습 및 검증을 위한 프레임 단위의 이미지 데이터로 추출이 진행되었다. 30(fps)의 프레임 속도를 가지는 추출된 실신 영상 데이터는 동작 및 움직임의 일정한 변화를 확인할 수 있는 5프레임 간격의 이미지 데이터로 추출되었으며, 모델 학습 및 검증에 영향을 줄 수 있는 데이터의 손실을 방지하기 위해 무손실 무압축 그래픽 포맷인 BMP 형태로 저장하였다. 3840(w) x 2160(h)의 해상도는 원본인 영상 데이터와 동일하게 유지되었다.

마지막으로 프레임 단위로 추출된 이미지 데이터에 대해 그림 3과 같은 일정한 규칙에 따라 실신의 종료점을 기준으로 실신 감시 모델을 위한 이미지 데이터의 선정이 이루어졌다. 실신의 종료점은 5프레임 간격의 이미지 데이터로 확인하였을 때, 실신이 발생하여 사람의 움직임이 없어지는 이미지 데이터로 설정하였다. 실신의 종료점을 기준으로 앞의 1초 구간을 실신 감시 모델을 위한 이미지 데이터로 사용했으며, 1초 즉 30프레임 길이의 구간은 5프레임 단위의 이미지 데이터로 추출되었고 실신의 종료점 이미지 데이터를 포함해서 1개의 실신 영상 데이터마다 6개의 실신 이미지 데이터가 생성되었다. 실신의 종료점을 기준으로 생성된 프레임 단위 이미지 데이터

는 실신이 발생한 객체에 대해 경계 박스(Bounding Box)를 표시하는 라벨링(Labeling) 작업을 거쳐 저장되었다. 따라서 학습 데이터셋과 검증 데이터셋에 분류된 380개와 100개의 실신 영상 데이터는 2,280장과 600장의 프레임 단위 실신 이미지 데이터로 생성되어 각각의 데이터셋에 추가되었다.

2.2 객체 인식 알고리즘 기반 실신 검출

딥러닝을 기반으로 한 이미지 및 영상 분석 기법은 크게 분류(Classification), 위치 검출(Localization), 객체 인식(Object Detection) 및 개별 객체 분할(Instance Segmentation)로 나눌 수 있다[6]. 분류의 경우 입력으로 주어진 데이터 안에서 단일 객체의 종류를 판단하는 것을 나타내며, 위치 검출은 주어진 데이터에서 단일 객체의 종류를 판단한 후 위치 정보까지 함께 출력한다. 객체 인식은 단일 및 다수의 객체를 대상으로 분류와 위치 검출이 동시에 진행된다. 개별 객체 분할은 객체 인식에서 확인된 객체에 대해 같은 종류의 객체여도 각각 모든 객체의 경계를 정확히 지정하는 과정이 존재한다.

제안된 실신 감시 시스템 연구는 딥러닝을 기반으로 한 이미지 및 영상 분석 기법 중 객체 인식을 사용하여 진행되었다. 실신은 하나의 행동으로 볼 수 있으며, 이미지 및 영상 기반의 행동 인식은 행동이 일어나는 구간에서 인접한 프레임 사이의 연관 관계를 확인하고 움직임 벡터와 같은 특징의 추출 및 분석을 통해 이루어진다. 하지만 영상의 일정 구간에서 프레임 사이의 연관 관계를 확인하고 특징의 추출 및 분석을 통해 수식과 기준을 정의하여 진행되는 행동 인식은 많은 데이터 연산이 필수적으로 요구되며, 빠른 판별을 통한 실시간 처리가 필요한 실신 감시에 적합하지 못할 수 있다. 따라서 제안된 연구에서는 2.1절 같이 실신이라는 행동이 발생했을 때의 동작, 즉 사람의 자세를 하나의 객체로 지정해서 연구를 진행하였다. 실신이라는 행동이 하나의 객체로 지정되어 객체 인식 알고리즘을 통한 검출이 진행될 경우, 실신이라는 행동이 일어나는 구간에서 인접한 프레임 사이의 연관 관계를 확인하고 추가적인 특징과 연산이 사용될 필요 없이 단일 프레임 내에서 실신의 분류와 위치가 검출되는 객체 인식이 진행된다. 따라서 행동 인식의 방법 중 상대적으로 낮은 연산량을 가질 수 있으며, 빠른 판별을 위한 실시간 처리가 요구되는 실신 감시에 있어 강점을 가질 수 있다.

실신 검출을 위한 객체 인식 알고리즘으로는 CSP 네트워크 구조가 적용된 객체 인식 알고리즘인 YOLOv4-CSP가 사용되었다[7]. YOLOv4-CSP는 CNN(Convolutional Neural Network)을 기반으로 하는 대표적인 1-Stage Detector 모델인 YOLO(You Only Look Once)의 한 종류이다[8]. 객체 인식 알고리즘은 크게 1-Stage Detector와 2-Stage Detector로 분류된다. 1-Stage Detector는 입력 데이터의 경계 박스와 분류 예측 가능성(Class Probability)을 하나의 문제로 판단하여 입력 데이터를 한번 읽음으로 객체의 위치 검출과 분류가 진행되며, 위치

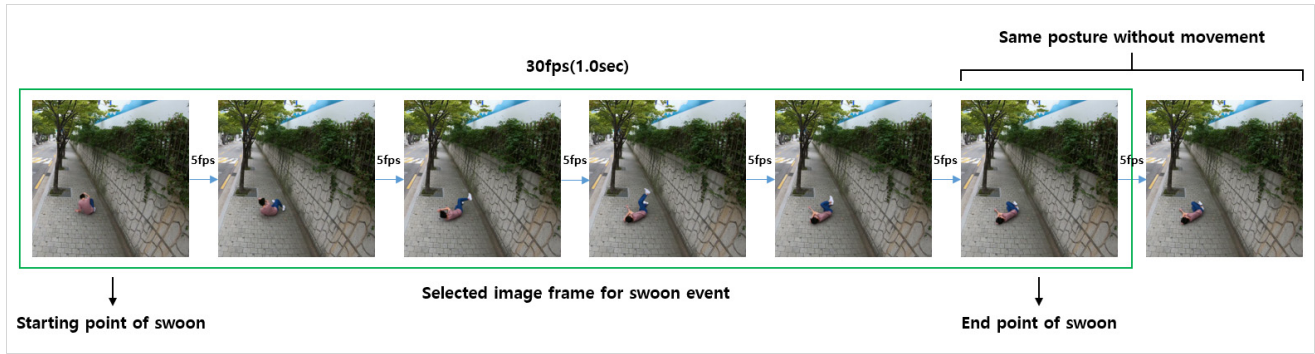


그림 3 실신 종료점 기준 실신 감시 모델 이미지 데이터 선정

Fig. 3 Selection of swoon monitoring model image data based on the end point of swoon

검출과 분류가 순차적으로 진행되는 2-Stage Detector와 비교했을 때 상대적으로 높은 처리 속도를 가진다. 또한, 네트워크 모델은 다양한 BoF와 BoS 기법의 적용을 통해 성능의 향상이 이뤄졌다[9]. CSP 네트워크 구조는 CNN을 기반으로 하는 YOLOv4의 백본망(Backbone Network) Darknet53에 적용되어 CSPDarknet53으로 새롭게 설계되었다[10]. CSP 네트워크 구조는 그림 4와 같이 계층의 분할을 통해 가중치가 갱신될 때 중복된 기울기 정보 사용을 방지해서 정확도의 손실을 줄이는 특성을 가진다. 더욱이 CSP 네트워크 구조가 가지는 cross-channel pooling 기법은 이미지 내 존재하는 다양한 해상도의 같은 객체를 인식하는데 사용되는 FPN(Feature Pyramid Network) 구조에서의 작업을 압축하고 소모되는 메모리 비용을 줄여준다[11].

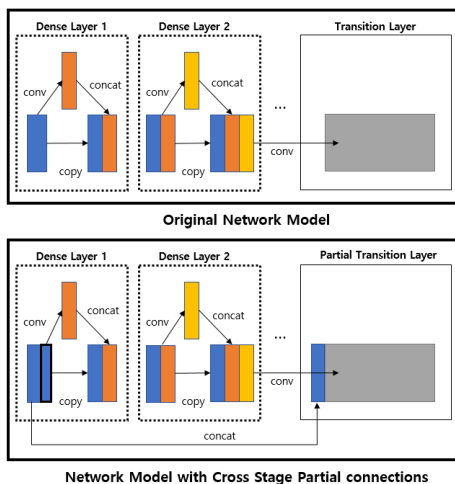


그림 4 CSP 네트워크 모델 구조

Fig. 4 Architecture of CSP network model

이처럼 앞서 언급된 모든 기법의 이점을 가지고 CSP 네트워크 구조를 통해 빠른 속도로 다양한 해상도의 객체에 대해 높은 분별력을 가지는 YOLOv4-CSP가 실신 검출에 있어 높은 강점을 가질 수 있다고 판단했으며, 빠른 처리 속도와 함께 정확한 인식이 중요한 실신 검출에 사용하였다. 또한, 2.1절에

서 제안한 실신의 시작점부터 종료점까지 프레임 단위 이미지 데이터 추출 규칙에 따라 선정된 데이터를 기반으로 학습 및 평가를 진행함에 따라 정확한 규칙과 기준을 가지고 실신 검출이 이뤄질 수 있도록 하였다.

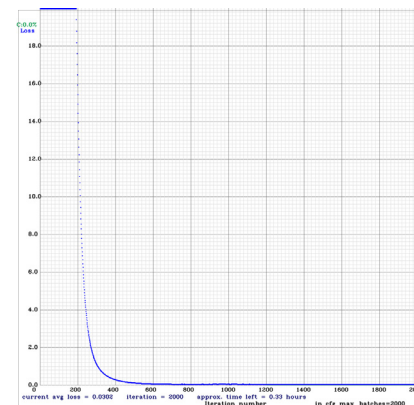


그림 5 모델 학습에 따른 손실 그래프

Fig. 5 Loss rate per iteration in training

표 2 학습에 사용된 기기 및 환경 설정

Table 2 Device performance and environment settings

CPU	Intel(R) Xeon(R) W-2133 CPU @ 3.60GHz
GPU	NVIDIA TITAN RTX
Memory	64.0GB
OS	Windows 10 x64
	CUDA 11.1 with cuDNN
	OpenCV 3.4

2.3 딥러닝 네트워크 모델 학습

실신 검출을 위한 딥러닝 네트워크 모델 학습은 YOLOv4-CSP의 백본망 CSPDarknet53을 통해 이루어졌다. 학습 데이터셋의 프레임 단위 실신 이미지 데이터 2,280장이 네트워크 모델 학습에 사용되었다. 사용된 프레임 단위 실신 이미지 데이터는 네트워크 모델에 640 x 640 입력 크기를 가지고 3개의 채널로 입력되었다. 모델 학습 결과에 영향을 주는 학습률(Learning Rate)의 경우 0.001로 설정하였다. Batch Size는 64로

설정되었으며, 2,000회의 Iteration까지 네트워크 모델 학습을 진행하여 약 50회의 Epoch이 이루어질 수 있도록 하였다. 딥러닝 네트워크 모델의 학습에 따른 손실 함수 그래프는 그림 5와 같으며, 사용된 기기와 환경 설정은 표 2와 같다.

3. 연구결과

연구에서 제안된 CSP 네트워크 구조가 적용된 객체 인식 알고리즘을 통한 실신 검출의 전체적인 과정과 구조는 그림 6과 같다. 실신 검출 성능 평가는 일반 이미지 데이터 600장, 실신 이미지 데이터 600장을 통해 진행되었다. 실신 이미지 데이터 600장은 2.1절에서 제작된 네트워크 모델 학습에 사용되지 않은 검증 데이터셋에 포함된 데이터이며, 일반 이미지 데이터는 검증 데이터셋의 실신 영상 데이터 100개에서 사람은 확인할 수 있지만, 실신 행동이 존재하지 않는 영상 구간 중 무작위로 추출되어 구성되었다.

실신 검출 성능 지표로는 실신 이미지 데이터 600장에 대해 TP(True Positive), FP(False Positive), FN(False Negative), Precision, Recall, F1-Score 및 mAP(mean Average Precision)가 사용되었다. 실신 분류 성능 지표는 실신 이미지 데이터 600장과 일반 이미지 데이터 600장에 대해 TP, TN, FP, FN, Specificity, Sensitivity 및 Accuracy가 사용되었다. 제안된 실신 감시 시스템 성능 평가는 표 3과 같다.

표 3 실신 감시 시스템 성능 평가

Table 3 Performance evaluation of swoon monitoring system

Performance of swoon detection	
All (Swoon)	600
TP (True Positive)	553
FP (False Positive)	21
FN (False Negative)	47
Precision	0.9634
Recall	0.9217
F1-Score	0.9421
mAP (mean Average Precision)	0.9255

Performance of swoon classification	
All	1200
Positive (Swoon)	600
Negative (Normal)	600
TP (True Positive)	553
TN (True Negative)	592
FP (False Positive)	8
FN (False Negative)	47
Specificity	0.9867
Sensitivity	0.9217
Accuracy	0.9542

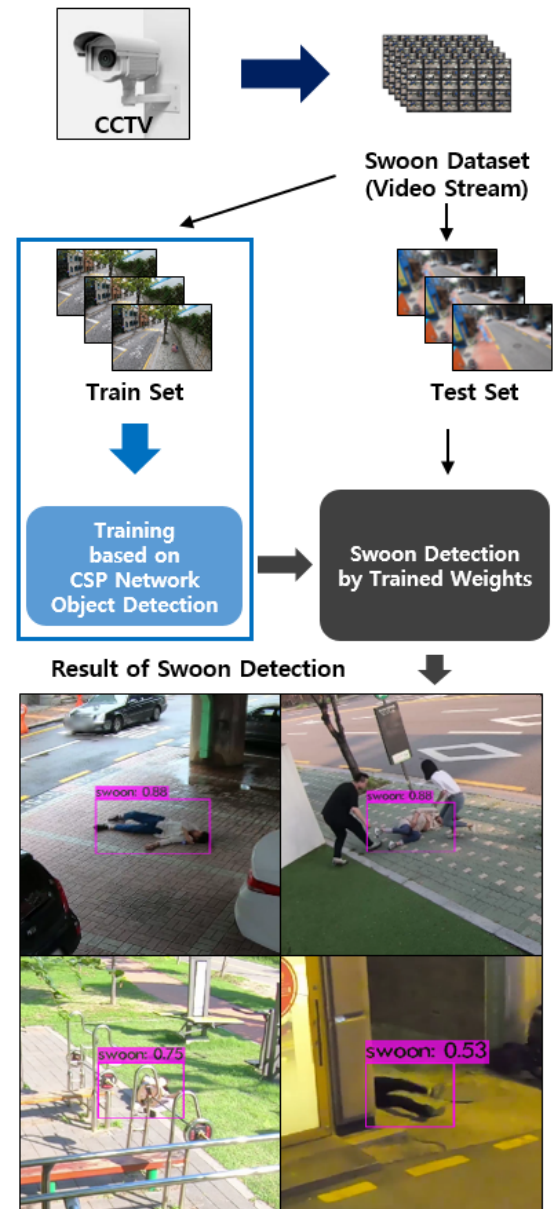


그림 6 객체 인식 알고리즘 기반 실신 검출 과정

Fig. 6 Swoon detecting process based on object detection algorithm

실신 검출 성능에서 Precision 값은 약 96.3%의 높은 수치를 보여주었다. 실신이라는 자세를 하나의 객체로 지정함에 따라 프레임 단위에서 객체 검출은 진행하는 YOLOv4-CSP 알고리즘은 지정된 객체에 대해서만 높은 이해도를 가진다. 더욱이 CSP 네트워크 구조와 함께 적용된 FPN 구조를 통해 다양한 스케일의 지정된 객체를 학습하였고, 이를 통해 지정되지 않은 객체에 대해서 낮은 오탐(Background Error)을 가지며, 높은 Precision 값을 보여준 것으로 판단된다. Recall 값은 약 92.2%로 상대적으로 낮은 수치를 보여주었다. 하지만 전체적인 학습 데이터는 약 2,000장의 많지 않은 수치로 데이터가 추가로 보완될 경우 높은 성능 값을 가질 수 있을 것으로 판단된다. 또한, F1-Score 값은 약 94.2%의 수치를 나타내었으며, IoU(Intersection over Union)를 기준으로 모델의 성능을 확

인할 수 있는 mAP 값은 약 92.6%의 수치를 보여주었다. 일반 이미지 데이터가 포함된 실신 분류 성능에서는 약 98.7%의 높은 Specificity 값을 보여주었는데, 이는 앞서 언급된 것과 같이 지정되지 않은 객체에 대해서 낮은 오탐을 가지는 객체 인식 알고리즘의 특성상 일반 이미지 데이터에 대해서도 낮은 오탐을 보여주어 나타난 결과로 판단된다. Recall 값과 같은 계산식을 가지는 Sensitivity는 앞선 결과와 같은 약 92.2%의 수치를 보여주었으며, 전체적인 성능을 확인할 수 있는 Accuracy의 경우 약 95.4%로 준수한 성능을 보여주었다. 또한, 그림 6의 실신 검출 결과 부분과 같이 다수의 객체가 존재하거나, 절반 이상의 가림이 발생해도 실신이 정상적으로 검출되는 것을 확인할 수 있다. 이를 통해서 전체적으로 준수한 성능을 통해 실신 검출이 이루어짐을 확인할 수 있었다.

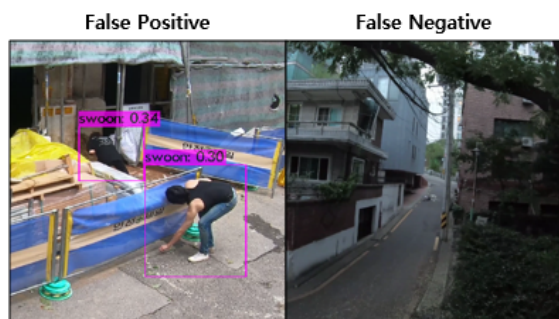


그림 7 실신 검출 오탐 및 미탐

Fig. 7 False positive and false negative of swoon detection

4. 결 론

실신은 발생했을 때의 확인 가능 여부가 매우 중요한 사고 중 하나이며, 즉각적이고 빠른 확인이 불가능하다면 큰 위험으로 다가올 수 있다. 따라서 본 연구에서는 전국적으로 점점 늘어나는 CCTV의 개수의 증가량 및 설치량을 토대로 CCTV의 실시간 영상에서 실신을 검출할 수 있도록 하는 실신 감시 시스템을 제안하였다. 실신의 검출은 객체 인식 알고리즘 YOLOv4-CSP가 사용되었다. 실신이 발생했을 때의 사람 동작을 하나의 객체로 지정해서 행동 인식에 사용했으며, 영상에서 행동 인식을 진행할 때 사용되는 프레임 사이의 연산을 위한 계산량을 줄이고 단일 프레임 내에서 실신 검출을 진행함에 따라 빠른 처리가 가능하도록 하였다. 이를 통해 약 94.2%의 F1-Score와 약 92.6%의 mAP 값을 가지는 실신 검출 성능을 확인했으며, 약 95.4%의 Accuracy 값을 가지는 실신 분류 성능을 확인하였다.

본 연구에서는 YOLOv4-CSP 객체 인식 알고리즘을 통해 실신 검출을 진행했으며, 이를 통해 CSP 네트워크 구조가 적용된 객체 인식 알고리즘이 실신 검출에 있어 적합한 성능을 보여줄 수 있음을 확인하였다. 향후 연구로는 학습에 사용되는 데이터를 추가함으로 더욱 높은 실신 검출 성능을 확보할 수

있는지 확인할 것이며, 딥러닝을 기반으로 하는 다른 객체 인식 알고리즘을 적용해서 성능의 비교 분석을 진행할 것이다. 또한, 그림 7과 같이 발생한 오탐을 줄이고 작은 객체에 대한 미탐을 방지할 수 있도록 사용된 객체 인식 알고리즘의 네트워크 모델 구조 수정을 진행할 예정이다. 마지막으로 연구의 학습 데이터는 전처리 기법이 전혀 이뤄지지 않은 3840(w) x 2160(h) 해상도의 raw-data로 진행되었는데, 여러 BoF 및 BoS 기법과 같이 딥러닝 네트워크 모델의 성능 향상에 도움이 될 것으로 예상되는 전처리 방법들을 연구하고 입력 데이터의 적용을 통해 성능의 향상을 목표로 할 예정이다.

Acknowledgements

Foundation of Korea(NRF) grant funded by the Korea government (MSIT)(No. 2017R1E1A1A03070297). This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2021-2018-0-01433) supervised by the IITP(Institute for Information & communications Technology Promotion).

References

- [1] WANG, Zhen-hai; XU, Bo, "A robust home alone faint detection based on wireless sensor networks," International Journal of Distributed Sensor Networks, 2015.
- [2] GARRIDO, Juan Enrique, et al., "Automatic Detection of Falls and Fainting," J. Univers. Comput. Sci., 2013.
- [3] ELHAMOD, Mohammad; LEVINE, Martin D., "Automated real-time detection of potentially suspicious behavior in public transport areas," IEEE Transactions on Intelligent Transportation Systems, 2012.
- [4] WANG, Chien-Yao, et al., "CSPNet: A new backbone that can enhance learning capability of CNN," In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020.
- [5] AI Open Innovation Hub, <http://www.aihub.or.kr/>
- [6] SZEGEDY, Christian; TOSHEV, Alexander; ERHAN, Dumitru. Deep neural networks for object detection. 2013.
- [7] WANG, Chien-Yao; BOCHKOVSKIY, Alexey; LIAO, Hong-Yuan Mark, "Scaled-yolov4: Scaling cross stage partial network," In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- [8] REDMON, Joseph, et al., "You only look once: Unified, real-time object detection," In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [9] BOCHKOVSKIY, Alexey; WANG, Chien-Yao; LIAO, Hong-Yuan Mark. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.
- [10] RAMACHANDRAN, Prajit; ZOPH, Barret; LE, Quoc V. Searching for activation functions. arXiv preprint arXiv:1710.05941, 2017.

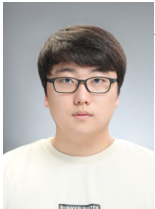
- [11] LIN, Tsung-Yi, et al., Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

저자소개



채정우 (Jung-woo Chae)

Jung-woo Chae received the B.S. degree in Electrical & Electronic Engineering and M.S. degree in Electrical & Medical Convergent Engineering all from the Kangwon National University, South Korea, in 2019 and 2021 respectively. He is currently working towards a Ph.D. degree in Interdisciplinary Graduate Program in BIT Medical Convergence from the Kangwon National University, South Korea.



김태경 (Tae-kyeong Kim)

Tae-kyeong Kim is currently working toward a B.S. degrees in Department of Electronic Engineering from Kangwon National University, South Korea.



조현종 (Hyun-chong Cho)

Hyun-chong Cho received his M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Florida, USA in 2009. During 2010– 2011, he was a Research Fellow at the University of Michigan, Ann Arbor, USA. From 2012 to 2013, he was a Chief Research Engineer in LG Electronics, South Korea. He is currently a professor with the Department of Electronics Engineering and Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, South Korea.