

# 무인 점포 사용자 이상행동을 탐지하기 위한 지능형 모션 패턴 인식 알고리즘<sup>☆</sup>

## Intelligent Motion Pattern Recognition Algorithm for Abnormal Behavior Detections in Unmanned Stores

최 영 준<sup>1</sup>                      나 지 영<sup>1</sup>                      안 준 호<sup>1\*</sup>  
Young-june Choi              Ji-young Na                      Jun-ho Ahn

### 요 약

최근 최저시급의 가파른 인상으로 인건비에 대한 부담이 늘어남과 함께 코로나19의 여파로 무인 상점의 점유율이 높아지고 있는 추세이다. 그로 인해 무인 점포를 타겟으로 하는 도난 범죄들도 같이 늘어나고 있어 이러한 도난 사고를 방지하기 위해 Just-Walk-Out 시스템을 도입하고 고비용의 LIDAR 센서, 가중치 센서 등을 사용하거나 수동으로 지속적인 CCTV 감시를 통해서 확인하고 있다. 하지만 이런 고가의 센서를 많이 사용할수록 점포 운영에 있어 비용 부담이 늘어나게 되고, CCTV 확인은 관리자가 24시간 내내 감시하기 어려워서 사용이 제한적이다. 본 연구에서는 이런 센서들이나 사람에 의지하는 부분을 해결할 수 있고 무인점포에서 사용할 수 있는 저비용으로 도난 등의 이상행동을 하는 고객을 탐지하여 클라우드 기반의 알림을 제공하는 인공지능 영상 처리 융합 알고리즘을 제안하고자 한다. 또한 본 연구에서는 mediapipe를 이용한 모션캡처, YOLO를 이용한 객체탐지 그리고 융합 알고리즘을 통해 무인 점포에서 수집한 행동 패턴 데이터를 바탕으로 각 알고리즘들에 대한 정확도를 확인하며 다양한 상황 실험을 통해 융합 알고리즘의 성능을 증명했다.

☞ 주제어 : 무인 점포, 센서, 저비용, 이상행동, 탐지, 융합 알고리즘

### ABSTRACT

The recent steep increase in the minimum hourly wage has increased the burden of labor costs, and the share of unmanned stores is increasing in the aftermath of COVID-19. As a result, theft crimes targeting unmanned stores are also increasing, and the "Just Walk Out" system is introduced to prevent such thefts, and LiDAR sensors, weight sensors, etc. are used or manually checked through continuous CCTV monitoring. However, the more expensive sensors are used, the higher the initial cost of operating the store and the higher the cost in many ways, and CCTV verification is difficult for managers to monitor around the clock and is limited in use. In this paper, we would like to propose an AI image processing fusion algorithm that can solve these sensors or human-dependent parts and detect customers who perform abnormal behaviors such as theft at low costs that can be used in unmanned stores and provide cloud-based notifications. In addition, this paper verifies the accuracy of each algorithm based on behavior pattern data collected from unmanned stores through motion capture using mediapipe, object detection using YOLO, and fusion algorithm and proves the performance of the convergence algorithm through various scenario designs.

☞ keyword : Unmanned Stores, Sensors, Low Costs, Abnormal Behaviors, Detect, Fusion Algorithm

## 1. 서 론

최근 최저시급 인상으로 인해 늘어난 인건비 부담과 코로나19의 여파로 인해 무인 점포들의 점유율이 늘어나고 있다. 한국에서 CU, GS25, 세븐일레븐, 이마트24 총 4곳의 무인 점포 개수는 2019년 208개에서 2022년 2783개로 약 15배 증가했다[1]. 무인 점포의 특성상 감시하는 인력이 상주하지 않는 특성 때문에 절도 뿐만 아니라 시설 훼손 등과 같은 생각지 못한 사고들이 늘어나고 있다.

<sup>1</sup> AI Computer Science and Engineering, Kyonggi University, Suwon, Gyeonggi-do, 16227, South Korea.

\* Corresponding author (Junho Ahn, jha@kyonggi.ac.kr)

[Received 7 September 2023, Reviewed 16 September 2023(R2 6 November 2023), Accepted 10 November 2023]

<sup>☆</sup> This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)(2021-0-01393)

이러한 사고와 관련해서 현재 뚜렷한 해결책이 나오지 않고 수동적이 감시나 고비용의 센서를 이용해서 제한적으로 대처하고 있다. 대형 무인 점포에서는 진열대에 각종 센서를 설치해 무게에 따라 물건을 가져갔는지 확인하고, 카메라를 사용해 무인 점포 이용자 각각을 추적해 절도가 일어날 상황에 대비하고 있다. 하지만 2018년 문을 열었던 ‘아마존 고’ 매장의 8곳은 폐점을 앞두고 있다. 그 이유에는 매장을 운영하는 데 드는 비용보다 유지하는 데 드는 비용이 더 많이 나가 점포 운영에 있어 비용적 부담이 증가했다는 점이 있다. 소형 무인 점포에서 절도 사건이 발생했을 때는 알아차리기 쉽지 않고 소액인 경우가 많다. 절도 사실을 알았다고 하더라도 점주가 직접 CCTV를 돌려보며 해당 사실을 찾아야 하기 때문에 시간이 지난 뒤 신고를 하게 된다면 절도범을 찾기 어렵다. 또한 앞에서 언급했듯이 도난 사고를 대비하기 위해 각종 센서와 같은 수많은 센서를 사용하는 것은 소형 점포에서 감당하기 어렵다.

본 연구는 이러한 무인 점포 운영의 어려움을 인지하고 모션 캡처 알고리즘과 객체 인식 알고리즘을 융합해서 무인점포 고객의 이상행동을 탐지하는 알고리즘을 제안한다. 제안된 알고리즘을 위해 모션 캡처 모델과, 객체 인식 모델, 융합 모델을 통해서 이상징후를 탐지하는 알고리즘을 개발했다. 모션 캡처 모델로 *mediapipe* 알고리즘을 활용해서 사람의 관절점을 탐지해 서있는 자세, 훔치는 모션, 달리는 자세, 의자에 앉거나 주저앉기, 움직임을 인식하는 모델을 개발했다. 객체인식 알고리즘으로 사람의 형태 및 손모양에 대한 모델을 개발했다. 이를 융합하여 모션캡처에서 발생하는 탐지오류를 줄여서 정확도를 높이는 알고리즘을 개발했다. 이를 통해, 무인 점포에서 발생하는 다양한 이상행동인 뛰거나 훔치거나 바닥에 주저 앉는 등의 개별 알고리즘의 및 융합 알고리즘의 성능을 검증하였다.

2장에서는 해외 및 국내의 무인점포 운영에 사용된 기술과 객체 인식, 모션 캡처에 관련된 알고리즘들에 대해 서술하였다. 3장에서는 제안된 알고리즘인 모션 패턴, 객체 인식, 융합 알고리즘에 대해 서술했다. 4장에서는 해당 개별 알고리즘에 대한 성능을 실험 및 분석했으며, 융합 알고리즘 및 상황에 따른 검증을 수행하였으며 마지막으로 논문의 결론으로 마무리하였다.

## 2. 관련 연구

### 2.1 해외 사례

‘아마존(amazon)’은 시에틀에 무인 식료품점인 ‘아마

존 고(amazon go)’를 열었다[2]. 아마존 고는 just walk out 기술이 적용된 매장으로 컴퓨터 비전, 딥러닝 알고리즘, 센서 퓨전, 자율주행 기술을 사용했다. 걸어 나가기만 하면 된다는 뜻으로 매장에서 물건을 집어 나가면 결제가 되는 서비스이다[3]. 소비자는 아마존 고 앱에서 회원가입을 하고 결제할 신용카드의 정보를 입력하면 입장할 수 있는 QR코드를 부여받는다. 매장 입구에서 QR코드를 찍고 들어가 원하는 물건을 집어서 나오면, 자동으로 결제가 된다. 카메라와 센서로 소비자 각각을 추적해 구별하고, 가상의 장바구니를 생성해 소비자가 구매하려는 물품들의 목록을 작성한다. 사용자가 집은 물건은 가상의 장바구니에 담기는데 다시 내려놓는 경우에도 반영이 된다. 유동인구가 많은 매장에서는 여러 변수들이 발생할 수 있는데, 아마존 고에서는 여러 기술을 융합 사용하여 이를 해결했다. 아마존은 소비자가 들어오는 순간부터 나가는 순간까지 추적하는데, 많은 사람이 몰려 특정 사람이 낮은 신뢰도로 인식되면 신뢰도가 낮은 사람을 표시해 나중에 다시 식별되도록 한다[4]. 또한 일반적인 소비자들과 다르게 행동하는 직원을 구분한다. 위치 추적 카메라는 위에서 아래로 촬영해 이미지에서 소비자의 팔과 물건 사이의 각도를 측정한다. 영상이 평면적이기 때문에 어려움이 있었다. 이를 해결하기 위해 영상의 사람을 막대 형태로 표현해 관절 간의 각도를 계산하기 쉽게 했다. 소비자가 물건을 집었을 때 선반에서 물품이 꺼내지므로 빈 곳이 생긴다. 다시 내려놓으면 빈 공간이 채워지므로 이를 계산하면 되지만, 물건을 밀어서 집어넣었을 때도 빈 곳이 생기므로 선반 위의 모든 항목을 고려해야 한다.

### 2.2 국내 사례

2019년 국내에서 아마존 고와 비슷한 방식으로 운영되는 이마트24 김포dc점이 운영을 시작했다. 출입 인증기를 통해 입장한다. 컴퓨터 비전과 센서 융합 등의 기술을 사용해 매장을 운영하며 여러 대의 카메라가 고객을 인식하여 추적한다. 무게 센서를 사용해 물품의 이동 여부를 알 수 있다[5]. 2021년 이마트24 코엑스점에서는 김포dc점과 동일하게 운영하는데, 6대의 라이다 카메라를 추가로 설치해 매장 운영의 정확도를 높였다[6].

최근 아이스크림 가게나 편의점등 무인으로 운영하는 점포를 많이 찾아볼 수 있다. 점포에서는 보안을 위해 사각지대 없이 CCTV를 배치하기도 하고 카드 인식후 매장에 들어갈수 있는 출입인증 기계를 사용하기도 한다. CCTV 설치의 경우 무인이기 때문에 CCTV가 있더라도

무인으로 운영되기 때문에 범위가 일어났을 때 바로 대응하기 어려워 물건을 훔쳐가는데 어려움이 없다. 출입인증 기계 설치의 경우에도 점포 이용자가 문을 닫지 않고 갔을 때 누구나 인증 없이 들어갈 수 있고, 앞사람의 바로 뒤에 따라 들어가게 된다면 카드인식을 피할수 있다.

### 2.3 모션패턴 인식을 위한 알고리즘

본 연구에서는 앞에서 서술한 아마존 고와 이마트24에서 적용하고 있는 도난 방지 시스템의 허점과 무인 점포 운영에 있어 비용적 부담을 줄여주기 위해 사용자의 모션패턴 인식 알고리즘을 제안한다. 이를 위해 객체 인식, 모션캡처 알고리즘들을 찾아보고 연구를 진행하였다.

우선 무인 점포 내 사용자 객체를 인식할 수 있는 ‘객체 탐지 알고리즘’이 필요하다. 점포 내 사용자는 여러 명이기 때문에 ‘다중 객체’에 대한 인식에 최적화된 알고리즘을 고려하였다. ‘RCNN’ 알고리즘은 bounding box를 기준으로 객체를 구별하는 대표적인 알고리즘이다. 입력된 이미지를 여러 장으로 분할하고 CNN 기반 모델을 이용해 이미지를 분석한다. 이미지의 특정 픽셀들을 잡아 다양한 모양의 bounding box들을 추출해 객체를 검출한다[7]. ‘YOLO’ 알고리즘 또한 CNN 기반의 객체 검출 모델로 RCNN과 마찬가지로 bounding box를 이미지에서 추출해 모델을 학습하고 가중치를 연산해 이를 기준으로 객체를 탐지한다[8].

다음으로 사용자의 행동 패턴을 분석할 수 있는 알고리즘이 필요하다. 본 연구에서는 사용자의 ‘관절들의 움직임’을 바탕으로 행동 패턴을 분석 가능한 알고리즘들을 고려하였다. ‘deeplabcut’ 알고리즘은 행동을 취하고 있는 동물의 자세를 추정하는 프레임워크로 심층 신경망을 기반으로 다중 객체의 자세 추정이 가능하다[9]. 하지만 이는 동물에 대한 자세 추정에 최적화되어 있다는 단점이 있다. ‘openpose’ 알고리즘은 Carnegie Mellon University에서 발표한 C++기반의 오픈소스 라이브러리로 사람의 몸체, 손 등의 관절 포인트들을 추출할 수 있다[10]. 실제로 openpose 라이브러리를 이용해 손에 대한 관절점을 추출해 학습하여 수화를 인식하는 연구들이 많이 진행되었으며 이와 비슷한 ‘mediapipe’ 프레임워크도 존재한다.

## 3. 제안 알고리즘

### 3.1 다중 객체 탐지 알고리즘

이번 단락에서는 무인 점포 내 모션 탐지를 위해 사용

한 알고리즘에 대해서 설명한다.

우선 무인 점포 내 사용자 객체를 탐지하기 위해 ‘YOLO’ 알고리즘을 사용하였다. YOLO는 CNN 기반의 object detection 모델로 새로운 데이터를 쉽게 학습하고 모델이 무겁지 않아 실시간으로 객체를 탐지하기에 적합하다. 또한 YOLO는 각 객체의 유일성을 보존해준다는 특징을 가지고 있어 무인 점포와 같이 수많은 사용자들이 밀집되어 있는 공간에서 각 객체를 탐지하기에 유용하며 이미지를 여러 장으로 분할하여 분석하는 RCNN과 달리 이미지를 한번만 보고 기존 RCNN보다 6배 좋은 성능을 보여주어 보다 빠르게 실시간으로 탐지할 수 있다 [11].



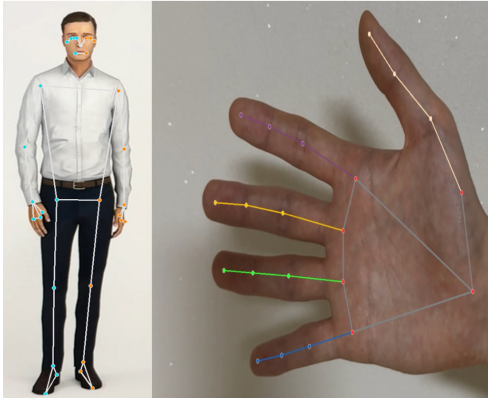
(그림 1) YOLO를 이용한 다중 객체 탐지  
(Figure 1) Multiple object detection using YOLO

다음 그림 1과 같이 YOLO를 이용하여 bounding box 내 객체를 탐지할 수 있고 객체가 아무리 많아도 각각을 구분하여 bounding box를 설정해 실시간으로 탐지 결과를 확인할 수 있다.

### 3.2 모션 캡처 알고리즘

YOLO를 이용해 무인 점포 내 사용자 및 손 객체를 탐지했다면 다음으로 해당 객체의 행동 패턴을 탐지하는 알고리즘이 필요하다. 본 연구에서는 객체의 행동 패턴을 각 관절의 움직임을 이용하여 분석하기 위해 앞에서 ‘deeplabcut’, ‘openpose’, ‘mediapipe’ 등의 다양한 알고리즘들을 찾아보았고 ‘사람 객체의 관절을 탐지할 수 있는가’와 ‘실시간으로 빠르게 패턴을 탐지할 수 있는가’의 기준을 모두 충족시킬 수 있는 ‘mediapipe’ 알고리즘, 즉 프레임워크를 선택하였다. 또한 해당 프레임워크는 C++, Python 등의 다양한 언어로 개발이 가능하기 때문에 활용도가 높고 속도가 빠르다[12, 13]. mediapipe는 사람의 얼굴, 몸체, 손 등의 객체에 Land Mark, 즉 관절점을 찾아주어 탐지한 객체의 움직임을 보여주는 프레임워크 중 하

나이대[12, 13]. 본 연구에서는 그 중 사람 몸체와 손 객체의 관절점을 이용하여 무인 점포 내 사용자 객체의 행동 패턴을 분석하고 탐지하도록 구현하였다.



(그림 2) mediapipe를 이용한 객체 탐지  
(Figure 2) Object detection using mediapipe

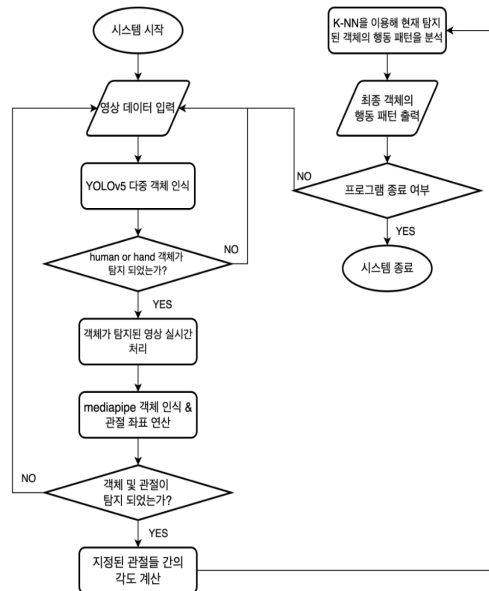
다음 그림 2과 같이 사람의 몸체는 총 33개의 관절점을 제공하고 손 객체는 총 21개의 관절점을 제공한다[12, 14]. 이는 bounding box만 제공하는 객체 인식 알고리즘과 같이 사용하여 보다 정확하게 객체의 움직임 자체를 탐지하는 것이 가능하다.

### 3.3 행동 패턴 학습을 위한 KNN 지도 학습

본 연구에서는 실시간 영상 처리를 위해 ‘OpenCV’ 알고리즘을 사용한다. 또한 모션 캡처 알고리즘을 이용한 행동 패턴을 분석하기 위해서는 패턴에 대한 모델이 필요하고 학습이 필요하다. 그리하여 본 연구에서 제시하는 시스템은 행동 패턴 모델 학습을 위해 python OpenCV 라이브러리에 내장된 ‘KNN 알고리즘’을 이용하여 모델을 학습시킨다. KNN은 지도 학습 방법 중 하나로 주변의 가장 가까운 K개의 데이터를 인지하고 학습시킬 데이터가 속할 군집이 어디인지를 판단해주는 ‘분류’ 알고리즘이다[15]. 또한 이러한 분류 알고리즘은 현재 다양한 분야에서 활용되고 있으며 여러 학습 방식 중에서 가벼운 편에 속한다[15]. 본 알고리즘에서는 모션 캡처 알고리즘에서 제공하는 관절점을 이용해 각 관절마다 각도 값을 계산하여 데이터셋을 만들고 이를 KNN 알고리즘에 적용시켜 인접한 3개의 데이터를 인지하고 학습시킬 데이터에 적합한 군집을 판단하도록 하였다.

### 3.4 인공지능 융합 알고리즘

이 단락에서는 시스템의 순서도를 통해 전체적인 시스템의 흐름을 설명하고 각 단계에서 알고리즘들의 역할을 서술한다.



(그림 3) AI 영상 인식 융합 알고리즘 Flow Chart  
(Figure 3) AI Image Recognition Fusion Algorithm Flow Chart

그림 3와 같이 우선 시스템에 영상 데이터를 입력받는다. 해당 영상 데이터는 객체 인식 알고리즘을 통해 사람 또는 손 객체가 존재하는지 판단하고 탐지하게 된다. 사람 또는 손 객체의 존재가 인식되지 않았다면 계속해서 영상 데이터를 다시 입력받는다. 이후 객체 인식 알고리즘을 통해 객체가 탐지된 영상은 영상 처리 알고리즘을 이용해 실시간 처리를 진행하게 되고 해당 영상에서 객체가 존재하는 bounding box 부분만을 추출한다. 모션 캡처 알고리즘은 추출된 bounding box 부분만을 보고 객체를 인식해 관절점을 생성한 뒤, 각 관절의 각도 값을 연산하게 된다. 특정 객체의 관절점과 각 관절들의 각도 값이 탐지되지 않았다면 다시 처음으로 돌아가 영상 데이터를 입력받는다. 연산이 완료된 각도 값들은 KNN 알고리즘으로 학습된 데이터셋을 통해 적합한 군집을 찾고 해당 군집의 라벨링 값을 구해 최종 객체의 행동 패턴 데이터를 출력하게 된다.

그림 4는 본 연구에서 제안한 AI 영상 인식 융합 알고리즘을 실행한 결과 화면이다. 첫 번째로 입력된 영상 데이터에서 사람 객체가 탐지된 것을 볼 수 있다. 그 다음 영상 처리 라이브러리 등을 활용해 bounding box 부분만 추출되었고 마지막으로 모션 캡처 알고리즘과 KNN 알고리즘을 통해 행동 패턴을 분석하고 최종적으로 ‘STEAL\_1’이라는 라벨링이 출력된 것을 볼 수 있다.



(그림 4) 객체인식과 모션캡처를 이용한 움직임 탐지  
(Figure 4) Moving behavior detection fusing object recognitions and motion captures

## 4. 실험 결과 및 분석

### 4.1 실험 환경 및 실험 설계

본 연구에서는 실제 무인 점포와 다수의 사용자 객체가 존재하는 수많은 점포에서 실험을 진행하였고 각 알고리즘 별로 객체 탐지와 행동 패턴 탐지에 대한 정밀도와 정확도 등을 확인하였다.

#### 4.1.1 객체 인식 알고리즘에 대한 실험 결과

우선 객체 인식 알고리즘에 대한 데이터 수집과 실험을 진행하였다. 본 연구에서 제시하는 시스템 내 YOLO는 사람 또는 손 객체를 탐지하는 역할을 수행하므로 사람과 손에 대한 데이터가 필요하다. ‘roboflow’는 bounding box와 라벨링이 완료된 여러 이미지 데이터를 제공하고 데이터 파일을 YOLO, Pascal, COCO 등 다양한 포맷으로 제공해주는 사이트이다[16]. 본 연구에서는 roboflow에서 human object dataset과 hand object dataset을 ‘YOLO v5 PyTorch’ 형태의 포맷으로 변환하여 학습하였고 학습이 완료된 상태에서 실시간 객체 탐지 실험을 진행하였다. 점포 내 사람 및 손 객체를 탐지하는 객체 인식 알고리즘

의 recall, precision 및 F1-Score는 Table 1과 같다.

(표 1) 객체 인식 알고리즘 재현율, 정밀도 및 F1-점수  
(Table 1) YOLO Object Detection Recall, Precision and F1-Score

Object	Recall	Precision	F1-Score
Human object detection	0.90	0.9783	0.9375
Hand shape detection	0.86	0.9556	0.9053

#### 4.1.2 모션 캡처 대한 실험 결과

본 연구에서는 모션 캡처 알고리즘을 이용한 사용자 행동 패턴 탐지에 대한 실험을 진행하였다. 실제 무인 점포나 다중 객체가 존재할 수 있는 다수의 점포에서 카메라 촬영을 통해 행동 패턴에 대한 데이터를 수집하였고 촬영하지 못한 특정 패턴에 대해서는 ‘AI-Hub’의 이상행동 영상 데이터를 일부 사용하였다[17]. 크게 ‘standing’, ‘runing’, ‘siting’, ‘stealing’, ‘catching’, ‘grabing’ 등의 행동 패턴 데이터를 수집하였다. 촬영한 영상에서 특정한 행동 패턴을 취하는 사용자 객체를 선별한 후, 모션 캡처 알고리즘을 이용해 객체의 각 관절들의 각도 값을 연산하여 라벨링을 포함하여 데이터셋을 구성하였다. 모션 캡처 알고리즘의 행동 패턴 별 recall, precision 및 F1-Score는 Table 2와 같다.

(표 2) 모션캡처 행동 패턴 탐지 재현율, 정밀도 및 F1-점수  
(Table 2) Motion capture behavior pattern detection recall, precision and f1-Score

Pattern	Recall	Precision	F1-Score
Standing	0.92	0.8846	0.9020
Stealing behavior	0.90	0.9184	0.9090
Running	1	0.8772	0.9346
Siting (chair)	0.96	0.8889	0.9231
Siting (ground)	0.86	0.9556	0.9053
Catching	0.92	0.8519	0.8846
Grabing	0.96	0.9231	0.9412
hand (poket)	0.82	0.8039	0.8119

### 4.1.3 융합 알고리즘에 대한 실험 결과

앞에서 객체 인식 알고리즘에 대한 실험 결과와 모션 캡처에 대한 실험 결과에 대해 서술하였다. 점포 내에서 사람이나 손 객체가 다른 객체에 가려지지 않은 상태에서는 비교적 좋은 실험 결과를 얻었다. 하지만 다른 객체에 일부가 가려진 상태이거나 조명과 거리 등의 영향으로 객체가 잘 보이지 않는 상태에서는 모션 캡처 알고리즘만으로는 탐지가 불안정한 모습을 일부 보였다. 본 연구는 앞에서 인공지능 융합 알고리즘을 제안하였고 해당 알고리즘에 대한 실험을 진행하였다. 결과적으로 객체 인식 알고리즘과 영상 처리 알고리즘을 사용해 객체가 존재하는 영역만 추출해 모션 캡처 알고리즘을 적용하면 일부가 가려진 객체 등의 탐지 정확도가 향상되었다. 결과는 Table 3과 같다.

(표 3) 융합 알고리즘 재현율, 정밀도 및 F1-점수  
(Table 3) Fusion algorithms recall, precision and f1-Score

Recall	Precision	F1-Score
0.91	0.9802	0.9438

### 4.2 상황에 따른 실험 결과

본 연구에서 제시하는 알고리즘은 객체 인식 알고리즘과 모션 캡처 알고리즘, 영상 처리 알고리즘을 융합한 융합 알고리즘으로 무인 점포 내 사용자 행동 패턴을 탐지한다. 본 연구는 무인 점포 내에서 일어날 수 있는 일부 상황을 제시하고 모션 캡처 알고리즘만 사용하였을 때와 알고리즘을 융합하여 사용하였을 때의 탐지 성능 실험하였으며 각 상황에서 원하는 행동 패턴을 탐지를 분석하였다. 결과는 Table 4과 같다. 모션 캡처 단일 알고리즘만 사용한 경우에 탐지가 불안정한 상황이 존재했는데 특히 물체에 가려진 객체나 중첩되어 있는 객체들은 정확한 탐지가 어려웠다. 하지만 객체 인식 알고리즘을 함께 사용하여 객체를 확실히 탐지한 뒤 해당 객체만 보여주는 영상 데이터로 다시 행동 패턴을 분석하였을 때, 대부분의 까다로운 상황에서 탐지 가능한 것을 확인할 수 있었다.

(표 4) Mediapipe만 검출 가능 및 검출 불가 및 융합 알고리즘

(Table 4) Detectable and Undetectable of Only mediapipe and Fusion Algorithm

Situation-based approach	Only mediapipe	mediapipe with YOLO	Fusion
Object standing in front of a shelf	○	○	○
Object standing in front of a shelf (Obscured object)	X	○	○
Object running in store	○	○	○
Object that steal things into its inner pocket	○	○	○
Object sitting on a chair	○	○	○
Object sitting on a floor	○	○	○
Object sitting on a floor (Obscured object)	X	○	○
Object picking up things (Obscured object)	X	○	○
Object putting hand in pocket of pants (Obscured object)	X	○	○
Hand object grabbing an something	○	○	○

## 5. 결론 및 향후 연구과제

본 연구에서는 수많은 LiDAR 센서, 가중치 센서 등의 사용으로 고 비용으로 어려움을 겪고 있는 무인 점포에 도움을 주고자 저비용으로 점포 내 사용자들의 행동 패턴을 분석하고 탐지하는 모션 탐지 융합 알고리즘을 제안했다. 다중 객체 탐지가 가능한 객체 인식과 영상 처리를 지원하고 KNN 알고리즘을 제공하는 영상 처리, 객체의 관절점을 제공해 행동 패턴을 탐지할 수 있는 모션 캡처 등을 융합한 알고리즘을 개발했다. 또한 실제 무인 점포와 수많은 사용자들이 존재하는 점포에서 행동 패턴에 대한 이미지 데이터를 수집하였고 학습 시킨 뒤 알고리즘을 실행하여 행동 패턴 별로 Recall, Precision, F1-Score를 구하였으며 다양한 시나리오를 설계해 실험을 통해 융합 알고리즘의 성능을 보였다. 이는 곧 무인 점포들의 센서 의존도를 낮추고 본 연구에서 제시한 융합 알고리즘이 이를 대체하게 하여 점포 운영에 있어 비용적인 부담을 줄이는데 도움을 줄 수 있을 것이다. 그리고 우리는 본 연구를 해당 융합 알고리즘을 통해 무인 점포뿐만 아

나라 무인 카페, 어린이 교통 사고 방지 CCTV, 군부대 내 거주자 탐지 CCTV 등의 시스템으로 확장해 나갈 예정이다.

## 참고문헌(Reference)

- [ 1 ] Kyungha Min, The number of unmanned convenience stores exceeds 3,300 Increased 6 times in 2 years,” etnews, 2023. <https://www.etnews.com/20230109000221>
- [ 2 ] Youtube, “Introducing Amazon Go and the world’s most advanced shopping technology,” 2016. <https://www.youtube.com/watch?v=NrmMk1Myrxc>
- [ 3 ] Just Walk Out, <https://justwalkout.com/>
- [ 4 ] Ryan Gross, “How the Amazon Go Store’s AI Works, Towards Data Science, 2019. <https://towardsdatascience.com/how-the-amazon-go-store-works-a-deep-dive-3fde9d9939e9>
- [ 5 ] Namkyeong Lee, Hyunjun Park, “Key Successful Factors for Unmanned Convenience Stores in the Fourth Industrial Revolution : Case of E-Mart24 Self-Store”, Vol. 26, No. 2, 2021. <http://dx.doi.org/10.9723/jksis.2021.26.2.073>
- [ 6 ] Jeongwon Yoo, “E-Mart 24 Smart Store... 3 things that the Gimpo DC branch does not have and the only the COEX branch has,” goodkyung, 2021. <http://www.goodkyung.com/news/articleView.html?idxno=157031>
- [ 7 ] Byungjoon Kim, Yongduck Seo, “Data Generation System for Flaw Detection of OLED Panel and Application of RCNN-based Defect Detection,” Vol. 20, No. 12, pp.57-63, 2022. <http://dx.doi.org/10.14801/jkiit.2022.20.12.57>
- [ 8 ] Gunwoo Do, Donghyeon Kim, Siwoong Jang, “Performance Comparison between Yolov5 and Yolov8 Models trained on Fire Image,” Vol. 27, No. 1, pp.586-588, 2023. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11498522>
- [ 9 ] Jihoon Lee, Minchan Shin, Junhee Park, Nammee Moon, “Deep Learning-Based Companion Animal Abnormal Behavior Detection Service Using Image and Sensor Data,” Journal of The Korea Society of Computer and Information, Vol. 27, No. 10, pp.1-9, 2022. <http://dx.doi.org/10.9708/jksci.2022.27.10.001>
- [10] Inhye Kim, Ilhong Jung, “A Study on Korea Sign Language Motion Recognition Using OpenPose Based on Deep Learning,” Journal of Digital Contents Society, Vol 22, No. 4, pp.681-687, 2021. <http://dx.doi.org/10.9728/dcs.2021.22.4.681>
- [11] Suyeon Han, Deawoo Park, “Cat Behavior Pattern Analysis and Disease Prediction System of Home CCTV Images using AI,” Journal of the Korea Institute of Information and Communication Engineering, Vol. 26, No. 9, pp.1266-1271, 2022. <http://doi.org/10.6109/jkiice.2022.26.9.1266>
- [12] Rakbin Song, Yuna Hong, Noyoon Kwak, “User Interface Using Hand Gesture Recognition Based on Mediapipe Hands Model,” Journal of Korea Multimedia Society, Vol. 26, No. 2, pp.103-115, 2023. <https://doi.org/10.9717/kmms.2023.26.2.103>
- [13] Bokdeuk Song, Seunghwan Lee, Hongkyw Choi, Sunghoon Kim, “Design and Implementation of a Stereoscopic Image Control System based on User Hand Gesture Recognition,” Journal of the Korea Institute of Information and Communication Engineering, Vol. 26, No. 3, pp.396-402, 2022. <http://doi.org/10.6109/jkiice.2022.26.3.396>
- [14] Mediapipe, <https://developers.google.com/mediapipe>
- [15] Giwook Cha, Wonhwa Hong, “Development of Optimal k-Nearest Neighbors (KNN) Model to Predict Demolition Waste Generation in Redevelopment Area,” J. Korean Soc. Living Environ. Sys., Vol. 30, No. 1, pp.20-29, 2023. <https://doi.org/10.21086/ksles.2023.2.30.1.20>
- [16] roboflow, <https://roboflow.com/>
- [17] AI-Hub, <https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=171>

## ◎ 저 자 소 개 ◎



### 최 영 준(Youngjune Choi)

2019년~현재 경기대학교 AI컴퓨터공학부 컴퓨터공학전공 학부생  
관심분야 : 딥러닝, 컴퓨터 비전, 백엔드  
E-mail : dudwnzero@kyonggi.ac.kr



### 나 지 영(Jiyoung Na)

2023년~현재 경기대학교 AI컴퓨터공학부 인공지능학과 학부생  
관심분야 : 빅데이터, 컴퓨터 비전, 딥러닝.  
E-mail : jiyong@kyonggi.ac.kr



### 안 준 호(Junho Ahn)

2009년~2013년 University of Colorado, Boulder, Computer science, Ph.D.  
2013년~2017년 ETRI 국가보안기술연구소  
2017년~2023년 한국교통대학교 컴퓨터학부 교수  
2023년~현재 경기대학교 AI컴퓨터공학부 인공지능전공 교수  
관심분야 : 인공지능, 딥러닝, 비전.  
E-mail : jha@kyonggi.ac.kr