

# 스마트 홈 사용자를 위한 라이다, 영상, 오디오 센서를 이용한 인공지능 이상징후 탐지 알고리즘<sup>☆</sup>

## Intelligent Abnormal Situation Event Detections for Smart Home Users Using Lidar, Vision, and Audio Sensors

김 다 현<sup>1</sup>                      안 준 호<sup>1\*</sup>  
Da-hyeon Kim                Jun-ho Ahn

### 요 약

최근 COVID-19가 확산하고 외출 자제 권고와 같은 방역지침에 따라 집에서 생활하는 시간이 늘고 있다. 이에 따라 집에서 생활하는 1인 가구가 증가하고 있지만 1인 가구는 다인 가구보다 집 안에서 위급한 상황이 발생할 때 외부에 알리기 어렵다. 본 연구는 집안에서 발생하는 다양한 상황을 라이다, 영상, 음성 센서로 수집하고 센서에 따른 데이터를 각각의 알고리즘을 통해 분석하였다. 이를 이용해 위급상황 등의 비정상 패턴을 분석하여 사람의 이상징후를 탐지하는 연구를 진행했다. 각 센서에 따른 사람의 이상징후를 탐지하는 인공지능 알고리즘을 연구하였으며 센서에 따른 이상징후 탐지 정확도를 측정했다. 또한, 본 연구는 다양한 상황에 대한 센서의 탐지 가능 여부를 실험하여 센서 간의 장단점을 보완한 융합 방식을 제안한다.

☞ 주제어 : 스마트 홈, 인공지능, 이상징후 탐지, 센서 융합, 알고리즘

### ABSTRACT

Recently, COVID-19 has spread and time to stay at home has been increasing in accordance with quarantine guidelines of the government such as recommendations to refrain from going out. As a result, the number of single-person households staying at home is also increasing. Single-person households are less likely to be notified to the outside world in times of emergency than multi-person households. This study collects various situations occurring in the home with lidar, image, and voice sensors and analyzes the data according to the sensors through their respective algorithms. Using this method, we analyzed abnormal patterns such as emergency situations and conducted research to detect abnormal signs in humans. Artificial intelligence algorithms that detect abnormalities in people by each sensor were studied and the accuracy of anomaly detection was measured according to the sensor. Furthermore, this work proposes a fusion method that complements the pros and cons between sensors by experimenting with the detectability of sensors for various situations.

☞ keyword : Smart home, Intelligent, Abnormal detections, Sensor fusing, Algorithm

## 1. 서 론

1인 가구는 집 안에서 발생하는 사고, 갑작스러운 질

병 등에 대한 타인의 도움을 구하기 위해 본인이 직접 외부에 이상징후를 알려야 한다. 1인 가구 수는 19년도 기준 전체 가구 10가구 중 3가구가 1인 가구 [1]일 정도로 꾸준히 증가하고 있다. 노령 인구의 고독사 역시 매년 증가 [2]하는 것으로 볼 때, 1인 가구는 다인 가구에 비해 집 안 내 사고, 질병에 취약하다. 전 세계에 대유행 중인 COVID-19로 인해 재택근무를 하는 직장인들이 많아지고 사회적 거리 두기와 외출 자제 권고 등 방역지침에 따라 집에서 체류하는 시간이 길어졌다. 배달 앱 이용 비율은 77% 증가하였으며 약 60% 정도가 평일과 주말에 집에서 체류하는 시간이 늘었다고 조사 [3]되었다. 집안에서 체류하는 시간이 길어질수록 집안에서 발생할 수 있는 위급사항에도 빈번하게 노출된다. 이때, 인공지능 스피커

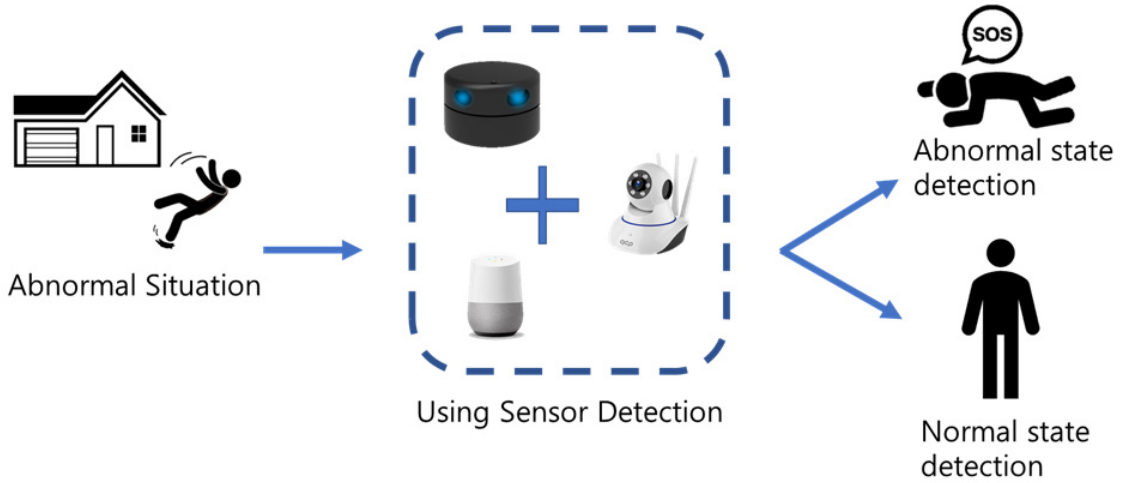
<sup>1</sup> Dept. of Software, Korea National University of Transportation, Chungju-si, 27469, Korea.

\* Corresponding author (jhahn@ut.ac.kr)

[Received 23 February 2021, Reviewed 12 April 2021, Accepted 12 May 2021]

☆ This research was supported by Korea National University of Transportation in 2021. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2020R111A3068274).

☆ 본 논문은 2020년도 한국인터넷정보학회 추계학술발표대회 우수 논문 추천에 따라 확장 및 수정된 논문임.



(그림 1) 이상징후 탐지 알고리즘 개념도

(Figure 1) Abnormal Situation Event Detection Algorithm Conceptual diagram

등 집 안 IoT 사물을 이용하여 위급상황과 이상징후를 판단하고 도움을 요청할 수 있다. 작년에 인공지능 스피커를 통한 단순한 도움 요청으로 구급대원을 출동시켜 독거 노인을 구했다 [4].

본 연구에서는 집 안에 있는 로봇 청소기, 인공지능 스피커, 홈 CCTV 등 다양한 IoT 사물의 센서를 활용하여 1인 가구의 패턴을 파악하고 수집된 패턴과 데이터에 따른 각각의 알고리즘과 융합 알고리즘을 통해 사용자의 비정상적인 패턴인 이상징후를 판단한다. 이때 이상징후는 갑작스러운 변화나 장시간의 무 변화 등 자연스럽게 않은 비정상적 패턴을 뜻한다. 음성, 영상, 라이더를 이용하여 패턴 분석하고 이상징후 판단하는 알고리즘을 제안한다. 하지만 단일 센서만으로 이상징후를 판단하는 경우 제한상황이 발생할 수 있다. 예를 들어 음성 센서의 경우 주변 TV, 라디오 등의 소리를 사용자가 내는 소리로 착각하여 사용자의 이상징후를 판단하기 힘들어진다. 영상 센서는 제한된 구역만 감지할 수 있으며 라이더의 경우 로봇 청소기 등에 부착된 저렴한 2D 라이더를 이용하기 때문에 정확한 상황판단이 힘들다. 따라서 본 연구에서는 다양한 상황에서 연구하여 상황에 따른 센서들의 사용 가능 여부를 조사하고 센서를 융합한 알고리즘을 제안한다. 연구에 대한 개념도는 그림 1을 보면 알 수 있다. 이는 이상징후라 판단되는 상황이 발생했을 때 센서에 따른 패턴을 분석하여 위급상황 등의 여부를 판단한다. 본 논문은 2장에서 관련 연구를 소개하고, 3장에는 센서에

따른 알고리즘 센서 별로 분류하여 제안하고 설명한 뒤 4장과 5장에서 실험 및 결과에 대한 평가를 진행한다.

## 2. 관련 연구

센서에 따른 각각에 알고리즘을 이용하여 이상징후를 탐지한다. 센서에 따른 사람, 사물 등을 탐지하고 이상징후를 탐지하는 것과 관련된 연구는 다음과 같다.

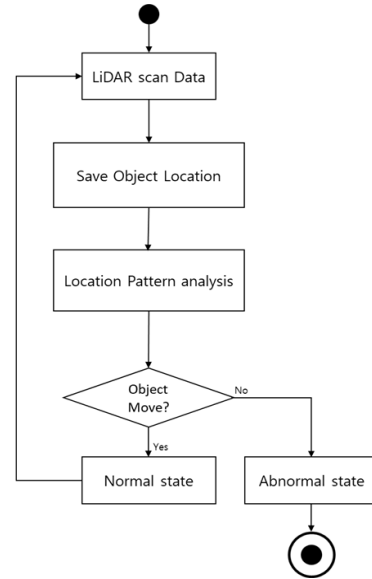
라이더를 이용하여 사람, 사물을 탐지하거나 움직임 추적하는 것과 관련된 연구가 진행되고 있다. 3D 라이더를 이용해 사람의 궤적을 분석하는 연구 [5]가 있다. 보행자의 궤적을 분석하여 도로에서 보행자의 행동을 예측하는 연구 [6]도 있다. 이때, 정상적이지 않은 궤적으로 움직임은 이상징후를 나타낸다. 3D 라이더와 같이 값비싼 라이더가 아닌 저렴하고 저해상도의 2D 라이더에 대해 인공지능 알고리즘을 이용한 연구가 많이 진행되고 있다. 2D 라이더의 스캔 결과 패턴이 국소적인 상태를 벽과 구분되는 물체라 판단하여 장애물을 탐지하는 연구 [7]가 있다. Recurrent Neural Network(RNN) 모델을 이용하여 라이더 센서의 인식된 사람의 패턴을 학습하여 사람을 탐지하고 움직임을 추적 [8]한다. 2D 라이더만 이용하여 사람을 탐지하기에는 정확도가 낮아 사람을 탐지하는 모델을 만들 때 카메라와 융합한 연구 [9, 10]도 있다. 이때 카메라와 라이더가 같은 방향을 향하게 고정된 뒤 카

메라 이미지에서 사람이 탐지되면 그때 라이다를 통해 얻은 스캔의 패턴을 사람이라 분류한다. 본 연구에서는 2D 라이다를 이용하여 고정된 객체와 움직이는 객체를 구분하고 움직이는 객체에 대해 패턴을 분석하여 비정상 패턴을 탐지할 수 있다.

영상을 통해 객체를 탐지하는 다양한 연구들이 진행되고 있다. 이때 딥러닝 모델을 설계하여 영상의 프레임에서 객체를 탐지한다. R-CNN [11]은 이미지에서 일정한 규칙으로 이미지를 여러 장 쪼개야 CNN 모델을 통과시킨다. 한 장의 이미지이지만 수천 번의 객체탐지를 수행하게 된다. 이러한 방법은 많은 학습시간을 요구하고 객체탐지 시 오랜 시간이 걸린다. 따라서 R-CNN 알고리즘에 Region of Interest(ROI) pooling layer를 이용하여 보완한 Fast R-CNN [12]이 있다. 다른 객체 탐지 알고리즘인 SPPnet [13]은 R-CNN의 Region Proposal이나 Feature map을 생성하여 속도에서 보완을 이루었다. 실시간으로 빠르게 처리가 가능한 YOLO [14]의 경우 한 장의 이미지에서 셀마다 객체를 탐지하고 탐지된 여러 개의 경계상자를 합쳐 객체를 탐지한다. 영상을 통해 사람의 패턴을 분석하여 제적을 분류하고 비정상 행동을 판단하는 연구 [15]도 있다. 본 논문은 객체를 탐지한 뒤 탐지된 객체를 분류하여 사람인 객체에 대해 패턴을 분석하여 이상징후를 탐지한다.

음성 분석을 통해 음성과 노이즈 신호를 분류하고 신호의 패턴을 분석하는 연구가 진행되고 있다. MFCC를 이용하여 GMM 기반의 음성 신호를 분류하는 연구 [16]가 있다. Dynamic Time Warping (DTW)를 기반으로 유아의 울음소리의 패턴을 분석하여 서로 다른 상황으로 분류한 연구 [17]도 있다. CNN을 이용하여 오디오 장면을 분류하여 주변 환경을 인식하는 연구 [18]도 있다. 음성 분석을 통해 비정상 상황을 판단하는 연구 [19, 20]도 있다. 사람이 있다고 판단되지만 장시간 어떤 소리도 들리지 않거나 갑작스러운 큰 소리가 들리는 등의 상황을 음성 분석을 통해 탐지한다.

본 연구에서는 2D 라이다와 관련된 알고리즘 [7]을 이용하여 사람과 사물을 구별하고 사람에 대한 행동 패턴을 분석하여 이상징후를 판단했다. 객체탐지 알고리즘 YOLO [14]과 OpenCV를 이용한 컴퓨터 비전 알고리즘을 이용하여 홈 CCTV를 활용한 이상징후를 판단을 연구했다. 기존 연구에서 개발한 음성 분석 알고리즘 [19]을 이용하여 음성 패턴을 분석하고 이상징후 판단을 연구했다. 또한, 각각의 알고리즘의 한계점을 보완하기 위해 융합 알고리즘을 제안하고 다양한 상황에 대해 센서의 이상징후 탐지 가능 여부를 실험하여 융합 방식을 제안한다.



(그림 2) 2D 라이다 알고리즘 구조  
(Figure 2) Flow chart of the 2D LiDAR algorithm

### 3. 알고리즘

#### 3.1 라이다 알고리즘

본 연구에서는 집 안 로봇 청소기에 부착된 2D 라이다를 활용하여 이상징후를 판단하기 위해 2D 라이다를 이용한 연구를 진행하였다. LiDAR는 Light Detection and Ranging의 줄임말로 펄스 레이저 형태의 빛을 사용하여 가변 거리를 측정하는 센서이다. 라이다 기술은 빛 펄스가 발사되어 물체의 표면에서 반사되면서 감지되는 시간을 측정하여 물체까지의 거리를 계산한다. 이때 2D LiDAR는 2차원 평면에서만 물체를 감지할 수 있다. 2차원의 전체 평면을 측정하기 위해 라이다는 360도 회전하며 평면을 스캔한다. 3D LiDAR를 이용하면 물체의 특징이 되는 지점을 3차원으로 얻을 수 있지만 수백만 원에서 수억 원까지의 고가로 상용화하기 힘들다. 하지만 2D 라이다는 약 십만 원 정도로 저렴하며 쉽게 활용할 수 있다. 라이다는 로봇의 자율주행에서 필수적인 SLAM 기술을 활용하기 위해 많이 사용된다. 라이다 센서마다 측정할 수 있는 범위가 다른데 연구에 사용된 YDLiDAR 사의 X4 LiDAR는 360도 회전하며 스캔하고 12cm ~ 10M까지의 범위를 측정할 수 있다. 따라서 한 번에 넓은 범위를 스캔할 수 있다는 장점이 있다. 본 연구에서는 그림 2과

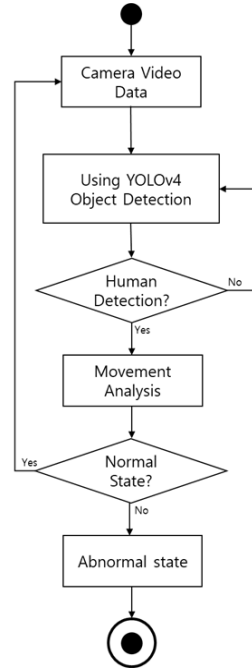
같은 2D 라이더를 사용하여 이상징후 탐지 알고리즘을 제안한다. 라이더 알고리즘은 물체의 위치를 파악하는 알고리즘 [7]을 이용하여 라이더의 스캔 데이터를 벽과 사람으로 구분하였다. 스캔 된 point cloud의 x, y 값이 일정하게 변화하고 큰 변화가 없는 경우, 해당 집합을 건물의 벽으로 판단하고 국소적인 point cloud 집합은 사람이라고 판단한다. 그 후 본 연구에서는 사람이라고 판단된 집합의 중심을 추정하여 물체의 위치로 x, y 값을 저장하게 된다. 저장된 사람의 위치 패턴을 분석하여 사람의 움직임 변화가 하루 이상에 장시간 없는 경우 이상징후가 있다고 판단한다.

### 3.2 영상 알고리즘

영상 알고리즘은 홈 CCTV를 활용하여 이상징후를 판단하였다. 홈 CCTV는 일반적인 단안 카메라로 집안의 상황을 실시간으로 촬영한다. 카메라는 고정된 위치에서 정해진 장소만 촬영한다는 단점이 있다. 하지만 정해진 범위 내에서 대해서는 정확한 이상징후 판단이 가능하다. 이때의 이상징후는 카메라 영상에서 탐지된 사람에게서 장시간 움직임이 포착되지 않고, 사람이 홈 CCTV 영상에 탐지되지 않으며 집 안 고정된 사물에도 장시간 변화가 없을 때 이상징후가 있다고 판단한다. 영상 알고리즘을 활용한 Flow chart는 그림 3과 같다. 홈 CCTV를 통해 녹화된 영상에서 실시간 탐지를 위해 5초에 한 번 객체를 탐지한다. 탐지 시 YOLO를 이용하여 학습시킨 모델을 사용하는데 다양한 YOLO 버전 중 YOLOv4를 사용하였다. YOLOv4는 이전에 개발된 알고리즘들에 비해 속도 및 정확도가 월등히 높다. 사용하는 모델은 COCO 데이터를 이용하여 학습된 모델로 객체를 탐지하여 객체가 어떤 것을 나타내는지 80개의 라벨로 분류하게 된다. 본 연구에서는 탐지된 객체를 바탕으로 객체가 가지는 경계상자의 값과 라벨의 이름을 이용하여 사람인지 사물인지 구분하고 이때, 사람으로 판단된 객체의 움직임을 분석한다. OpenCV를 이용하여 객체가 비정상적 패턴을 탐지한다.

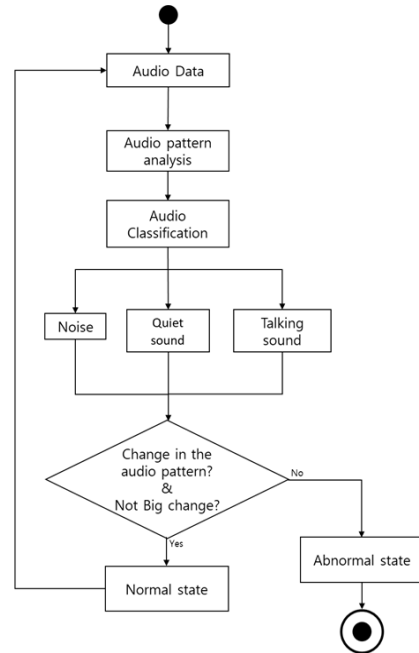
### 3.3 음성 알고리즘

음성 알고리즘은 기존 연구에서 개발한 알고리즘 [19]을 이용한다. 알고리즘은 음성 패턴을 분석한 뒤 분류한다. 이때, 음성 패턴은 노이즈 낀 소리, 조용한 소리, 말하는 소리로 분류된다. 음성은 라이더나 영상보다 훨씬 더 넓은 범위의 탐지가 가능하며 벽이나 장애물에 큰 영향을 받지 않는다. 하지만 집 안의 사물에서 발생하는 대화



(그림 3) 영상 알고리즘 구조

(Figure 3) Flow chart of the vision algorithm



(그림 4) 음성 알고리즘 구조

(Figure 4) Flow chart of the audio algorithm

소리나 생활하며 발생하는 노이즈로 인해 잘못된 결과를 낼 수도 있다. 따라서 본 연구에서는 이상징후 판단 시 장 시간에 같은 패턴이 지속하는 경우 이상징후가 있다고 판단한다. 실제 사람이 말을 하는 상황에서는 분류에 변화가 생기지만 변화 없이 같은 패턴으로 계속 분류가 된다는 것은 이상징후가 생겼다고 볼 수 있다. 예를 들어 사람의 실제 대화 소리는 대화가 끝나면 조용한 소리나 노이즈 낀 소리로 바뀌지만 계속되는 대화 소리는 티브이, 라디오 등의 소리로 판단할 수 있다. 티브이, 라디오 등이 켜질 동안 재생되는 것을 보고 사용자에게 이상징후가 생겼다고 판단하게 된다. 또한, 패턴의 변화가 자연스럽게 않고 갑작스럽게 변하는 경우 이상징후가 있다고 판단했다. 이는 대화 소리와 같은 큰 소리가 짧게 탐지되고 조용한 소리가 계속되는 경우 등이 있다. 그림 4는 음성 패턴 알고리즘의 Flow chart이다.

### 3.4 융합 알고리즘

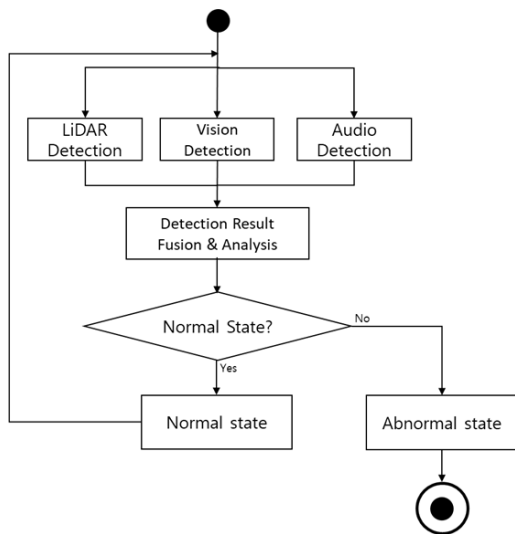
본 연구에서는 각각의 알고리즘의 단점을 보완하기 위해 융합 알고리즘을 제안한다. 2D 라이다 알고리즘의 경우 탐지에 정확도가 떨어지며 라이다의 스캔 범위를 벗어나서 물체가 움직이는 경우 알고리즘을 통해 탐지할 수 없다. 큰 물체나 기둥과 같은 장애물이 있는 곳은 통과하지 못해 장애물에 많은 영향을 받는다. 또한, 저해상도의 라이다를 사용하기 때문에 탐지가 되지 않는 경우가

발생할 수 있다. 영상 알고리즘은 홈 CCTV가 설치된 고정된 장소에서 탐지하기 때문에 탐지범위를 벗어나는 경우 탐지할 수가 없다. 음성 알고리즘의 경우 생활 속 다양한 노이즈로 인해 탐지의 정확도가 낮아지며 오판의 가능성이 크다. 본 연구에서는 라이다와 영상, 음성 알고리즘을 동시에 사용함으로써 넓은 범위를 탐지하며 정확도를 높였다. 알고리즘 Flow chart는 그림 5와 같으며 다양한 상황에서 각 센서가 탐지한 결과를 분석하여 이상징후를 판단한다.

## 4. 실험 및 결과

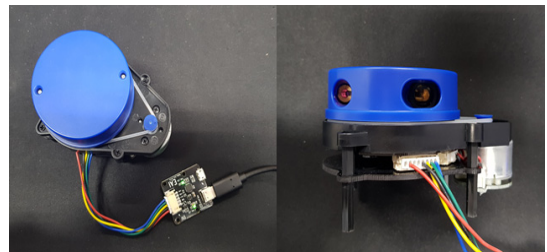
본 연구에서는 각 센서의 알고리즘을 통하여 이상징후를 탐지하는 실험을 진행하고 각 알고리즘에 따른 결과를 구하였다.

라이다를 통해 이상징후를 판단하는 실험과 실험 결과이다. 그림 6은 실험에 이용한 YDLiDAR 사의 X4 2D 라이다이다. 라이다를 실행시키면 파란색 부분이 360도 회전하며 탐지하게 된다. 이때, 라이다는 ROS를 이용하거나 YDLiDAR 사에서 제공하는 응용 프로그램을 이용하여 라이다를 실행시키면 스캔 된 point cloud 값을 얻을 수 있다. 라이다는 360도 전 방향 스캔하며 6~12Hz의 스캔 주파수를 갖는다. 본 연구에서는 Ubuntu 18.04 OS 환경에 Ros Melodic 버전을 이용하여 라이다를 사용하였다.



(그림 5) 융합 방식 구조

(Figure 5) Flow chart of the fusion method



(그림 6) X4 2D 라이다

(Figure 6) X4 2D LiDAR sensor

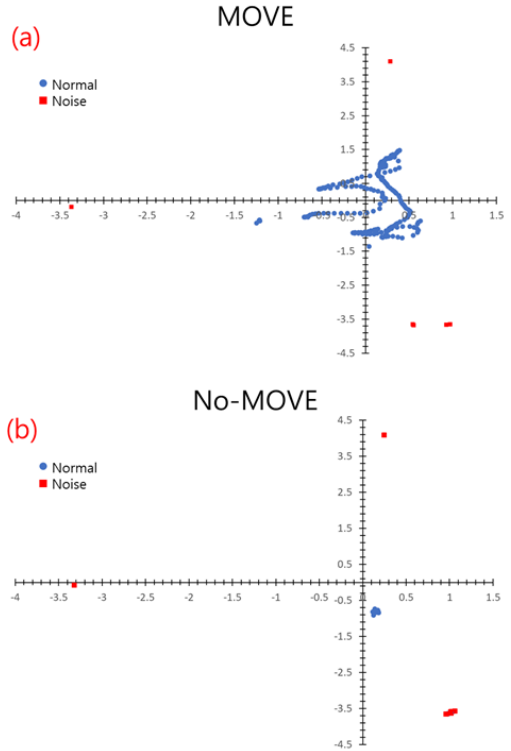
본 연구의 실험에서 라이다를 이용하여 탐지한 결과는 그림 7과 같고 이때 벽과 물체를 구분하여 물체가 무엇인지 추정한다. 이때 물체로 추정된 point cloud의 집합은 초록과 빨강으로 이루어진 원으로 나타나며 벽으로 추정되는 부분은 파란색 선으로 나타나고 있다. 그리고 라이다가 스캔하여 얻어내는 point cloud는 빨간색 점으로 나타나고 동시에 라이다만으로 Mapping이 가능한 Hector

Mapping을 이용하여 Mapping 하였다. Mapping 결과는 회색으로 채워져 있으며 채워진 부분은 사람이 움직일 수 있는 범위이다. 본 연구에서는 라이다에 스캔이 가능한 범위 내에서 움직이며 라이다가 물체를 인식하는지 확인했다. 또한, 코드를 통해 얻어내는 물체의 point cloud 집합의 중심을 위젯값으로 기록하였다. 그림 7의 가장 위쪽 그림은 움직이는 객체를 정확하게 탐지한 것이다. 이때 Map과 같이 표현된 우측 그림을 통해 객체가 벽 쪽에서 움직이고 있다는 것을 알 수 있다. 2번째 줄에 있는 그림에서 Map이 같이 나타난 우측 그림을 통해 벽 부분이 객체로 탐지된 것을 볼 수 있다. 마지막 줄의 그림은 라이다가 스캔한 point cloud의 빨간색 점을 제외하고 물체 추정과 벽 추정을 나타낸 그림이다.

본 연구에서는 탐지된 객체를 나타내는 원의 중심 좌표를 구하여 이를 표시한 그래프는 그림 7과 같다. 그림 8에 잘못 추정된 위치는 빨간색 점으로 표시하고 정상적인 객체의 위치는 파란색 점으로 표시하였다. 그림 8에 (a)는 객체가 움직이고 있을 때의 위치 변화를 나타내고 (b)는 객체가 움직이고 있지 않을 때의 위치 변화를 나타낸 것이다. 이때 본 연구에서는 움직임이 적거나 객체가 탐지되지 않는 경우 이상징후가 발생했다고 판단했다.



(그림 7) 라이다를 이용하여 물체 추정 결과  
(Figure 7) The Result of the object detections using the LiDAR sensor



(그림 8) 라이다를 통해 탐지된 객체의 움직임, a)는 객체의 움직임 패턴 결과, b)는 객체가 움직이지 않을 패턴 결과

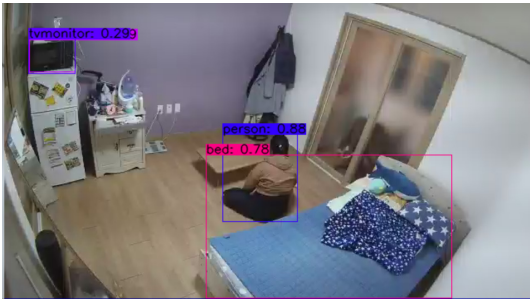
(Figure 8) Movement of object detected through LiDAR, a) is movement of object patten result, b) is not movement of object patten result

(표 1) 라이다 알고리즘의 탐지 성능

(Table 1) The detection performance result of the LiDAR algorithm

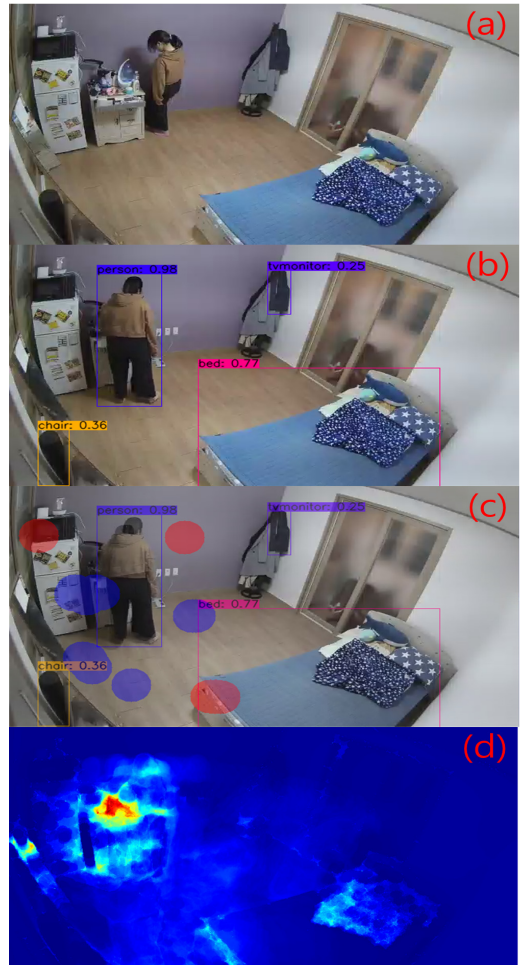
Algorithms	Move	Accuracy
2D LiDAR	○	0.954
Algorithms	×	0.866

본 연구에서는 총 216개의 위치 정보를 얻었으며 이에 따라 움직일 때의 정확도와 움직이지 않을 때의 정확도는 표 1과 같다. 본 연구에서는 사람으로 분류된 객체에 움직임이 하루 이상 No-MOVE와 같이 나타나는 경우 이상징후가 발생했다고 판단했다.



(그림 9) YOLOv4를 이용한 객체탐지 결과  
(Figure 9) The Result of the vision object detections using YOLOv4

영상을 통해 이상징후를 판단하는 실험은 다음과 같다. 그림 9는 YOLOv4 모델을 사용하여 실험 영상에서 객체를 탐지하는 그림이다. 그림9에서는 탐지한 객체가 어떤 객체인지 분류하여 라벨을 붙이고 어느 정도의 비율로 해당 라벨로 분류되었는지 같이 표시되고 있다. 영상 데이터는 홈 CCTV를 이용하여 실제 1인 가구와 같은 환경에서 수집했다. 홈 CCTV는 1초에 30프레임을 얻을 수 있으며 눕거나, 움직이는 등 다양한 상황을 연출하여 영상을 수집하였다. 본 연구의 알고리즘은 YOLOv4 모델을 통해 객체를 탐지한 뒤 기존에 연구된 OpenCV를 사용한 컴퓨터 비전 알고리즘으로 움직임과 패턴을 분석하였다. 연산의 과정을 줄이기 위해 모든 프레임에서 객체를 탐지하는 것이 아닌 5초에 1번씩 객체를 탐지하였다. 탐지된 객체가 0.5 미만의 탐지 정확도를 가지면 해당 객체를 잘못 탐지된 것으로 분류했다. 이때 사람으로 탐지된 객체에 비정상 패턴이 탐지되는 경우 이상징후가 발생했다고 판단하였다. 이는 사람의 궤도가 갑작스럽게 변하는 경우이다. 본 연구의 실험에서 수집한 영상 데이터에 영상 알고리즘에 적용한 결과는 그림 10와 같다. 그림 10의 (a)는 원본 영상이고 (b)는 5초에 한 번 YOLO를 통해 객체가 탐지되고 있는 모습이다. (a)와 (b) 그림은 5초간의 차이로 인해 다르게 나타나고 있다. 그림 10의 (c)와 (d)는 OpenCV를 활용하여 객체의 패턴을 분석하고 객체의 이동 경로를 추적하는 것을 보여준다. (c)에 파란 동그라미는 사람으로 탐지된 객체의 이동 경로를 보여주며 빨간 동그라미는 사물이 탐지될 때 사람과 구분하기 위해 나타나고 있다. (d)는 프레임 별 비교를 통해 움직임이 있는 객체를 나타내는 것으로 이때 움직임이 있는 사람 객체가 있는 곳에 특이점이 나타난 것을 볼 수 있다. 이러한 움직임이 비정상적일 경우 이상징후라 탐지했다. 본 연구



(그림 10) 영상 알고리즘을 이용하여 객체탐지 결과  
(Figure 10) Examples of the object detection result using the vision algorithm

에서는 수집한 영상 중 객체가 탐지된 120개의 데이터에서 알고리즘의 정확도를 Recall, Precision, Accuracy를 구하고 이를 표 2에 나타냈다.

(표 2) 영상 알고리즘의 이상징후 탐지 성능  
(Table 2) The result of user event detections using the vision algorithm

Algorithms	Recall	Precision	Accuracy
Vision Algorithms	0.92	0.95	0.89

(표 3) 음성 알고리즘의 탐지 성능

(Table 3) The performance result of the audio algorithm

Algorithms	Type	Accuracy
Audio Pattern Algorithms	Talk	0.71
	Noise	0.56
	Quit	0.85

음성 센서를 이용하여 넓은 범위에서 이상징후를 판단할 수 있다. 특정 패턴의 소리가 하루 이상 지속 되는 경우 집 안에서 이상징후가 발생했다고 판단했다. 본 연구의 알고리즘은 음성 데이터의 패턴을 분석하여 데이터가 어떠한 소리 유형에 가까운지 분류했다. 분류 가능한 소리 유형에는 대화 소리, 노이즈 낀 소리, 조용한 소리가 있다.

분류 가능한 3가지 소리에 따라 유튜브 동영상과 직접 녹음한 음성파일을 이용하여 실험하였다. 각각의 소리마다 5개의 씹의 음성 파일을 수집하여 나타낸 결과는 그림 11과 같다. 이때 대화 소리의 경우 음성 그래프가 가장 빈번하게 오가는 것을 볼 수 있다. 조용한 소리의 경우 대화 소리와 노이즈 낀 소리가 거의 나타나고 있지 않다. 각 유형에 따른 알고리즘 정확도를 계산하여 표 3으로 나타내었다. 정확도는 해당 유형에서 나올 수 없는 패턴을 통해 계산하였다. 본 연구에서 자연스럽게 패턴이 변화하는 것이 아닌 대화 소리와 같이 큰 소리가 들린 후 갑자기 조용한 소리가 들리거나 하루 이상 같은 패턴의 소리가 반복되는 경우 이상징후가 발생했다고 판단했다.

알고리즘을 융합하여 이상징후 탐지 정확도를 높이기 위해 각 상황에 따른 알고리즘 및 센서 사용 가능 여부를 실험하였다. 표 4는 집 안에서 발생할 수 있는 9개의 정상적인 상황과 3개의 비정상적인 상황에 대해서 알고리즘의 이상징후 탐지 가능 여부를 나타내고 있다. 본 연구에서 제안하는 융합 알고리즘은 각 센서 알고리즘이 상호 보완적으로 사용된다. 만약 청소기를 사용하는 상황이라면 음성 센서는 이상징후를 탐지할 수 없지만, 영상 센서를 이용해 사람의 행동 패턴을 분석하여 이상징후를 탐지할 수 있다. 이와 마찬가지로 집 안이 어두운 경우 영상 알고리즘을 사용할 수 없지만, 라이다와 음성 알고리즘을 통해 이상징후를 탐지할 수 있으며 집 주변에 도로 등이 존재해 음성 알고리즘의 잘못된 판단이 많은 상황이라면 다른 센서를 통해 이상징후 여부를 더 정확히 탐지할 수 있다.

(표 4) 융합 방식을 통한 이상징후 탐지 가능 여부

(Table 4) Unusual user event detections using the fusion method

Situation	LiDAR	Vision	Audio
집 안에서 운동하는 상황	×	○	○
TV를 시청하는 상황	○	○	×
식사를 하고 있는 상황	○	○	○
청소기를 사용하고 있는 상황	×	○	×
밤에 잠을 자는 상황	○	×	○
낮에 잠을 자는 상황	○	○	○
전화하고 있는 상황	○	○	○
TV를 하루 이상 켜 두는 상황	○	○	×
책을 읽는 상황	○	○	×
사람이 쓰러져 기절 상황	○	○	○
집 안에서 넘어진 상황	×	○	○
집 안에서 연기가 발생한 경우	○	×	○

## 5. 결 론

본 연구를 통해서 라이다, 영상, 음성 알고리즘을 이용하여 1인 가구에서 발생하는 사용자의 비정상적인 패턴인 이상징후를 탐지했다. 라이다의 경우 사람의 움직임 패턴을 분석하여 움직이는 경우 대략 0.95의 탐지 정확도를 나타내고 움직이지 않는 경우 대략 0.87의 탐지 정확도를 나타내고 있다. 카메라를 이용한 경우 대략 0.89의 정확도를 나타내며 음성 센서를 이용한 경우 각 상황에 따라 말하는 경우 약 0.71의 정확도를 나타내며 노이즈가 낀 경우 0.56의 정확도를 나타내고 조용한 경우에는 0.85의 정확도를 나타내고 있다. 하지만 실제 상황에서 사각지대나 제한상황이 발생할 경우를 단일 센서의 탐지 정확도는 더 낮아질 수 있다. 단일 알고리즘뿐만 아니라 융합 알고리즘을 제안하여 이상징후를 잘못 판단할 경우 각 센서가 서로 보완할 수 있게 하였다. 차후 연구를 통해 각 센서를 통한 탐지의 정확도 및 융합 알고리즘의 정확도를 높여서 1인 가구에서 발생할 수 있는 위험을 예방하는데 도움을 줄 수 있을 것이다.

## 참고문헌(Reference)

[1] The Statistics Korea, "2020 Statistics show single-person households", 2020.  
[http://kostat.go.kr/portal/korea/kor\\_nw/1/1/index.board?bmode=read&aSeq=386517](http://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=386517)



- [ 2 ] Korea Mutual-aid News, “The soaring rise of elderly lonely deaths, Urgent need for a social safety net”, 2020.  
<https://www.kongje.or.kr/news/articleView.html?idxno=880>
- [ 3 ] The Science Times, “How has your lifestyle changed since COVID-19?”, 2020.  
<https://www.sciencetimes.co.kr/news/%EC%BD%94%EB%A1%9C%EB%82%98-%EC%9D%B4%ED%9B%84-%EB%9D%BC%EC%9D%B4%ED%94%84-%EC%8A%A4%ED%83%80%EC%9D%BC%EC%9D%80-%EC%96%B4%EB%96%BB%EA%B2%8C-%EB%B3%80%ED%96%88%EB%82%98#.YcVUs4HVA8Y.link>
- [ 4 ] Yonhapnews, “AI speaker saved for the elderly living alone”, 2020.  
<https://www.yna.co.kr/view/AKR20200803102200052>
- [ 5 ] Csaba Benedek, “3D people surveillance on range data sequences of a rotating Lidar”, Pattern Recognition Letters, Vol. 50, pp. 149-158, 2014.  
<https://doi.org/10.1016/j.patrec.2014.04.010>
- [ 6 ] Z. Ma, J. Bigham, S. Poslad, B. Wu, X. Zhang and E. Bodanese, “Device-Free, Activity During Daily Life, Recognition Using a Low-Cost Lidar”, IEEE Global Communications Conference (GLOBECOM) ,pp. 1-6, 2018. <https://doi.org/10.1109/GLOCOM.2018.8647251>
- [ 7 ] M. Przybyła, “Detection and tracking of 2D geometric obstacles from LRF data,” 2017 11th International Workshop on Robot Motion and Control (RoMoCo), pp. 135-141, 2017.  
<https://doi.org/10.1109/RoMoCo.2017.8003904>
- [ 8 ] C. Premebida, O. Ludwig and U. Nunes, “Exploiting LIDAR-based features on pedestrian detection in urban scenarios”, 12th International IEEE Conference on Intelligent Transportation Systems, pp. 1-6, 2009.  
<https://doi.org/10.1109/ITSC.2009.5309697>
- [ 9 ] Dan Jia, Mats Steinweg, Alexander Hermans, Bastian Leibe, “Self-Supervised Person Detection in 2D Range Data using a Calibrated Camera”, arXiv, 2020.  
<https://arxiv.org/abs/2012.08890>
- [10] Tslil, Or and Elbaz, Amit and Feiner, Tal and Carmi, Avishy, “Representing and updating objects’ identities in semantic SLAM”, IEEE 23rd International Conference on Information Fusion (FUSION), Rustenburg, South Africa, pp. 1-7, 2020.  
<https://doi.org/10.23919/FUSION45008.2020.9190524>
- [11] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, arXiv, 2014.  
<https://arxiv.org/abs/1311.2524>
- [12] Ross Girshick, “Fast R-CNN”, International Conference on Computer Vision (ICCV), 2015.  
<https://arxiv.org/abs/1504.08083>
- [13] He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition”, Lecture Notes in Computer Science, pp. 346-361, 2014.  
[http://dx.doi.org/10.1007/978-3-319-10578-9\\_23](http://dx.doi.org/10.1007/978-3-319-10578-9_23)
- [14] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection”, arXiv, 2016.  
<https://arxiv.org/abs/1506.02640>
- [15] S. Coşar, G. Donatiello, V. Bogomy, C. Garate, L. O. Alvares and F. Brémont, “Toward Abnormal Trajectory and Event Detection in Video Surveillance”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 27, No. 3, pp. 683-695, March 2017.  
<https://doi.org/10.1109/TCSVT.2016.2589859>
- [16] L. Vuegen, B. V. Broeck, P. Karsmakers, J. Gemmeke, B. Vanrumste, H. V. hamme, “An MFCC GMM approach for event detection and classification”, IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events, 2013.  
[https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS2349843&context=L&vid=Lirias&search\\_scope=Lirias&tab=default\\_tab&lang=en\\_US](https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS2349843&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US)
- [17] Xilin Yu, Laishuan Wang, Xian Zhao, Chunmei Lu, Xi Long, Wei Chen, “An Investigation into Audio Features and DTW Algorithms for Infant Cry Classification”, Association for Computing Machinery, No. 06, pp. 54-59, 2019. <https://doi.org/10.1145/3375923.3375929>
- [18] Huy Phan, Lars Hertel, Marco Maass, Philipp Koch, Radoslaw Mazur, Alfred Mertins, “Improved Audio Scene Classification Based on Label-Tree Embeddings and Convolutional Neural Networks”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 25, 2017.

<https://doi.org/10.1109/TASLP.2017.2690564>

- [19] Ju-Ho Jung, Jun-Ho Ahn, "Intelligent User Pattern Recognition based on Vision, Audio and Activity for Abnormal Event Detections of Single Households", Korean Society of Computer Information, Vol. 24, No. 5, pp. 59-66, 2019.

<https://doi.org/10.9708/JKSCl.2019.24.05.059>

- [20] R. Bardeli and D. Stein, "Uninformed Abnormal Event Detection on Audio," Speech Communication, 10. ITG Symposium, Braunschweig, Germany, pp. 1-4, 2012.  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=6309608>

## ● 저 자 소 개 ●



### 김 다 현(Da-hyeon Kim)

2018년 3월~현재 한국교통대학교 컴퓨터정보기술공학부 소프트웨어전공  
관심분야 : 딥러닝, 패턴인식  
E-mail : 1826059@ut.ac.kr



### 안 준 호(Sun-shin Lee)

2009년~2013년 University of Colorado, Boulder, Computer science, Ph.D.  
2013년~2017년 ETRI 국가보안기술연구소  
2017년~현재 한국교통대학교 컴퓨터정보기술공학부 소프트웨어전공 교수  
관심분야 : 인공지능, 사물인터넷  
E-mail : jhahn@ut.ac.kr