

# 상황인식 기반 위험인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템

이동섭\*<sup>1</sup>, 전채욱\*<sup>2</sup>, 박주환\*\*<sup>3</sup>, 옥승호\*\*\*<sup>4</sup>

## Context-Aware Security Robot System for Dangerous Person Discrimination with Digital Twin Integration

Dong-Sup Lee\*<sup>1</sup>, Chae-Uk Jeon\*<sup>2</sup>, Joo-Hwan Park\*\*<sup>3</sup>, and Seung-Ho Ok\*\*\*<sup>4</sup>

“이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임” (IITP-2026-RS-2020-II201791, 80%) “본 논문은 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임” (20%)

### 요약

최근 인공지능(AI)과 자율주행 기술의 발전으로 이동형 방법 로봇이 자율 순찰과 실시간 위험 탐지를 수행하며 기존 감시 체계의 한계를 보완하고 있다. 그러나 기존 시스템은 영상 전달과 단순 객체 인식에 의존하여 위험 상황을 정밀하게 판단하는 데 한계가 있다. 이러한 문제를 보완하기 위해 본 논문에서는 자세 추정과 객체-얼굴 인식 정보를 융합하여 무기 소지 여부를 판별하고 결과를 3차원 디지털 트윈 환경에서 시각화할 수 있는 상황인식 기반 위험인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템을 제안한다. 실험 결과, 무기 탐지 정확도는 mAP50 기준 95.55%, 위험인물 판별 정확도는 1~2m 거리 구간에서 82.5%를 기록하였으며, Jetson Orin NX 기반 on-device AI 환경에서 준실시간 위험 인식이 가능함을 확인하였다.

### Abstract

Recent advances in Artificial Intelligence (AI) and autonomous driving technologies have enabled mobile security robots to perform autonomous patrols and real-time threat detection, complementing the limitations of existing surveillance systems. However, existing systems rely on video transmission and simple object recognition, limiting their ability to accurately assess threat situations. To address these issues, this paper proposes a context-aware dangerous-person discrimination and digital twin security robot system that integrates pose estimation with object and facial recognition information to discriminate weapon possession and visualize the results within a 3D digital twin environment. Experimental results showed a weapon detection accuracy of 95.55% (mAP50) and a dangerous person discrimination accuracy of 82.5% within a 1 - 2 meter range, demonstrating near real-time threat recognition on a Jetson Orin NX -based on-device AI platform.

### Keywords

digital twin, security robot, threat detection, weapon recognition, real-time monitoring, ai surveillance

\* 동의대학교 인공지능학과

- ORCID<sup>1</sup>: <https://orcid.org/0009-0003-0328-5742>

- ORCID<sup>2</sup>: <https://orcid.org/0009-0004-5031-9447>

\*\* (주)아가비타

- ORCID<sup>3</sup>: <https://orcid.org/0009-0008-7726-1889>

\*\*\* 동의대학교 인공지능로봇연구소(교신저자)

- ORCID: <http://orcid.org/0000-0002-9036-0872>

· Received: Sep. 24, 2025, Revised: Nov. 27, 2025, Accepted: Nov. 30, 2025

· Corresponding Author: Seung-Ho Ok

Dept. of Robot Engineering, Dong-eui University, Korea

Tel.: +82-51-890-2265, Email: osh@deu.ac.kr

## 1. 서 론

최근 인공지능(AI), 자율주행, 센서 융합 등 첨단 정보기술의 발전으로 이동형 방법 로봇이 도시 안전 및 보안 분야에 도입되기 시작했다. 기존의 CCTV 기반 감시나 인력 순찰 방식은 고정된 시야와 인력 한계로 인해 감시 사각지대가 발생하고, 실시간 대응에도 한계가 존재한다[1][2]. 이러한 한계를 극복하기 위해, 자율주행과 센서 융합, AI 기반 실시간 데이터 분석 기능을 갖춘 이동형 방법 로봇이 차세대 보안 솔루션으로 주목받고 있다[2][3].

이러한 기술 발전을 기반으로, 국내외에서는 이동형 방법 로봇을 활용한 보안 시스템 개발이 활발히 이루어지고 있다. 국내에서는 HL만도가 개발한 ‘골리(Goalie)’와 SK텔레콤, 뉴빌리티, SK쉴더스가 공동으로 개발한 순찰 로봇이 다양한 도시 환경에 도입되어 자율주행, 열화상 카메라, 라이다 센서, AI 기반 객체 인식 등 첨단 기술을 적용하고 있다. 해외에서는 미국 나이트스코프(Knightscope)사의 K5 로봇이 대표적으로, 360도 영상 감시, 라이다 센서, GPS, 음성 경고, 실시간 데이터 전송 등 다양한 기능을 탑재하여 상업시설, 공공장소, 주차장 등에서 자율 순찰과 이상 상황 감지 임무를 수행하고 있다. 또한, 현대자동차의 공장 안전 순찰 로봇, 일본-미국 대학 캠퍼스, 국내 대학 연구단지 등에서도 실내외 공간을 대상으로 24시간 순찰, 실시간 영상 전송, 출입자 인증, 응급상황 대응 등 다양한 임무를 수행하는 사례가 보고되고 있다.

국내외 연구에서는 이러한 이동형 방법 로봇이 실제 현장에 적용되어, LiDAR와 카메라 기반 실시간 객체 인식, 경로 순찰, 이상행동 감지, 관리자 알림 등 기존 CCTV 시스템의 한계를 보완함을 실험적으로 검증하였다[2][4]. 더불어 AI와 로봇 기술의 융합은 도시 환경과 실내 공간에서 실시간 데이터 수집, 예측적 대응, 반복적 순찰, 위험 지역 접근성 향상 등 보안 서비스의 자동화와 효율화에 기여하고 있다[3][4].

하지만 이동형 방법 로봇과 같은 첨단 기술의 도입에도 불구하고, 복잡한 도시 및 실내 환경에서는 실시간 상황 인식, 직관적 모니터링, 다양한 데이터

의 통합 분석 등 한계가 여전히 존재한다[5][6].

이에 따라 최근에는 방법 시스템의 감시 및 대응 능력을 한층 더 향상시키기 위해 디지털 트윈(Digital twin) 기술이 적극적으로 도입되고 있다. 디지털 트윈은 현실 세계의 데이터를 실시간으로 수집하여 가상 환경에서 이를 시뮬레이션하고 시각화하는 기술로, 보안 및 감시 시스템에 적용할 경우 더욱 정밀하고 직관적인 감시 체계를 구축할 수 있다. 특히, 디지털 트윈을 활용하면 기존 CCTV나 로봇 센서가 수집한 데이터를 3D 가상 공간에서 실시간으로 반영할 수 있어 보다 효과적인 감시 및 대응이 가능하다[7][8].

디지털 트윈 기반의 모니터링 시스템은 여러 장점을 갖는다. 첫째, 3D 가상 환경을 통해 실시간으로 상황을 시각적으로 표현할 수 있어 직관적인 모니터링이 가능하다. 둘째, 로봇과 연계하여 위험 요소를 분석하고 실시간 경보 기능을 강화할 수 있다. 셋째, 기존 보안 시스템과 연동하여 데이터의 정확성을 높이고, 원격지에서도 감시 및 대응이 가능하도록 설계할 수 있다[9][10]. 이러한 기술적 발전은 기존 CCTV 및 인력 기반 순찰 방식의 한계를 보완하며, 보다 신속하고 정확한 위협 탐지가 가능하게 한다[8][9].

최근 연구에 따르면, 디지털 트윈은 데이터 통합, 실시간 시뮬레이션, 네트워크 및 다양한 응용 분야에서 활용도가 높으며, 감시 및 방법 시스템에서도 중요한 역할을 할 수 있음이 강조되고 있다[8][10]. 또한, 실제 국내외 사례에서도 디지털 트윈을 기반으로 한 원격 감시 및 위협 분석 시스템이 활발히 연구·적용되고 있다[7][10].

본 논문에서는 실내 환경에서도 상황을 해석하고 위협 상황을 식별할 수 있는 상황인식 기반 위협인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템을 제안한다. 제안된 시스템은 이동형 방법 플랫폼을 기반으로 사람의 위치 추정·신원 인식·무기 탐지를 통합 수행하는 AI 기반 통합 인식 시스템과, 인식 결과를 3차원 가상 공간에 실시간 반영하는 디지털 트윈 연동 체계로 구성된다. 이를 통해 현실 공간에서 수집된 데이터가 실시간으로 디지털 트윈 환경에 반영되어, 보안 관리자가 공간 내 상황과 위

협 상황을 직관적으로 파악할 수 있도록 설계되었다. 또한, 본 연구는 사람의 신원과 무기 소지 여부를 통합적으로 인식하는 AI 기반 위협 인식 구조를 구현함으로써, 기존 영상 중심 보안 시스템의 공간적 한계를 보완하고 실제 상황을 해석할 수 있는 통합 방법 플랫폼의 가능성을 제시하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 제안 시스템의 구조와 동작 원리를 설명하고, 제3장에서는 사람 위치 추정, 무기 탐지, 얼굴 인식 기능의 실험 설계 및 구현을 다룬다. 제4장에서는 디지털 트윈 연동 결과와 탐지 정확도 평가를 제시하며, 제5장에서는 결론 및 향후 연구 방향을 논의한다.

## II. 본 론

최근 자율 순찰 로봇은 실내 환경에서 보안 및 감시 효율성을 향상시키는 핵심 기술로 주목받고 있다. 인공지능(AI) 기반 객체 인식과 자율주행 기술이 결합되면서 로봇이 일정 구역을 자율적으로 순찰하고 사람을 인식하며 이상 상황을 감지하는 다양한 연구가 이루어지고 있으나, 기존 시스템들은 여전히 실시간 위협 인식과 3차원 시각화 측면에서 한계를 보인다. 본 장에서는 기존 시스템의 구조적 한계를 분석하고, 이러한 문제를 개선하기 위해 제안한 디지털 트윈 기반 지능형 방법 로봇 시스템의 구성과 동작 흐름을 기술한다.

기존 연구에서 제안된 방법 로봇 시스템은 크게 두 가지 형태로 구분된다. 첫 번째 유형은 영상 기반 자율 추적 로봇(Vision-based autonomous tracking robot)으로, 로봇이 지정된 구역을 순찰하며 카메라로 획득한 영상을 원격지 서버로 전송하여 관리자가 이를 모니터링하는 방식이다[2]. 그러나 이 시스템은 단순 객체 인식과 추적 기능에 초점이 맞추어져 있어 무기나 위협 행위의 탐지는 불가능하며, 로봇의 시야 밖 영역이나 복잡한 공간 구조에서는 감시 시각지대가 발생한다. 두 번째 유형은 스마트 자율 순찰 로봇(Smart autonomous patrol robot)으로, LiDAR와 카메라, 초음파 센서 등을 복합적으로 활용해 자율 순찰 및 장애물 회피 기능을 수행한다[3]. 그러나 주로 2차원 기반의 영상 분석을 사용하

므로 공간적 이해가 제한적이며, 사람의 신원 인식이나 위협 행위 탐지와 같은 고차원 지능형 판단 기능이 포함되어 있지 않다. 표 1은 기존 방법 로봇 시스템과 본 연구에서 제안하는 시스템의 주요 기능을 비교한 것이다. 기존 시스템은 공통적으로 자율 순찰 기능은 제공하지만, AI 기반 위협 인식, 얼굴 인식에 의한 출입자 식별, 디지털 트윈 기반 3차원 시각화와 같은 통합 기능은 지원되지 않는다. 반면 본 연구의 시스템은 이러한 한계를 극복하기 위해 자율 순찰, 사람 및 무기 탐지, 신원 인식, 그리고 실시간 3차원 시각화 기능을 통합하여 구현하였다. 이를 통해 실제 환경에서 발생하는 위협 상황을 실시간으로 탐지하고, 관리자가 3D 공간상에서 직관적으로 상황을 파악할 수 있도록 설계하였다.

표 1. 기존 방법 로봇 시스템과 제안 시스템의 기능 비교  
Table 1. Feature comparison between existing systems and the proposed system

Feature	Case 1: Vision-based autonomous tracking robot	Case 2: Smart autonomous patrol robot	Ours
Autonomous patrol	Yes	Yes	Yes
Threat detection (Weapon)	No	No	Yes
Digital twin visualization	No	No	Yes
3D spatial monitoring	No	Limited (2D/2.5)	Yes

### 2.1 시스템 개요

본 연구에서 제안하는 상황인식 기반 위협인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템은 이동형 로봇 플랫폼을 기반으로 사람의 위치 추정·신원 인식·무기 탐지를 통합 수행하는 AI 기반 통합 인식 시스템과 실시간 시각화를 위한 디지털 트윈 연동 체계로 구성된다. 제안된 시스템은 단순한 객체 인식 단계를 넘어, 자세 추정과 객체·얼굴 인식 정보를 융합하여 위협 상황을 인지할 수 있는 구조를 갖는다. 로봇은 실시간 영상 데이터를 수집

하여 얼굴 인식을 통해 구성원 여부를 판단하고, 손 좌표 기반의 자세 추정과 무기 탐지 결과를 결합하여 실제 무기 소지 여부를 판별한다. 특히 본 연구에서는 단순히 영상 내 무기 존재 여부만을 판단하는 기존 방식과 달리, 사람의 관절 좌표 정보를 활용한 상황 기반 무기 판별을 통해 오탐(False positive)을 최소화하였다. 최종적으로 인식된 사람의 위치·신원·위험 여부 정보는 디지털 트윈 환경에 실시간으로 연동되며, 3차원 가상 공간 내 가상 인물을 통해 시각적으로 표현된다. 이를 통해 현실 공간의 인식 결과를 가상 공간에 실시간으로 반영함으로써, 위험 상황을 직관적으로 시각화하고 실시간 감시 효율성을 향상시킨다.

그림 1은 본 연구에서 제안하는 방법 로봇 시스템의 전체 구조를 나타낸다. 그림의 왼쪽 영역은 실제 환경(Real world)을, 오른쪽은 가상 환경(Virtual world)을 나타낸다. 로봇은 실제 환경에서 사람을 감지하고 얼굴 인식, 자세 추정, 무기 탐지를 수행하며, 그 결과를 디지털 트윈 공간으로 전송한다.

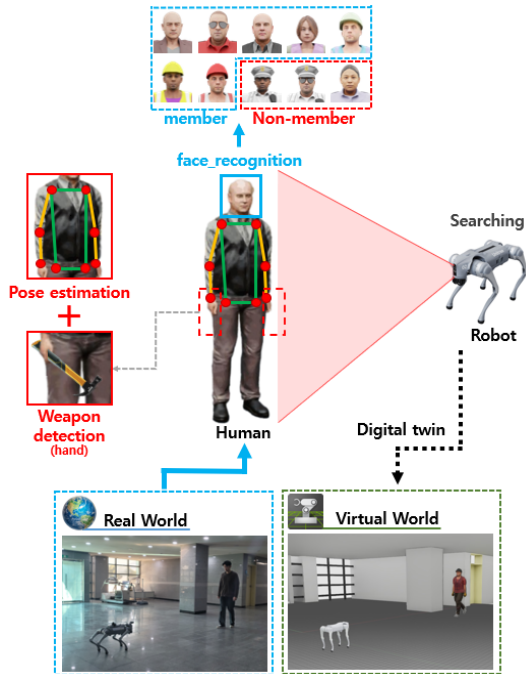


그림 1. 상황 인식 기반 위험인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템 개요

Fig. 1. Overview of the context-aware security robot system for dangerous person discrimination with digital twin integration

디지털 트윈 환경에서는 실제 건물 구조와 동일한 3차원 공간이 구현되며, 감지된 사람은 상태에 따라서 다른 색상으로 표시된다. 구성원은 녹색, 외부인은 빨간색, 위험인물은 파란색으로 시각화되어 보안 관리자가 직관적으로 상황을 파악할 수 있다.

## 2.2 시스템 흐름도

본 시스템의 전체 동작 절차는 사람 위치 추정, 신원 인식, 위험인물 판별, 디지털 트윈 시각화의 네 단계로 구성된다. 그림 2는 본 연구에서 제안하는 시스템의 동작 과정을 순차적으로 나타낸 것이다.

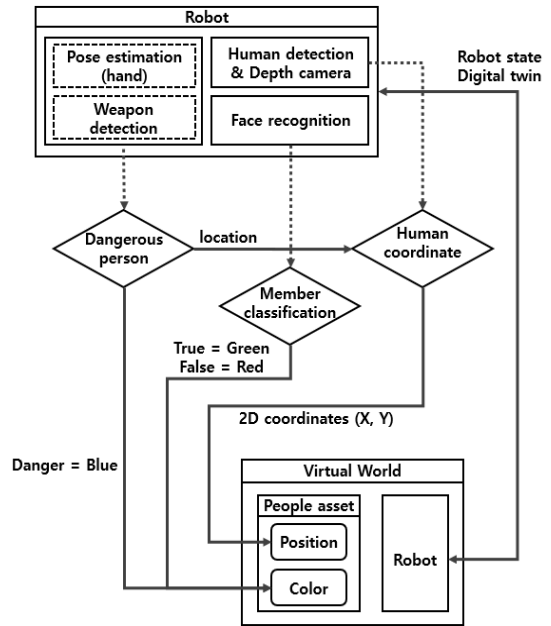


그림 2. 상황 인식 기반 위험인물 판별 및 디지털 트윈 방법 로봇 시스템의 동작 흐름도

Fig. 2. Operation flow of the context-aware security robot system for dangerous person discrimination with digital twin integration

시스템의 첫 번째 단계에서는 로봇이 탑재된 깊이 카메라를 이용해 주변 환경을 인식한다. 객체 탐지 모델을 기반으로 입력 영상에서 사람을 검출하고, 감지된 객체의 중심 좌표( $u, v$ )와 깊이 정보( $d$ )를 이용해 실제 3차원 좌표( $X, Y, Z$ )를 계산한다.

$$X = \frac{(u - c_x) \cdot d}{f_x}, Y = \frac{(v - c_y) \cdot d}{f_y}, Z = d \quad (1)$$

이때 카메라의 광학 중심( $c_x, c_y$ )과 초점 거리( $f_x, f_y$ )를 고려하여 공간 좌표를 변환하고, 회전 행렬 R과 병진 벡터 T를 적용하여 로봇 기준 좌표계로 매핑한다. 이러한 변환 과정을 통해 로봇은 탐지된 사람의 위치를 실시간으로 추정하며, 이는 디지털 트윈 환경 내 가상 인물의 위치로 반영된다.

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T \quad (2)$$

사람의 위치가 파악되면, 시스템은 두 번째 단계로서 얼굴 인식 기반 신원 인식 절차를 수행한다. 탐지된 인물의 얼굴 영역은 HOG(Histogram of Oriented Gradients) 기반 검출기를 통해 추출되고, CNN(Convolutional Neural Network)을 활용하여 얼굴 특징을 128차원 임베딩 벡터로 변환한다. 변환된 벡터는 사전에 등록된 데이터베이스(Member's) 내 구성원 벡터와 유클리드 거리를 기준으로 비교된다. 이때 유사도가 임계값 이하일 경우 동일 인물로 판단되어 신원이 일치한 것으로 간주되며, 데이터베이스 등록 여부에 따라 구성원(Member) 또는 외부인(Non-member)으로 분류된다. 그림 3은 얼굴 인식을 통해 구성원 여부를 판별하고 인물의 신원을 식별하는 과정을 보여준다.

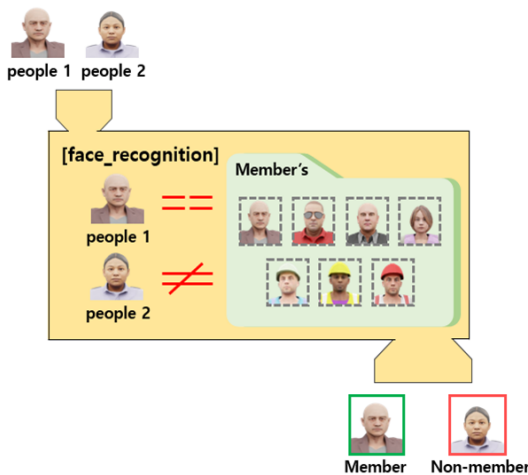


그림 3. 얼굴 인식 기반 구성원 판별 및 신원 식별 과정  
Fig. 3. Membership discrimination and identity recognition process based on facial recognition

인식 결과는 디지털 트윈 환경에서 녹색(구성원)과 빨간색(외부인) 가상 인물로 실시간 반영된다.

세 번째 단계에서는 자세 추정과 무기 탐지 과정이 수행된다. 자세 추정 모델을 이용해 인식된 인물의 주요 관절 좌표(어깨, 팔꿈치, 손목, 무릎 등)를 검출하고, 손 좌표를 중심으로 무기 탐지 모델의 결과를 결합한다. 무기 탐지 모델은 칼, 총, 망치, 방망이 등 네 가지 무기류를 대상으로 학습되었으며, 탐지된 무기 객체의 좌표가 손 좌표 반경 내에 위치할 경우 실제 무기 소지로 판단한다. 특히 본 연구에서는 무기 객체가 단순히 영상에 존재하는 것만으로 위험인물로 판단하지 않고, 손 좌표 근처에서 1초간 연속적으로 탐지될 경우에만 실제 무기 소지로 판정한다. 이를 통해 배경 물체나 일시적 감지 오류로 인한 오탐을 방지한다.

마지막 단계에서는 탐지된 모든 정보를 디지털 트윈 환경과 연동하여 시각화한다. 사람의 위치, 신원, 위험 여부는 실시간으로 3D 가상 공간에 가상 인물 생성 및 색상 변화를 통해 반영되며, 구성원은 녹색, 외부인은 빨간색, 무기를 소지한 위험인물은 파란색으로 표시된다. 이와 같은 시각화 방식을 통해 보안 관리자는 실제 현장에 가지 않고도 실시간으로 상황을 파악할 수 있으며, 공간적 맥락을 고려한 즉각적인 대응이 가능하다. 결과적으로 제안된 시스템은 기존 영상 기반 모니터링의 한계를 실험적으로 보완할 수 있음을 입증하였다.

### III. 실험

본 장에서는 제안된 방법 시스템의 핵심 기능을 검증하기 위한 실험 과정을 다룬다. 실험은 세 단계로 구성되었다. 첫 번째 단계에서는 이동형 방법 로봇을 활용하여 실시간으로 사람의 위치를 탐지하였다. 두 번째 단계에서는 자세 추정과 무기 탐지를 결합하여 무기 소지 여부를 분석하고 위험인물을 판별하는 실험을 진행하였다. 마지막으로, 얼굴 인식 기술을 적용하여 인물의 신원을 식별하고, 데이터베이스 등록 여부에 따라 구성원과 외부인을 판별하는 실험을 수행하였다. 모든 실험은 Intel RealSense D435i 깊이 카메라와 YOLOv8 기반 객체 탐지 모델

과 자세 추정 모델을 활용하였으며, 대학 캠퍼스 건물 내부 복도의 실제 환경에서 진행되었다. 실험에는 서로 다른 신원을 가진 두 명의 인물이 동시에 등장하는 환경을 구성하여, 다인 상황에서도 안정적인 탐지와 인식이 가능한지를 확인하였다.

### 3.1 실시간 사람 위치 추정 실험

첫 번째 실험은 로봇이 자율 순찰 중에 감지한 사람의 위치를 실시간으로 추정하고, 이를 로봇 기준 좌표계로 변환하여 정확도를 검증하는 데 목적이 있다. 로봇은 실내 경로를 따라 이동하면서 주변 인물을 탐지하고, 깊이 카메라를 통해 획득한 거리 정보를 기반으로 상대 좌표(X, Y, Z)를 계산하였다. 계산된 좌표는 2차원 좌표계에 시각적으로 표시되어 실제 감지 위치와의 일치 여부를 확인하였다. 실험 결과, 로봇이 탐지한 실제 사람의 이동 경로와 계산된 좌표가 일관되게 대응하였으며, 두 명의 인물이 동시에 존재하더라도 각각의 위치가 명확히 분리되어 표현되었다. 이를 통해 제안된 시스템이 실내 구조에서도 사람의 위치를 안정적으로 추정할 수 있음을 확인하였다.

그림 4는 사람의 위치를 로봇 기준 좌표로 변환하고 이를 좌표계에 표시한 결과를 보여준다. 좌측은 실험환경에서 로봇에 설치된 카메라 이미지이며, 우측은 로봇의 위치를 중심으로 카메라에 인식된 사람의 상대 좌표이다. 이러한 좌표 정보는 이후 단계에서 디지털 트윈 환경 내 가상 인물 생성에 활용되어, 현실 공간의 인식 결과를 3차원 가상 공간과 연동하는 기반으로 사용된다.

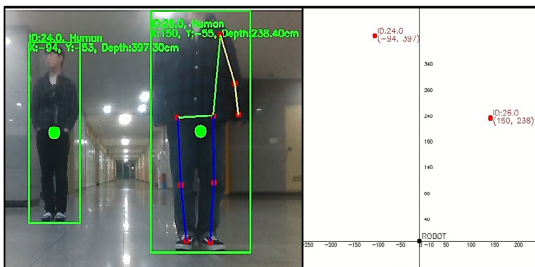


그림 4. 사람 인식 및 로봇 기준 좌표 변환 결과  
Fig. 4. Human detection and transformation into the robot-relative coordinate system

### 3.2 자세 추정 및 무기 탐지를 통한 위험인물 판별 실험

두 번째 실험은 이동형 방법 로봇이 사람을 인식한 후, 자세 추정과 무기 탐지 결과를 결합하여 실제로 무기를 소지한 위험인물을 판별할 수 있는지를 검증하는 것이다. 실험은 두 명의 인물이 동시에 등장하는 환경에서 진행되었으며, 한 인물은 일반 상태로, 다른 인물은 손에 무기를 들고 이동하는 시나리오로 구성하였다. 이를 통해 제안된 시스템이 다인 환경에서도 안정적으로 위험인물을 구별할 수 있는지를 확인하였다.

두 명 이상의 환경에서는 단순 객체 탐지 모델을 적용할 경우, 인물 간 경계가 겹치거나 손 좌표가 중첩되어 잘못된 무기 판별이 이루어지는 문제가 발생한다. 이러한 문제를 해결하기 위해 본 연구에서는 자세 정보와 시간적 일관성을 결합한 프레임 연속 검증 방식을 적용하였다. 이 방식은 각 인물의 손 좌표를 기준으로 무기 객체의 위치를 추적하며, 동일한 손에서 1초간 연속적으로 탐지될 때만 실제 무기 소지로 판단한다. 그림 5는 이러한 중첩 인식 문제와 제안된 개선 방식의 효과를 시각적으로 보여준다.

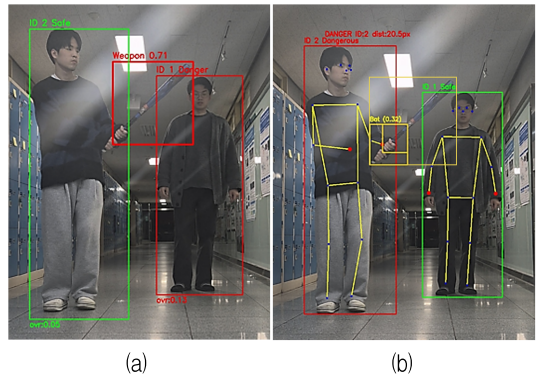


그림 5. 다인 환경에서의 중첩 및 경계 겹침 문제와 제안 방식의 비교 (a) 객체 탐지 모델만 적용, (b) 제안된 방식  
Fig. 5. Comparison of overlapping and boundary interference in multi-person environments  
(a) Using only the object detection model,  
(b) Proposed method

왼쪽은 기존 단순 객체 탐지 모델을 적용한 결과

로, 인물 간 경계가 겹치는 구간에서 한 인물의 손 근처에 있는 물체가 다른 인물의 무기로 잘못 인식된 사례를 나타낸다. 반면 오른쪽의 제안 방식은 자세 기반 손 좌표 추적과 연속 프레임 검증을 결합하여 각 인물의 손 위치를 독립적으로 추적함으로써 이러한 오류를 효과적으로 방지하였다. 이를 통해 제안된 시스템은 다인 환경에서도 실제 무기 소지만을 안정적으로 판별할 수 있음을 확인하였다.

이후 로봇은 인물의 자세 정보를 이용하여 손 좌표를 기준으로 무기 존재 여부를 분석하였다. 동일한 손에서 1초간 연속적으로 무기가 감지될 경우 실제 무기 소지로 판정하였으며, 이를 통해 스마트폰이나 우산처럼 유사한 형태의 비위험 물체로 인한 오탐을 효과적으로 줄였다. 그림 6은 이러한 손 좌표 기반 탐지 방식의 결과를 보여준다. 그림 6(a)는 일반적인 상황으로 무기가 탐지되지 않은 경우이며, 그림 6(b)는 손 주변에서 무기가 탐지되어 위험인물로 판별된 사례를 나타낸다.

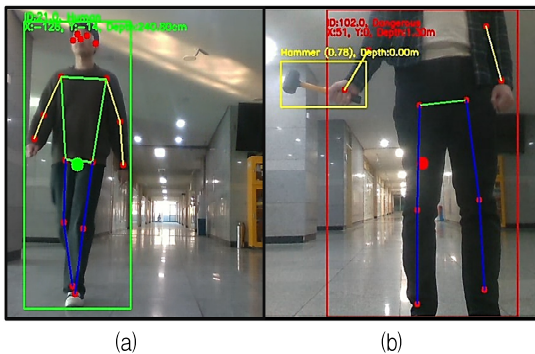


그림 6. 손 좌표 기반 무기 탐지 결과  
 (a) 무기가 탐지되지 않은 상황, (b) 무기가 탐지된 상황  
 Fig. 6. Weapon detection results based on hand position  
 (a) Weapon not detected, (b) Weapon detected

다음으로 그림 7은 네 가지 무기류(칼, 망치, 총, 방망이)에 대해 제안된 시스템을 적용한 탐지 결과를 시각적으로 나타낸다. 각 무기 유형은 실제 손에 들린 상태에서 전반적으로 안정적인 탐지 결과를 보였으며, 형태적 차이에 따라 세부적인 인식 성능에는 다소 편차가 나타났다. 칼은 방향 변화에도 mAP50 기준 평균 95.55%의 탐지 결과를 보였고, 총은 세밀한 형상으로 인해 평균 검출률(Recall)

86.17% 수준을 유지하였다. 망치와 방망이는 크기가 크고 형태가 단순하여 각각 평균 92.96%, 93.26%의 높은 탐지 정확도를 기록하였다. 이러한 결과는 제안된 자세 추정 기반 무기 탐지 방식이 무기 종류에 관계없이 실제 손 소지 여부를 안정적으로 식별할 수 있음을 보여주며, 세부 성능 수치는 4장에서 정량적으로 분석하였다.

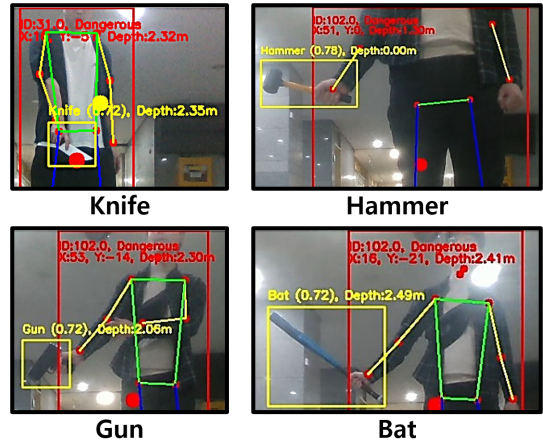


그림 7. 무기별 탐지 결과  
 Fig. 7. Detection results by weapon type

### 3.3 얼굴 인식을 통한 구성원 판별 실험

본 실험은 이동형 방범 로봇이 탐지한 인물의 얼굴 정보를 이용하여, 인물의 신원을 식별하고 데이터베이스 등록 여부에 따라 구성원과 외부인을 판별하는 성능을 검증하기 위해 수행되었다. 실험은 대학 캠퍼스 건물 내부 복도 환경에서 진행되었으며, 총 10명의 인물 데이터를 활용하였다. 이 중 7명은 데이터베이스에 등록된 구성원으로, 나머지 3명은 등록되지 않은 외부인으로 설정하였다. 실험 시 두 명의 인물이 동시에 등장하는 환경을 구성하여, 로봇이 실시간으로 얼굴을 인식하고 신원을 식별하는 과정을 검증하였다. 로봇은 탐지된 얼굴 특징을 데이터베이스와 비교하여 구성원 여부를 판별하였으며, 그 결과는 디지털 트윈 환경에 반영되어 시각적으로 표현되었다.

그림 8은 두 인물의 얼굴 인식 결과를 통해 신원을 식별하고 구성원 여부를 구분한 모습을 보여준다. 왼쪽 인물은 데이터베이스에 등록된 구성원으로

인식되어 녹색 프레임으로 표시되었고, 오른쪽 인물은 등록되지 않은 외부인으로 판별되어 빨간색 프레임으로 표시되었다. 이러한 결과는 제안된 시스템이 실시간으로 인물의 신원을 식별하고, 구성원 여부를 정확히 판별하여 그 결과를 디지털 트윈 환경에 효과적으로 반영할 수 있음을 보여준다.

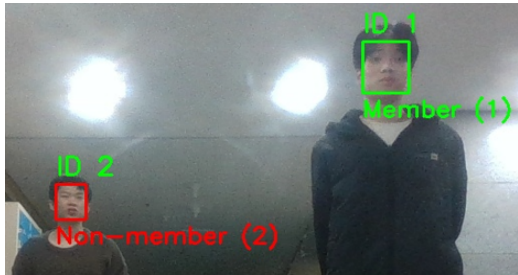


그림 8. 신원 인식 결과  
Fig. 8. Identity recognition results

#### IV. 실험 결과 및 성능 평가

본 장에서는 제안된 상황인식 기반 위협인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템의 실험 결과와 정량적 성능 분석을 다룬다. 특히, 실제 환경에서 감지된 사람의 신원·자세·무기 정보를 통합적으로 해석하여 위협 상황을 판단하는 능력과, 이를 디지털 트윈 환경에 실시간으로 반영하는 성능을 검증하였다. 앞서 3장에서 수행된 실험 데이터를 기반으로, 실제 환경에서 감지된 사람의 위치, 신원, 위험 여부를 가상 환경에 연동하고, 무기 탐지 및 거리별 판별 성능을 평가하였다. 가상 환경 구현에는 NVIDIA Isaac Sim을 사용하였으며, 실험 데이터는 Unitree Go2 로봇과 Jetson Orin NX 플랫폼을 통해 수집되었다. 본 장에서는 디지털 트윈 연동 결과를 시각적으로 제시하고, 무기 탐지 모델의 정량적 성능과 거리 변화에 따른 판별 정확도를 분석하였다.

##### 4.1 디지털 트윈 연동 결과

로봇이 감지한 사람의 위치 데이터를 기반으로, 가상 환경에서도 동일한 상대 위치에 가상 인물을 생성하여 현실과 가상 공간의 동기화가 제대로 이

루어지는지 확인하였다. Isaac Sim의 Omni.Anmi.People 기능을 활용해 가상 인물을 생성하고, 로봇이 인식한 2D 상대 좌표를 3D 좌표로 변환하여 실시간으로 반영하였다. 그림 9는 이러한 디지털 트윈 연동 결과를 나타낸다. 좌측은 실제 환경에서 로봇이 사람을 감지한 모습이며, 우측은 동일한 상대 좌표를 기반으로 가상 공간에 생성된 가상 인물의 위치를 보여준다. 실제와 가상 공간의 좌표가 일치하며, 로봇의 시점 변화에도 위치 동기화가 유지되는 것을 확인하였다.

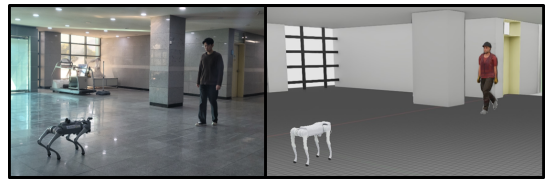


그림 9. 디지털 트윈 연동 결과  
Fig. 9. Digital twin integration results

다음으로, 구성원 판별 실험의 결과를 가상 공간에 반영하였다. 그림 10은 신원 인식 결과를 디지털 트윈 환경에 시각적으로 적용한 예시이다. 데이터베이스와의 비교를 통해 인식된 인물의 상태에 따라, 구성원은 초록색, 외부인은 빨간색으로 표시되었다.

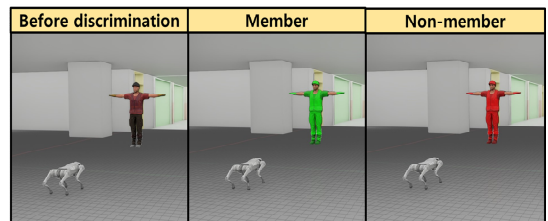


그림 10. 신원 인식 결과에 따른 가상 인물 색상 반영  
Fig. 10. Visualization of virtual human color representation based on identity recognition results

그리고 일반적인 상황에서는 모든 인물이 동일하게 표시되지만, 위협인물이 탐지된 경우에는 해당 인물만 가상 환경에 파란색 가상 인물로 생성되어 시각적으로 강조된다. 그림 11은 일반 상황(a)과 위협인물이 탐지된 위험 상황(b)을 비교하여, 실제 환경(좌측), 로봇 기준 2D 위치(중앙), 디지털 트윈 환경(우측)에 반영된 가상 인물의 변화를 보여준다. 일반 상황에서는 모든 인물이 그대로 표시되지만,

위험 상황에서는 무기를 소지한 인물만 파란색으로 강조되어 생성되며, 그 위치와 상태가 실시간으로 업데이트된다. 이와 같이, 제안된 시스템은 현실에서 인식된 사람의 위치, 신원, 위험 여부 정보를 가상 환경에 연동하여, 보안 담당자가 공간 전체의 상황을 직관적으로 파악하고 신속한 대응이 가능하도록 지원한다.

#### 4.2 성능 분석

무기 탐지 모델의 성능은 Precision, Recall, mAP50, mAP50-95 지표를 통해 정량적으로 평가하였다. YOLOv8 모델을 이용하여 4종의 무기(망치, 칼, 총, 방망이)에 대해 100 Epoch 학습을 진행하였으며, 각 항목별 성능 결과는 표 2에 제시하였다.

표 2. 무기 탐지 모델 성능  
Table 2. Weapon detection model performance

Model	Precision	Recall	mAP50	mAP50 -95
Hammer	0.9007	0.8694	0.9296	0.7130
Knife	0.9059	0.9034	0.9555	0.7528
Gun	0.8799	0.8617	0.9160	0.7012
Bat	0.8979	0.8801	0.9326	0.7186

칼(Knife)의 탐지 성능이 가장 우수하여 mAP50 기준 95.55%, mAP50-95 기준 75.28%의 높은 값을 기록하였다. 반면, 총기(Gun)는 mAP50 기준 91.60%, mAP50-95 기준 70.12%로 상대적으로 낮은 성능을 보였으며, 이는 학습 데이터의 불균형과 총기와 유사한 형태의 객체와의 구분이 어려운 환경적 요인에 기인한 것으로 판단된다. 이러한 결과는 무기 종류와 데이터 특성에 따라 탐지 성능에 차이가 있음을 시사하며, 향후 데이터의 다양성 확보와 알고리즘 정밀도 향상이 필요함을 보여준다.

또한, 사람과 로봇 간 거리 변화가 탐지 정확도에 미치는 영향을 분석하기 위해, 거리를 1m 이내, 1~2m, 2~3m, 3m 이상 네 구간으로 나누어 성능을 비교하였다. 거리별 위험인물 판별 정확도는 표 3에 제시하였다.

1m 이내에서는 인식이 원활하지 않았으며, 1~2m 구간에서 가장 높은 정확도(82.5%)를 나타냈다. 그러나 2~3m 구간에서는 정확도가 32.5%로, 3m 이상에서는 17.5%까지 감소하는 등 거리 증가에 따라 탐지 성능이 현저히 저하되는 경향을 보였다. 이러한 성능 저하는 거리 증가로 인해 영상 내 객체 크기가 작아지고, 무기나 신체 특징의 세부 식별이 어려워지기 때문으로 분석된다.

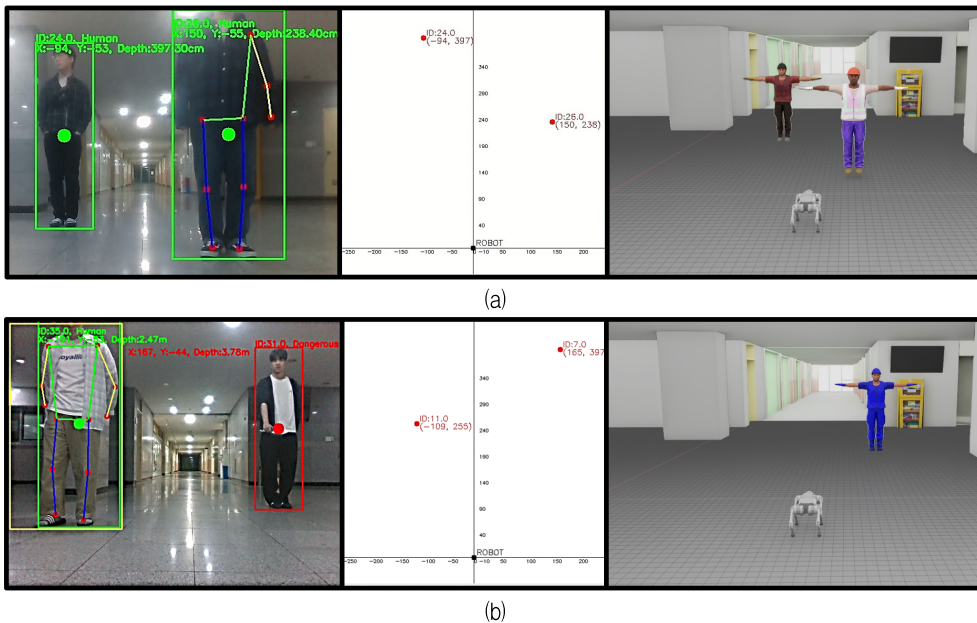


그림 11. 일반 상황과 위험 상황의 디지털 트윈 시각화 비교 (a) 일반 상황, (b) 위험 상황  
Fig. 11. Comparison of digital twin visualization between normal and dangerous situations  
(a) Normal situation, (b) Dangerous situation

표 3. 거리별 위협인물 판별 정확도

Table 3. Accuracy of dangerous person discrimination by distance

Distance	Discrimination accuracy
Within 1m	Not recognized
1m ~ 2m	0.825
2m ~ 3m	0.325
Over 3m	0.175

종합적으로, 제안된 시스템은 근거리(1~2m) 구간에서 가장 우수한 탐지 성능을 보였으나, 장거리 상황에서는 인식률이 감소하는 한계를 보였다. 따라서 향후 연구에서는 고해상도 영상 처리 기술과 탐지 알고리즘의 정밀화, 그리고 데이터셋의 다양성과 균형 확보를 통해 장거리 환경에서도 안정적인 탐지가 가능하도록 개선할 필요가 있다. 한편, Jetson Orin NX에서 YOLOv8 기반 무기 탐지 모델의 평균 추론 속도는 약 133.26ms/frame(약 7.5fps)으로 측정되었다. 이는 일반적인 영상 처리에서 요구되는 30fps 수준의 실시간성에는 미치지 못하지만, 위협인물 판별 및 경보 발령과 같은 보안 로봇 응용에는 충분히 활용 가능한 수준임을 확인하였다. 이를 통해 본 시스템이 실제 환경에서 on-device AI 기반의 준실시간 위협인물 탐지 및 대응이 가능함을 실증하였다.

## V. 결론 및 향후 과제

본 논문에서는 사람의 행동 의도와 위협 상황을 통합적으로 인식할 수 있는 상황인식 기반 위협인물 판별 및 디지털 트윈 연동형 방법 로봇 시스템을 설계하고 구현하였다. 제안된 시스템은 YOLOv8 기반 객체 탐지, 얼굴 인식, 자세 추정, 무기 탐지 기능을 통합하여 사람의 위치, 신원, 위협 여부를 실시간으로 분석하고 이를 디지털 트윈 환경에 3차원 가상 공간으로 시각화함으로써 실시간 보안 상황의 통합적 인식과 직관적 대응이 가능하도록 하였다.

실험은 대학 캠퍼스 건물 내부 환경에서 Unitree Go2 사족보행 로봇과 NVIDIA Jetson Orin NX를 이용한 on-device AI 환경에서 수행되었다. 사람 위치 추정, 신원 인식, 무기 탐지 기능이 정상적으로 통

합되어 실시간으로 동작하였으며, 실제 환경에서 획득된 데이터를 Isaac Sim 기반 디지털 트윈 환경과 연동함으로써 현실과 가상의 상태가 동기화됨을 확인하였다.

성능 분석 결과, 제안된 시스템은 Knife(칼)에서 mAP50 기준 95.55%, Gun(총기)에서 91.60%의 높은 탐지 정확도를 보였으며, 1~2m 구간에서 가장 우수한 위협인물 판별 성능(82.5%)을 기록하였다. Jetson Orin NX에서의 평균 추론 속도는 약 133.26ms/frame(7.5fps)로 측정되어, 완전한 실시간성(30fps)에는 미치지 못하지만 보안 로봇 응용 수준에서는 준실시간 대응이 가능한 성능임을 입증하였다.

본 연구에서는 기존 방법 로봇 시스템이 수행하지 못했던 인식의 통합성과 상황 이해 능력을 중심으로 한 새로운 보안 로봇 구조를 제시하였다. 사람의 신원 인식, 자세 추정, 무기 탐지 기능을 결합하여 단순한 객체 인식을 넘어 실제 위협 행위를 판별할 수 있는 상황 인식 체계를 구현하였으며, 이를 통해 로봇이 주변 환경의 맥락을 해석하고 능동적으로 대응할 수 있는 기반을 마련하였다. 또한 인식된 결과를 실시간으로 3차원 디지털 트윈 환경에 반영함으로써 물리적 공간과 가상 공간 간의 상태를 동기화하고, 보안 관리자가 직관적으로 상황을 파악할 수 있는 시각화 체계를 구축하였다. 이러한 구조는 기존 CCTV 중심의 수동적 감시 방식에서 벗어나, 로봇이 중심이 되는 능동적 보안 관리 체계로 발전할 수 있는 가능성을 보여주었다.

향후 연구에서는 고해상도 카메라 및 다중 센서 융합을 통한 장거리 인식 정밀도 향상, 데이터셋 다양화 및 균형 확보, 알고리즘의 경량화 및 실시간 최적화를 통해 시스템의 완성도를 높일 계획이다. 또한, 디지털 트윈과 로봇 간의 양방향 상호작용을 구현하여 스마트 시티·캠퍼스·산업 현장 등에서의 실시간 보안 자동화로 확장 적용할 예정이다.

## References

- [1] S. Lee, "A study of legislation for AI patrol robot", *Yonsei Law Review*, Vol. 40, pp. 277-306, Oct. 2022. <https://doi.org/10.33606/YLA.40.10>.
- [2] E. S. Woo, T. H. Kim, and D. H. Lee, "Design

- of an Indoor Security Robot System based on Autonomous Navigation and Object Detection", Journal of KIIT, Vol. 21, No. 8, pp. 19-25, Aug. 2023. <https://doi.org/10.14801/jkiit.2023.21.8.19>.
- [3] N. Basilio, "Recent Trends in Robotic Patrolling", Current Robotics Reports, Vol. 3, pp. 65-76, May 2022. <https://doi.org/10.1007/s43154-022-00078-5>.
- [4] J. H. Chang, K. I. Na, and H. C. Shin, "Trend of Technology for Outdoor Security Robots based on Multimodal Sensors", Electronics and Telecommunications Trends, Vol. 37, No. 1, pp. 1-9, Feb. 2022. <https://doi.org/10.22648/ETRI.2022.J.370101>.
- [5] G. Xu, A. S. Khan, A. J. Moshayedi, X. Zhang, and Y. Shuxin, "The Object Detection, Perspective and Obstacles In Robotic: A Review", EAI Endorsed Transactions on AI and Robotics, Vol. 1, pp. 1-9, Oct. 2022. <https://doi.org/10.4108/airo.v1i1.2709>.
- [6] H. Bavl, J. L. Sanchez-Lopez, C. Cimorelli, A. Tourani, and H. Voos, "From SLAM to Situational Awareness: Challenges and Survey", Sensors, Vol. 23, No. 10, pp. 1-34, May 2023. <https://doi.org/10.3390/s23104849>.
- [7] H. Wu, P. Ji, H. Ma, and L. Xing, "A Comprehensive Review of Digital Twin from the Perspective of Total Process: Data, Models, Networks and Applications", Sensors, Vol. 23, No. 19, pp. 1-20, Oct. 2023. <https://doi.org/10.3390/s23198306>.
- [8] Y. Mousavi, Z. Gharineiat, A. A. Karimi, K. McDougall, A. Rossi, and S. G. Barsanti, "Digital Twin Technology in Built Environment: A Review of Applications, Capabilities and Challenges", Smart Cities, Vol. 7, No. 5, pp. 2594-2615, Sep. 2024. <https://doi.org/10.3390/smartcities7050101>.
- [9] S. Choi, J. Woo, J. Kim, and J. Y. Lee, "Digital Twin-Based Integrated Monitoring System: Korean Application Cases", Sensors, Vol. 22, No. 14, pp. 1-14, Jul. 2022. <https://doi.org/10.3390/s22145450>.
- [10] M. El-Hajj, "Leveraging Digital Twins and Intrusion Detection Systems for Enhanced Security in IoT-Based Smart City Infrastructures",

Electronics, Vol. 13, No. 19, pp. 1-24, Oct. 2024. <https://doi.org/10.3390/electronics13193941>.

## 저자소개

### 이 동 섭 (Dong-Sup Lee)



2025년 2월 : 동의대학교  
로봇·자동화공학(공학사)  
2025년 3월 ~ 현재 : 동의대학교  
인공지능학과 석사과정  
관심분야 : Autonomous Driving,  
Digital Twin, Robotics

### 박 주 환 (Joo-Hwan Park)



2024년 2월 : 동의대학교  
로봇·자동화공학(공학사)  
2026년 2월 : 동의대학교  
인공지능학과(공학석사)  
2026년 3월 ~ 현재 : 아가비타  
연구원  
관심분야 : Digital Twin,

Sim2Real

### 전 채 욱 (Chae-Uk Jeon)



2024년 2월 : 동의대학교  
로봇·자동화공학(공학사)  
2026년 2월 : 동의대학교  
인공지능학과(공학석사)  
2026년 3월 ~ 현재 : 동의대학교  
인공지능학과 박사과정  
관심분야 : Vision Language

Model, Robotics Control

### 옥 승 호 (Seung-Ho Ok)



2008년 2월 : 경북대학교  
전자공학과(공학석사)  
2014년 2월 : 경북대학교  
전자공학부(공학박사)  
2014년 3월 ~ 2017년 8월 :  
삼성전자 책임연구원  
2017년 9월 ~ 현재 : 동의대학교

로봇자동화공학과 부교수(인공지능로봇연구소 소장)  
관심분야 : On-Device AI, System-on-Chip, VLSI