

모바일 인터페이스 기반 사용자 맞춤형 자율주행 로봇 시스템

고운별*¹, 김건웅*², 정현준*³, 정동원*⁴

Mobile Interface-based Personalized Autonomous Robot System

Unbyeol Ko*¹, Geonwoong Kim*², Hyunjun Jung*³, and Dongwon Jeong*⁴

본 연구는 2026년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음
(2023-0-00065)

요 약

실제 환경에서 사람 추종 로봇은 특정 대상을 지속적으로 재식별해야 하지만, 개인화된 등록과 즉각적인 제어 수단이 부족하다. 이 논문은 사용자가 모바일 앱으로 추종 대상을 직접 등록하고 제어할 수 있는 사용자 맞춤형 자율주행 로봇 시스템을 제안한다. OSNet 기반 Re-ID 모델이 다섯 방향의 전신 사진에서 대표 특징 벡터를 생성하고, 검출 - 추적 - 주행 인지 파이프라인과 MQTT 기반 실시간 제어를 하나의 시스템으로 통합하였다. 실험 결과, 동일 인물과 타인의 평균 유사도가 0.8069와 0.4992로 안정적으로 구분되었으며, 실제 주행에서 등록된 사용자만을 약 1.5m 거리로 추종하고 Jetson Xavier NX에서 온디바이스 실시간 동작을 확인하였다. 향후 거리 적응형 추종과 LiDAR 센서 융합을 통해 대상 상실과 장애물 회피 문제를 개선하고자 한다.

Abstract

In real-world environments, a person-following robot must continuously re-identify a specific target, yet personalized registration and immediate control remain lacking. This paper proposes a user-customized autonomous robot system in which a user registers and controls the target via a mobile app. An OSNet-based Re-ID model generates a representative feature vector from five-direction full-body images, integrated with a detection - tracking - navigation pipeline and MQTT-based control in a single system. Experiments showed average similarities of 0.8069 and 0.4992 for the same and different persons, and the robot followed only the registered user at about 1.5 m in real time on a Jetson Xavier NX. Future work will address target loss and obstacle avoidance via distance-adaptive following and LiDAR fusion.

Keywords

personalized, re-identification, autonomous robot, human-following, mobile Interface, OSNet

* 국립군산대학교 소프트웨어학과(*^{3,4} 공동교신저자)

- ORCID¹: <https://orcid.org/0009-0008-9432-5573>

- ORCID²: <https://orcid.org/0009-0009-1812-9949>

- ORCID³: <https://orcid.org/0000-0002-6717-1395>

- ORCID⁴: <https://orcid.org/0000-0001-9881-5336>

• Received: Aug. 12, 2026, Revised: Sep. 04, 2026, Accepted: Sep. 07, 2026

• Co-corresponding Author1: Hyunjun Jung

Dept. of Software Science & Engineering, Kunsan National University, 558, Daehak-ro, Gunsan, Jeollabuk-do, Republic of Korea

Tel.: +82-63-469-8911, Email: junghj85@kunsan.ac.kr

• Co-corresponding Author2: Dongwon Jeong

Dept. of Software Science & Engineering, Kunsan National University, 558, Daehak-ro, Gunsan, Jeollabuk-do, Republic of Korea

Tel.: +82-63-469-8911, Email: djeong@kunsan.ac.kr

I. 서 론

현재 글로벌 로봇 시장은 빠르게 성장하고 있으며, 2024년 기준 전 세계 산업용 로봇 설치 시장 규모는 약 167억 달러에 도달하였다[1]. 한국, 미국, 일본 등 주요 국가에서는 노동력 부족 문제를 해결하기 위해 로봇 자동화가 필수적으로 요구되고 있으며, 로봇 산업은 스스로 학습하고 경험을 통해 작동하는 Physical AI로 패러다임이 전환되고 있다. 특히 사람을 추종하는 자율주행 로봇은 개인 보조, 물류 운송, 공공시설 관리, 고령자 지원 등 다양한 분야에서 활용 가능성이 높아 주목받고 있다[2][3]. 이러한 추종 로봇이 실제 환경에서 활용되기 위해서는 대상 인식, 자율 주행, 사용자 인터페이스가 하나의 시스템으로 통합되어 안정적으로 동작해야 한다.

그러나 실제 환경에서의 사람 추종은 여전히 어려운 과제이다. 기존 대상 추적 연구의 상당수는 방해 요소가 적고 배경이 단순한 통제된 실험 환경을 중심으로 평가되어 있으며, 다수의 보행자가 존재하는 혼잡한 실제 환경에서는 빈번한 가림(Occlusion)과 유사한 외형의 인물로 인해 추종 대상을 상실하는 문제가 발생한다[4][5]. 이는 추종 로봇이 단순한 인물 검출을 넘어, 특정 대상을 지속적으로 재식별할 수 있는 능력을 갖추어야 함을 의미한다.

이러한 문제에 대응하기 위한 연구들이 진행되고 있다. 세 가지 운용 모드를 갖춘 폐기물 수집·이송 로봇[6]과 의복 색상 기반의 추종 방식[7]이 제안되었으나, 전자는 추종 대상의 개인화된 식별 기능이 부재하고 후자는 사전 등록 색상에 의존한다는 한계가 있다. 최근에는 Re-ID(Person Re-Identification) 기술을 추종 로봇에 적용하여 부분 가림 대응[5], 온라인 연속 학습 기반 재식별[8], 경량 모델을 통한 임베디드 실시간 처리[9] 등이 연구되었으나, 사용자가 추종 대상을 직접 등록하고 제어하는 인터페이스는 다뤄지지 않았다. 한편 모바일 기기 기반의 로봇 제어 연구[10]는 원격 조종과 모니터링 수준에 머물러 있으며, 자율 기능만으로는 돌발 상황에서의 충돌을 완전히 방지하기 어렵다는 점에서 [11] 사용자의 즉각적인 개입 수단 또한 시스템 차

원에서 요구된다.

이상의 기존 연구들은 공통적으로 다음과 같은 한계를 보인다. 첫째, 추종 대상을 사용자가 동적으로 등록하는 개인화 기능이 부재하다. 둘째, 색상 마커나 고정된 특징에 의존하여 가림과 유사 외형 인물이 존재하는 복잡한 환경에서의 강인성이 부족하다. 셋째, 사용자가 추종 시작과 정지를 즉각적으로 제어할 수 있는 인터페이스가 미비하여 돌발 상황 대응에 취약하다.

이러한 문제를 해결하기 위해 본 논문에서는 모바일 인터페이스 기반의 사용자 맞춤형 자율주행 로봇 시스템을 제안한다. 제안 시스템은 사용자가 모바일 앱에서 다섯 방향의 전신 사진을 등록하면 경량 Re-ID 모델이 대표 특징 벡터를 생성하여 로봇이 해당 사용자만을 선택적으로 추종하도록 하며, MQTT(Message Queuing Telemetry Transport) 기반 실시간 통신을 통해 사용자가 주행을 즉각적으로 시작·정지하고 로봇의 시야를 원격으로 모니터링할 수 있는 제어 수단을 함께 제공한다. 전체 시스템은 Jetson Xavier NX 임베디드 환경과 ROS2 기반 아키텍처 위에 구현되었으며, 실환경 주행 실험을 통해 등록된 사용자만을 선별적으로 추종하고 일정 거리를 안정적으로 유지함을 확인하였다.

본 논문의 기여는 다음과 같다. 첫째, 모바일 앱 기반의 대상 등록, Re-ID 기반 식별, 추종 주행, 즉각적인 사용자 제어를 하나의 시스템으로 통합한 아키텍처를 설계하고 구현하였다. 둘째, 검출, 재식별, 추적, 주행으로 이어지는 인지 파이프라인을 임베디드 온디바이스 환경에서 실시간으로 동작하도록 구성하였다. 셋째, 실환경 실험을 통해 제안 시스템의 사용자 선별 추종과 제어 동작을 시스템 수준에서 검증하였다. 2장에서는 관련 연구를, 3장에서는 제안 시스템의 설계를, 4장에서는 실험 및 평가를 기술하며, 5장에서는 결론을 맺는다.

II. 관련 연구

2.1 사람 추종 로봇 시스템

대형 공공시설의 청소 작업을 자동화하기 위해

수동 조작, 자율 주행, 작업자 추종의 세 가지 운용 모드를 갖춘 폐기물 수집·이송 로봇이 제안된 바 있다[6]. 이 시스템은 작업자를 따라 이동하며 폐기물 운반을 보조함으로써 추종 로봇의 산업적 활용 가능성을 보여주었다. 그러나 추종 대상을 구분하는 개인화된 식별 기능이 없어, 다수의 인물이 존재하는 환경에서 특정 작업자만을 선택적으로 추종하기 어렵고, 현장에서 변경하는 것도 고려되지 않았다.

스테레오 카메라 기반의 이동 로봇이 대상의 의복 색상과 질감을 사전에 등록하고 이를 단서로 특징인을 검출하고 추종하는 방식도 연구되었다[7]. 이는 비전 기반 특징인 추종의 초기 접근으로서 의의가 있으나, 의복 정보의 사전 등록이라는 준비 과정을 요구하고 색상이라는 단일 단서에 의존하므로, 유사한 색상의 의복을 착용한 인물과 혼동되거나 조명에 따른 색상 변화에 취약하여 실용성에 한계가 있다.

이들 연구는 각각 운용 모드의 다양화와 비전 기반 추적이라는 측면에서 추종 로봇의 가능성을 보여주었다. 그러나 두 연구 모두 추종 대상이 사전에 정의되거나 별도의 표식에 의존하는 구조로 설계되어, 사용자가 추종 대상을 동적으로 등록하고 변경하는 기능은 다루지 않았다.

2.2 Person Re-Identification 기반 대상 인식

Re-ID는 마커와 같은 인위적 표식 없이 대상의 외형 특징을 벡터로 표현하고 유사도를 비교하여 동일 인물 여부를 판별하므로, 추종 중 대상을 일시적으로 놓치더라도 재등장 시 다시 식별할 수 있다는 점에서 추종 로봇의 대상 인식에 적합하다. 혼잡한 실환경에서 발생하는 가림과 유사 외형 인물로 인한 대상 상실 문제[4]에 대응하기 위해, 외형 특징 기반의 인물 재식별(Re-ID) 기술을 추종 로봇에 적용하려는 연구가 진행되어 왔다.

부분 가림 상황에서의 추종 문제를 다룬 연구[5]는 대상의 일부만 관측되는 조건에서도 추종을 지속하는 방법을 제시하였다. 온라인 연속 학습 기반의 Re-ID 프레임워크를 활용한 연구[8]는 추종 중 수집되는 데이터로 외형 표현을 점진적으로 갱신함

으로써, 사전 학습 데이터와 실제 운용 환경 간의 차이로 인한 성능 저하를 완화하고 외형 변화와 가림, 유사 외형 인물이 존재하는 상황에서의 재식별 성능을 향상시켰다. 또한 다양한 스케일의 특징을 동적으로 융합하는 경량 신경망인 OSNet이 제안되어[9], 적은 파라미터 수로도 높은 재식별 성능을 달성함으로써 임베디드 환경에서의 실시간 Re-ID 적용 가능성을 넓혔다.

이들 연구는 재식별 알고리즘의 성능 향상에 초점을 두고 있으며, 사용자가 추종 대상을 직접 등록하고 로봇의 동작을 제어하는 시스템 수준의 통합은 연구 범위에 포함하지 않았다. 즉, 우수한 재식별 성능이 확보되더라도 이를 실제 사용자가 활용할 수 있는 형태로 연결하는 인터페이스 계층은 여전히 공백으로 남아 있다.

특히 추종 대상의 등록이 사전에 준비된 데이터나 운용 중 자동 수집에 의존하므로, 새로운 사용자가 현장에서 자신을 추종 대상으로 즉시 등록하는 시나리오는 지원되지 않는다.

2.3 모바일 인터페이스 기반 로봇 제어

모바일 기기를 로봇 운용의 인터페이스로 활용하려는 연구로, 안드로이드 기기와 ROS를 연동하여 실외 로봇과 상호작용할 수 있는 휴대용 HRI 프레임워크가 제안되었다[10]. 이 연구는 휴대 기기와 로봇 간의 연동 구조를 체계화하여, 스마트폰이 별도의 전용 장비 없이 로봇 제어와 모니터링 수단이 될 수 있음을 보여주었다. 그러나 범용적인 제어·연동 프레임워크의 구축에 초점을 두고 있어, 추종 대상 등록과 같은 개인화 기능이나 로봇의 자율 기능과 사용자 제어를 하나의 운용 흐름으로 통합하는 관점은 고려되지 않았다.

한편, 사람과 같은 공간에서 동작하는 서비스 로봇에는 안전성 확보가 필수적이다. 서비스 로봇의 안전 제어에 관한 연구[11]에 따르면, 센서 기반의 사전 충돌 회피 기법은 환경 정보의 불완전성과 엄격한 실시간 처리 요구로 인해 예측 불가능한 돌발 상황에서의 충돌을 전부 방지하기 어렵다. 이는 자율 기능에만 의존하는 대신, 사용자가 로봇의 동작

을 즉각적으로 개시하고 중지할 수 있는 개입 수단이 함께 제공되어야 함을 시사한다. 그러나 추종 로봇을 대상으로 이러한 개입 수단을 모바일 인터페이스와 결합하여 다룬 사례는 찾아보기 어렵다.

2.4 기존 연구와의 비교

표 1은 앞서 살펴본 기존 연구들과 제안 시스템을 여섯 가지 기능 항목을 기준으로 비교한 것이다. 비교 항목은 Re-ID 기반 대상 식별(Re-ID), 사용자의 동적 대상 등록(Target registration), 사용자 인터페이스 제공(User interface), 모바일 연동(Mobile integration), 즉각적인 시작/정지 제어(Instant control), 임베디드 온디바이스 처리(On-device)이다.

각 항목은 해당 기능의 지원 여부에 따라 ○(지원), △(부분 지원), ×(미지원)으로 표기하였으며, 연구의 성격상 해당 항목이 평가 범위 외인 경우는 - (해당 없음)으로 구분하여 불공정한 비교를 방지하였다. 예를 들어 OSNet[9]은 재식별 모델 자체에 관한 연구로서 로봇의 제어나 인터페이스는 평가 대상이 아니므로 해당 항목들을 -로 표기하고, 폐기물 수집 로봇[6]은 전용 조작기 기반의 수동 모드를 제공하므로 사용자 인터페이스 항목에서 완전한 지원이 아닌 부분 지원(△)으로 비교한다.

먼저 폐기물 수집 로봇[6]은 전용 조작기 기반의 수동 모드를 통해 제한적인 사용자 인터페이스를 제공하지만, 추종 대상의 식별과 등록 기능은 지원

하지 않는다. 색상 마커 방식[7]은 온디바이스 추적 이 가능하나 마커에 의존하므로 나머지 항목을 모두 만족하지 못한다. Re-ID 연구[8][9]는 외형 기반 대상 식별을 지원하고 임베디드 적용 가능성을 보였으나, 대상 등록·제어를 위한 인터페이스는 연구 범위에 포함하지 않았다. 반대로 ROS-Mobile[10]은 모바일 인터페이스와 즉각적인 제어를 제공하지만, Re-ID 기반 식별이나 추종 기능은 결합되지 않았다.

주목할 점은 이러한 공백이 상호 보완적인 형태로 나타난다는 것이다. 인식 계열 연구[8][9]는 사용자와의 접점 없이 식별 성능만을 확보하였고, 인터페이스 계열 연구[10]는 식별 능력 없이 제어 수단만을 제공하여, 어느 한쪽만으로는 사용자 맞춤형 추종이 성립하지 않는다. 결과적으로 기존 연구들은 여섯 항목 중 일부만을 개별적으로 만족하며, 전체를 하나의 시스템으로 통합한 사례는 부재하다. 이러한 기존 연구의 한계점들을 바탕으로, 본 논문은 사용자가 모바일 앱을 통해 자신의 사진을 등록하면 Re-ID 모델이 외형 기술자를 추출하고, 로봇이 해당 사용자만을 추종하며, 사용자가 주행을 즉시 제어할 수 있는 자율주행 로봇 시스템을 제안한다. 구체적으로, 사용자가 동적으로 등록한 사진으로부터 대표 특징 벡터를 생성하여 대상을 식별하고, 자율주행 중에는 MQTT 기반의 즉각적 제어와 실시간 카메라 모니터링을 제공하며, 검출·재식별·추적·주행으로 이어지는 파이프라인은 영상을 외부로 전송하지 않고 보드 위에서 온디바이스로 동작한다.

표 1. 기존 연구와의 기능 비교

Table 1. Functional comparison with existing studies

Study	Re-ID	Target registration	User interface	Mobile integration	Instant control	On-device
Waste-collecting robot [6]	×	×	△	×	×	○
Clothing-color following [7]	×	×	×	×	×	○
Online continual Re-ID [8]	○	×	×	×	×	△
OSNet [9]	○	-	-	-	-	△
ROS-Mobile [10]	×	×	○	○	○	-
Proposed system	○	○	○	○	○	○

III. 모바일 연동 Re-ID 기반 사용자 맞춤형 추종 로봇 시스템

3.1 통합 아키텍처

제안 시스템은 모바일 앱을 통한 사용자 등록·제어, 임베디드 환경에서의 실시간 인지, 그리고 추종 주행 제어를 하나의 파이프라인으로 통합한 사용자 맞춤형 자율주행 로봇 시스템이다. 그림 1은 제안 시스템의 통합 아키텍처를 나타내며, 시스템은 네 가지 계층으로 구성된다.

첫째, HMI(Human-Machine Interface) 계층은 React Native 기반 모바일 앱으로 구성되며, 사용자 사진 등록, 자율주행 시작/정지 명령, 수동 조종, 로봇 상태 모니터링 기능을 담당한다. 둘째, Intelligence 계층은 YOLOv11 기반 보행자 검출과 OSNet 기반 Re-ID 매칭을 수행하여 등록된 사용자를 식별한다. 셋째, Navigation 계층은 ROS2의 NAV2 스택을 기반으로 식별된 대상의 위치를 목표점으로 하는 경로를 생성하고, 대상과 일정 거리를 유지하는 추종 주행 제어를 담당한다. 넷째, Hardware Control Plane은 상위 계층의 주행 명령을 받아 Arduino와 VESC를 통해 모터와 서보 모터를 구동하는 저수준

제어 계층이다. 각 계층은 Main Integration Hub를 중심으로 연결되며, Main Integration Hub는 센서 데이터를 융합하고 MQTT 프로토콜을 통해 모바일 앱과의 실시간 통신을 수행하면서 계층 간 데이터 흐름을 중계한다. 즉, 영상 인식 결과와 로봇 상태 정보, 사용자의 제어 명령이 모두 이 허브를 거쳐 해당 계층으로 전달된다. 이러한 계층형 구조는 각 계층의 독립적인 개발과 수정을 가능하게 하며, 특정 모듈의 교체가 전체 시스템에 미치는 영향을 최소화한다.

이러한 계층 구조가 동작하는 하드웨어 기반은 저전력 임베디드 환경에서 GPU 가속 추론이 가능한 Jetson Xavier NX 보드를 중심으로 구성된다. Intelligence 계층과 Navigation 계층의 연산은 Xavier NX에서 수행되며, 영상 인식을 위한 카메라가 인지 파이프라인의 입력을 제공한다. Hardware Control Plane에 해당하는 Arduino는 IMU 센서 데이터를 실시간으로 융합하여 PID 제어 기반의 차체 평형 유지와 서보 모터 구동을 담당하는 저수준 제어기로 동작하고, 구동 모터는 VESC를 통해 제어된다. 전원은 LiPo 배터리로 공급되며, 모니터와 키보드는 개발 및 현장 디버깅 용도로만 사용되어 시스템 운용에는 관여하지 않는다.

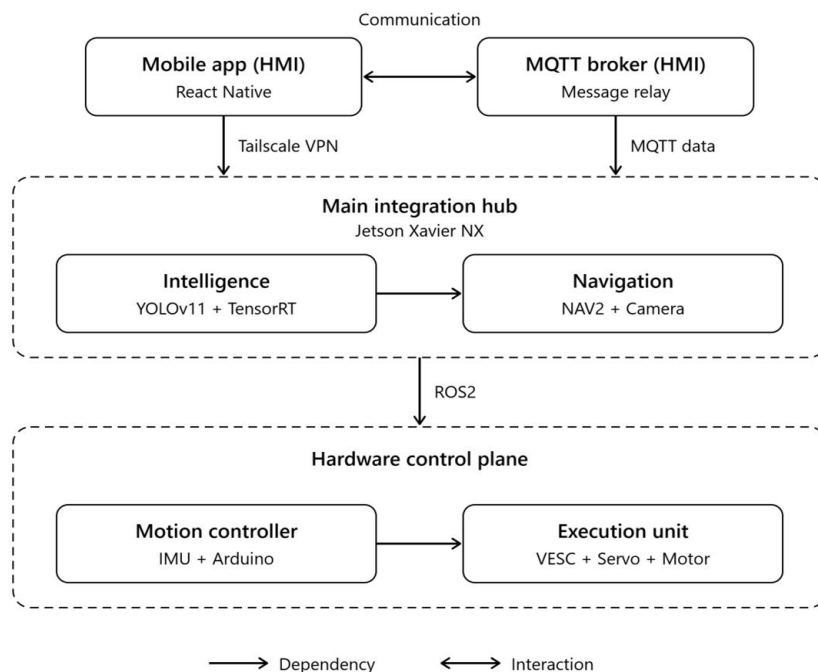


그림 1. 제안 시스템의 통합 아키텍처

Fig. 1. Integrated architecture of the proposed system

3.2 지능형 인지 파이프라인

인지 파이프라인은 그림 2와 같이 보행자 검출, 대상 재식별, 대상 추적의 세 단계로 구성되며, 각 단계의 출력이 다음 단계의 입력으로 연결되는 순차 구조로 동작한다.

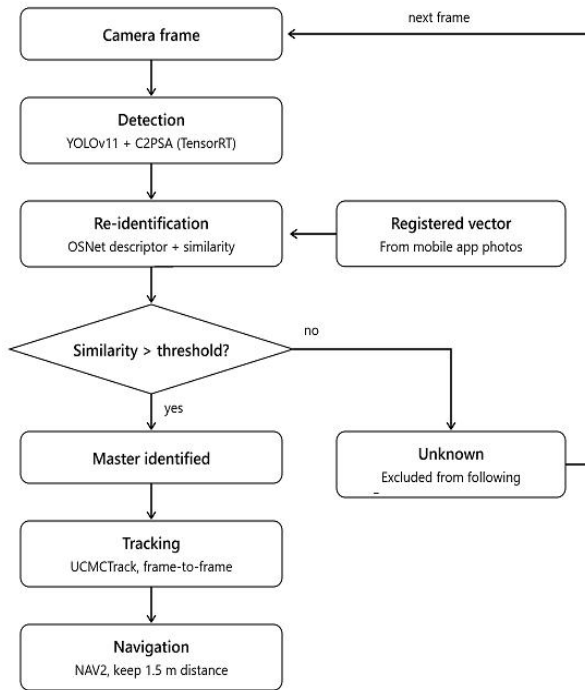


그림 2. 인지 파이프라인 처리 절차

Fig. 2. Processing procedure of the perception pipeline

첫째, 검출 단계에서는 카메라로부터 입력된 매 프레임에 대해 공간적 주의집중 모듈(C2PSA)이 통합된 YOLOv11이 보행자를 검출한다[12]. C2PSA 모듈은 특징 맵의 공간적 중요 영역에 가중치를 부여하여, 배경 잡음이 많은 환경에서도 보행자 영역에 집중된 특징 추출을 가능하게 한다. 검출 모델은 TensorRT로 최적화되어 Jetson Xavier NX의 제한된 연산 자원에서 실시간 추론이 가능하도록 구성하였다.

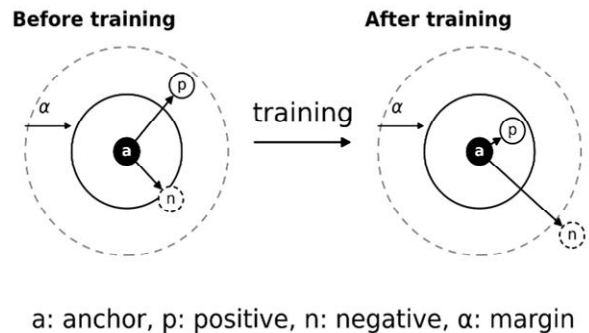
둘째, 재식별 단계에서는 검출된 각 보행자 영역에서 OSNet 기반 Re-ID 모델(osnet_ain_x1_0)[9]이 외형 기술자를 추출하고, 그림 2에 나타난 바와 같이 모바일 앱에서 사전에 등록된 사용자의 특징 벡터를 입력받아 두 벡터 간의 유사도를 계산한다. 유사도는 0과 1 사이의 값으로 산출되며, 1에 가까울수록 동일인일 확률이 높음을 의미한다. 이후 유사

도가 임계값을 초과하는지에 따라 처리 경로가 분기된다. 임계값을 초과하는 인물은 추종 대상(MASTER)으로 판별되어 추적 단계로 전달되고, 그 외의 모든 인물은 미등록 인물(UNKNOWN)로 분류되어 추종 대상에서 제외된다. 이러한 분기 구조는 등록된 사용자가 프레임에 존재하지 않는 경우 로봇이 어떤 인물도 추종하지 않도록 보장하며, 유사한 외형의 인물을 잘못 추종하는 위험을 낮춘다.

셋째, 추적 단계에서는 MASTER로 판별된 대상에 한하여 UCMCTrack 알고리즘[13]이 위치를 프레임 간 연속적으로 추적한다. 재식별이 매 프레임의 독립적인 판별이라면, 추적은 프레임 간 위치의 연속성을 활용하여 대상의 궤적을 유지하므로, 일시적인 검출 실패나 유사도 변동이 발생하더라도 추종의 연속성이 보존된다. 추적된 대상의 위치는 NAV2 스택의 주행 목표로 전달되며, NAV2가 생성한 주행 명령을 통해 로봇은 대상과 약 1.5m의 거리를 유지하며 추종한다.

3.3 Re-ID 모델 학습

Re-ID 모델의 학습에는 Market-1501 데이터셋을 사용하였다[14]. Market-1501은 총 32,668장의 이미지와 1,501명의 고유 인물 ID를 포함하며, 6대의 카메라에서 촬영된 다양한 각도의 이미지로 구성되어 가림 현상과 배경 잡음을 자연스럽게 포함한다. 이는 제한된 임베디드 하드웨어 환경에서 모델의 일반화 성능을 확보하기에 적합한 데이터셋이다.



a: anchor, p: positive, n: negative, α : margin

그림 3. Triplet loss 기반 특징 공간 최적화
Fig. 3. Feature space optimization based on triplet loss

손실 함수는 Cross-Entropy Loss와 Triplet Loss[15]

를 결합한 하이브리드 구조를 사용한다.

Cross-Entropy Loss는 각 인물 이미지를 정확한 ID로 분류하도록 학습시키며, Triplet Loss는 그림 3과 같이 동일 인물 간의 특징 거리는 가깝게, 서로 다른 인물 간의 거리는 멀게 배치하여 특징 공간을 최적화한다. 두 손실의 결합을 통해 분류 정확성과 특징 공간의 판별력을 동시에 확보할 수 있다.

데이터 증강 기법으로는 세 가지를 적용하였다. 첫째, Random Erasing은 이미지의 일부 영역을 무작위로 지워 가림 현상에 대응하는 강인성을 확보한다. 둘째, Color Jitter는 밝기와 색상을 무작위로 변형하여 조명 변화에 대한 강인성을 부여한다. 셋째, Horizontal Flip은 좌우 반전을 통해 대상의 방향 변화에 대한 불변성을 확보한다. 이 세 기법은 추종 로봇의 실제 환경 운용에서 발생하는 변동 요인에 대응하도록 선정하고, 최적화에는 Adam Optimizer를 사용하여 적응형 학습률 조정과 안정적인 수렴을 달성하였으며, 모델의 학습 및 평가는 Torchreid 프레임워크를 기반으로 수행하였다.

3.4 모바일 인터페이스

모바일 앱은 React Native 기반으로 구현되었으며, 로그인, 메인 대시보드, 타겟 등록, 실시간 카메

라, 활동 로그의 다섯 가지 화면으로 구성된다. 화면 구성은 대상 등록, 주행 제어, 상태 확인의 세 운용 단계를 최소한의 조작으로 수행할 수 있도록 설계하였으며, 그림 4는 이 중 핵심 기능인 메인 대시보드와 타겟 등록 화면을 나타낸다.

로그인 화면은 회원가입과 이메일 기반 인증을 처리하여, 여러 사용자가 하나의 로봇을 사용하는 환경에서도 각자의 등록 사진과 특징 벡터가 독립적으로 관리되도록 한다. 그림 4(a)는 메인 대시보드로, 자율주행 시작/정지 버튼, 수동 조종 패드, 타겟 등록 진입 기능을 하나의 화면에 통합하였으며, 상단의 상태 표시부를 통해 대상 인식 상태와 유사도, 조향각, 스로틀 값을 실시간으로 확인할 수 있다. 수동 조종 패드는 자율주행 모드와 독립적으로 동작하여, 자율 추종이 어려운 구간에서는 직접 조종으로 전환할 수 있다. 그림 4(b)는 타겟 등록 화면으로, 로봇 카메라로 다섯 방향에서 촬영한 전신 사진을 등록하며, 등록과 인식이 동일한 카메라 재식별 일관성과 방향 변화에 대한 강인성이 확보된다. OSNet 모델이 각 사진에서 전신 외형 특징 벡터를 추출한 뒤, 특정 방향에 치우치지 않는 단일 표현을 얻기 위해 이를 평균하여 대표 특징 벡터를 생성하고, 데이터베이스에 저장하여 로봇에 전송한다.

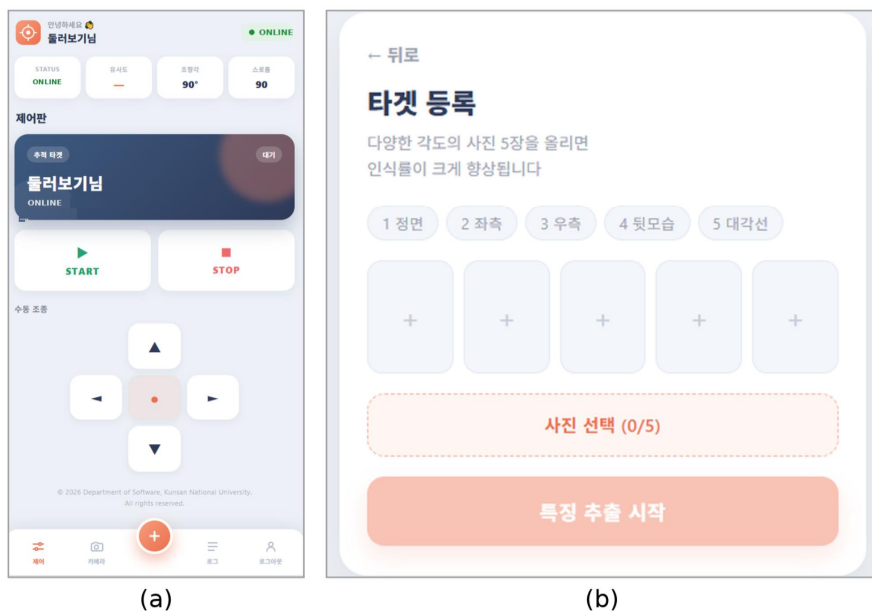


그림 4. 모바일 인터페이스 화면 구성

Fig. 4. Screen composition of the mobile interface

3.5 데이터 관리

데이터베이스는 표 2와 같이 사용자 정보(users), 업로드 사진(owner_images), Re-ID 특징 벡터(owner_descriptors), 로봇 상태(robot_status), 제어 이력(control_logs)의 다섯 개 테이블로 설계하였다. users 테이블을 중심으로 사용자와 대표 특징 벡터는 1:1, 사용자와 사진·로봇 상태·제어 이력은 각각 1:N 관계로 구성되어, 사용자별 등록 정보와 운용 이력을 일관성 있게 관리한다.

특징 벡터를 사용자당 하나의 대표 벡터로 유지하는 것은 평균 기반 대표 벡터 생성과 대응되는 설계로, 등록 사진이 추가되거나 갱신될 때 대표 벡터가 일관성 있게 재생성되도록 한다. 로봇 상태 테이블에는 온라인 여부와 조향각, 유사도 점수 등이 주기적으로 기록되어 모바일 앱 상태 표시부의 데이터 원천이 되며, 제어 명령이 control_logs에 기록되므로 운용 이력의 추적과 오류 상황의 사후 분석이 가능하다.

표 2. 데이터베이스 테이블 구성

Table 2. Database table composition

Table	Key attribute	Relation
users	Email, password, name, created at	—
owner_images	Image path, uploaded at	1:N with users
owner_descriptors	Re-ID feature vector, image count, updated at	1:N with users
robot_status	Online status, steering angle, similarity score	1:N with users
control_logs	Control mode, issued at	1:N with users

IV. 실험 평가 및 사용자 평가

4.1 실험 환경

표 3은 실험 환경을 나타낸다. Re-ID 모델은 Market-1501 데이터셋으로 120 epoch 학습하였으며,

추론과 주행 실험은 Jetson Xavier NX가 탑재된 실제 로봇 플랫폼에서 수행하였다. 추종 대상 판별의 유사도 임계값은 0.75로 설정하였다.

표 3. 실험 환경

Table 3. Experimental environment

Category	Specification
Embedded board	NVIDIA Jetson Xavier NX
Detection model	YOLOv11 (C2PSA), TensorRT
Re-ID model	osnet_ain_x1_0 (Torchreid)
Training dataset	Market-1501
Training epochs	120
Similarity threshold	0.75
Middleware	ROS2, NAV2, MQTT

4.2 Re-ID 모델 평가

학습된 모델의 성능은 데이터셋 기반 평가와 실제 환경 비교군 평가로 검증하였다. 그림 5의 데이터셋 기반 평가에서 동일 인물 쌍의 평균 유사도는 0.8069, 서로 다른 인물 쌍의 평균 유사도는 0.4992로 나타나, 두 분포 간에 약 0.31의 분리 간격이 확보되었다. 이는 임계값 0.75를 기준으로 동일 인물과 타인을 안정적으로 구분할 수 있음을 의미한다



그림 5. 데이터셋 기반 평가 결과

Fig. 5. Dataset-based evaluation results

실제 환경 비교군 평가에서는 의복 색상이 판별에 미치는 영향을 확인하기 위해, 동일 인물/같은 옷, 동일 인물/다른 옷, 다른 인물/같은 옷, 다른 인물/다른 옷의 네 가지 경우를 테스트하였다. 그림 6

은 각 경우의 유사도와 판별 결과를 나타낸다. 동일 인물은 옷 색상이 같을 때 임계값을 초과하는 유사도로 MASTER로 판별되었으며, 나머지 경우는 임계값 이하로 구분되어 UNKNOWN으로 분류된다.

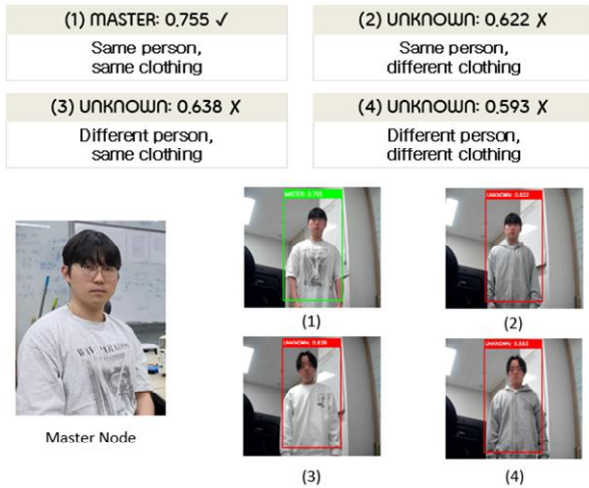


그림 6. 실환경 비교군 평가 결과

Fig. 6. Real-world control group evaluation results

4.3 실제 추종 주행 실험

제안 시스템의 통합 동작을 검증하기 위해 실제 환경에서 추종 주행 실험을 수행하였다. 로봇은 등록된 사용자만 MASTER로 인식하여 약 1.5m의 거리를 유지하며 추종하고 프레임 내에 2~3명의 인물이 동시에 존재하는 실환경에서도 로봇은 등록된 사용자만을 MASTER로 판별하여 추종하였으며, 미등록 인물에 대해서는 추종을 수행하지 않았다. 그림 7은 실제 주행 실험 장면을 나타낸다.



그림 7. 실제 주행 실험 장면

Fig. 7. Actual driving experiment scene

인지 파이프라인은 Jetson Xavier NX에서 약 9.8 FPS로 동작하여 실시간 추종에 필요한 처리 속도를 확보하였으며, 모바일 앱의 시작/정지 명령이 MQTT를 통해 로봇에 전달되어 주행이 즉각적으로 제어됨을 확인하였다. 또한 실험 전 과정에서 사용자는 앱의 실시간 카메라 화면을 통해 로봇의 시야와 추종 상태를 원격으로 확인할 수 있으며, 등록·주행·모니터링·개입으로 이어지는 전체 운용 흐름이 모바일 인터페이스 하나로 수행됨을 검증하였다.

4.4 시스템 한계 및 원인 분석

실험 과정에서 다음과 같은 한계가 확인되었다. 첫째, 추종 대상과의 거리가 유효 범위를 벗어나는 경우 대상을 상실하는 문제가 있다. 대상이 지나치게 멀어지면 영상 내 크기가 작아져 검출과 재식별의 신뢰도가 떨어지고, 반대로 지나치게 가까워지면 대상의 신체 일부만 프레임에 포착되어 외형 기술자 추출이 불안정해진다. 현재 검출 단계에서는 대상의 크기 변화를 보정하는 별도의 알고리즘을 적용하지 않았으며, 다양한 크기의 특징을 함께 학습하는 다중 스케일 융합이라는 OSNet의 구조적 특성을 통해 크기 변화에 부분적으로 대응하고 있다. 다만 원거리에서 대상이 작게 관측되는 경우의 신뢰도 저하는 구조적 특성만으로는 완전히 해소되지 않으므로, 향후 얼굴 특징 등 추가적인 식별 단서의 결합을 통한 보완이 필요하다. 둘째, 현재 시스템은 카메라 영상 기반의 대상 인식에 집중되어 있어 주행 경로 상의 장애물을 감지하고 회피하는 기능이 미흡하다. 장착된 LiDAR를 인지 파이프라인과 융합하지 않아, 추종 중 예기치 않은 장애물에 대한 대응은 사용자의 수동 개입에 의존한다. 셋째, 구동부가 일반형 바퀴로 구성되어 경사로나 요철이 많은 노면에서는 안정적인 주행이 어려워, 현재 시스템의 운용 범위는 평탄한 환경으로 제한된다.

V. 결론 및 향후 과제

본 논문은 모바일 앱 기반의 대상 등록, Re-ID 기반 사용자 식별, 추종 주행, 즉각적인 사용자 제어를 하나의 임베디드 시스템으로 통합한 사용자

맞춤형 자율주행 로봇 시스템을 제안하였다. 제안 시스템은 사용자가 모바일 앱에서 다섯 방향의 전신 사진을 등록하면 OSNet 기반 Re-ID 모델이 대표 특징 벡터를 생성하고, YOLOv11 기반 검출, UCMCTrack 기반 추적, NAV2 기반 주행 제어로 이어지는 인지 파이프라인을 통해 등록된 사용자만을 선택적으로 추종하도록 구성되었다.

실험에서는 데이터셋 기반 평가에서 동일 인물 쌍의 평균 유사도 0.8069, 타인 쌍의 평균 유사도 0.4992로 약 0.31의 분리 간격이 확인되었으며, 실환경 주행 실험에서는 로봇이 등록된 사용자만을 약 1.5m의 거리를 유지하며 추종하고 미등록 인물은 추종하지 않음을 확인하였다. 인지 파이프라인은 Jetson Xavier NX 임베디드 환경에서 약 9.8 FPS로 동작하여 외부 서버 없이 온디바이스 실시간 추종이 가능함을 보였다.

다만 현재 시스템은 대상과의 거리가 유효 범위를 벗어나면 대상을 상실할 수 있고, 장애물 회피가 사용자의 수동 개입에 의존하며, 일반형 바퀴 구동부로 인해 운용 범위가 평탄한 환경으로 제한되는 한계를 가진다. 이에 따라 향후 연구에서는 거리 적응형 추종 로직의 적용, LiDAR와 카메라의 융합을 통한 장애물 감지·회피 기능의 구현, 험지 주행이 가능한 구동계로의 확장을 수행하고, 아울러 보행 패턴 등 색상 외 특징의 결합과 명령 지연시간 등 정량 지표의 확대 측정을 통해 다양한 환경에서 시스템의 강인성을 검증하고자 한다.

이 연구는 사용자 개인화 인식과 모바일 제어를 임베디드 추종 로봇에 통합한 시도로서, 별도의 마커나 전용 장비 없이 스마트폰만으로 로봇을 등록하고 제어할 수 있는 사용자 맞춤형 로봇 서비스의 기반을 제공하였다. 향후 물류 현장의 운반 보조, 공항·병원 등 공공시설에서의 짐 운반, 고령자 및 교통약자의 이동 지원과 같은 실제 서비스 분야에 활용되고, 특정 사용자를 선별적으로 보조하는 서비스 로봇의 접근성 향상에 크게 기여할 것으로 예상된다.

References

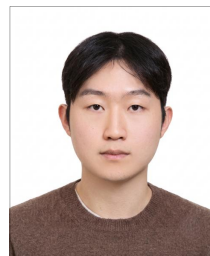
- [1] International Federation of Robotics (IFR), "WorldRobotics 2025 report - INDUSTRIAL ROBOTS - released by IFR", <https://ifr.org/ifr-press-releases/news/global-robot-demand-in-factories-double-s-over-10-years>. [accessed: May 09, 2026]
- [2] M. J. Islam, J. Hong, and J. Sattar, "Person-following by Autonomous Robots: A Categorical Overview", *The International Journal of Robotics Research*, Vol. 38, No. 14, pp. 1581-1618, Dec. 2019. <https://doi.org/10.1177/0278364919881683>.
- [3] A. Eirale, M. Martini, and M. Chiaberge, "Human Following and Guidance by Autonomous Mobile Robots: A Comprehensive Review", *IEEE Access*, Vol. 13, pp. 42214-42253, Mar. 2025. <https://doi.org/10.1109/ACCESS.2025.3548134>.
- [4] H. Ye, Y. Zhan, W. Situ, G. Chen, J. Yu, Z. Zhao, K. Cai, A. Ajoudani, and H. Zhang, "TPT-Bench: A Large-Scale, Long-Term and Robot-Egocentric Dataset for Benchmarking Target Person Tracking", *The International Journal of Robotics Research*, pp. 1-16, Jun. 2026. <https://doi.org/10.1177/02783649261447308>.
- [5] H. Ye, J. Zhao, Y. Pan, W. Chen, L. He, and H. Zhang, "Robot Person Following Under Partial Occlusion", *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, London, UK, pp. 7591-7597, May 2023. <https://doi.org/10.1109/ICRA48891.2023.10160738>.
- [6] T. W. Seo, H. J. Park, H. B. Lim, M. C. Kim, C. M. Kang, and H. S. Jung, "Utilizing autonomous robots for developina cleaning automation system in large public facilities", *Proc. Autumn Conference of the Korean Society of Manufacturing Technology Engineers (KSMTE)*, Jeju, Korea, p. 335, Dec. 2024.
- [7] T. Yoshimi, M. Nishiyama, T. Sonoura, H. Nakamoto, S. Tokura, H. Sato, F. Ozaki, N. Matsuhira, and H. Mizoguchi, "Development of a Person Following Robot with Vision Based Target Detection", *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, pp. 5286-5291, Oct. 2006.

- <https://doi.org/10.1109/IROS.2006.282029>.
- [8] H. Ye, J. Zhao, Y. Zhan, W. Chen, L. He, and H. Zhang, "Person Re-Identification for Robot Person Following with Online Continual Learning", IEEE Robotics and Automation Letters, Vol. 9, No. 11, pp. 9151-9158, Nov. 2024. <https://doi.org/10.1109/LRA.2024.3440087>.
- [9] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang, "Omni-Scale Feature Learning for Person Re-Identification", Proc. IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, pp. 3702-3712, Oct. 2019.
- [10] S. S. H. Hajjaj and K. S. M. Sahari, "Developing a Portable Human-Robot Interaction (HRI) Framework for Outdoor Robots Through Selective Compartmentalization", Arabian Journal for Science and Engineering, Vol. 44, No. 11, pp. 9779-9786, Nov. 2019. <https://doi.org/10.1007/s13369-019-04009-z>.
- [11] L. Zhang, P. Shi, S. Wang, and X. Su, "A Survey of Safety Control for Service Robots", Journal of Automation and Intelligence, Vol. 5, No. 2, pp. 112-125, Jun. 2026. <https://doi.org/10.1016/j.jai.2025.10.002>.
- [12] G. Jocher and J. Qiu, "Ultralytics YOLO11", Version 11.0.0, 2024. <https://github.com/ultralytics/ultralytics>. [accessed: Jul. 17, 2026]
- [13] K. Yi, K. Luo, X. Luo, J. Huang, H. Wu, R. Hu, and W. Hao, "UCMCTrack: Multi-Object Tracking with Uniform Camera Motion Compensation", Proc. AAAI Conference on Artificial Intelligence, Vancouver, Canada, Vol. 38, No. 7, pp. 6702-6710, Mar. 2024. <https://doi.org/10.1609/aaai.v38i7.28493>.
- [14] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable Person Re-identification: A Benchmark", Proc. IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, pp. 1116-1124, Dec. 2015. <https://doi.org/10.1109/ICCV.2015.133>.

- [15] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, pp. 815-823, Jun. 2015. <https://doi.org/10.1109/CVPR.2015.7298682>.

저자소개

고 윤 별 (Unbyeol Go)



2022년 3월 ~ 현재 :

국립군산대학교 소프트웨어학과
학사과정

관심분야 : 데이터베이스,
자율주행, 퍼지컬인공지능

김 건 웅 (Geonwoong Kim)



2021년 3월 ~ 현재 :

국립군산대학교 소프트웨어학과
학사과정

관심분야 : 데이터베이스,
자율주행, 퍼지컬인공지능

정 현 준 (Hyunjun Jung)



2008년 3월 : 삼육대학교

컴퓨터과학과(학사)

2010년 3월 : 숭실대학교

컴퓨터학과(공학석사)

2017년 9월 : 고려대학교

컴퓨터·전파통신공학과(공학박사)

2017년 8월 ~ 2020년 8월 :

광주과학기술원 블록체인인터넷경제연구센터 연구원
2021년 3월 ~ 현재 : 국립군산대학교 소프트웨어학과 교수
관심분야 : 블록체인, 데이터 사이언스, 센서 네트워크,
사물인터넷, 머신러닝

정 동 원 (Dongwon Jeong)



1997년 2월 : 군산대학교
컴퓨터과학과(학사)

1999년 2월 : 충북대학교
전자계산학과(석사)

2004년 2월 : 고려대학교
컴퓨터학과(박사)

2005년 4월 ~ 현재 :

국립군산대학교 소프트웨어학과 교수

관심분야 : 데이터베이스, 시맨틱 서비스, 빅데이터,
사물인터넷, 퍼지컬인공지능, 엡지컴퓨팅, 지능형 융합
서비스