

# ROS2 기반 로봇의 딥러닝 기반 사람 인식 및 위치 중앙 정렬 방법

박재원, 유경민, 남승우, 백의준, 김명섭

고려대학교

{2018270614, rudals2710, nam131119, pb1069, tmskim}@korea.ac.kr

## A Deep Learning-based Human Detection and Position Centering Method for ROS2-based Robots

Jae-Won Park, Gyeong-Min Yu, Seung-Woo Nam, Ui-Jun Baek, Myung-Sup Kim\*

Korea University

### 요약

ROS 2(Robot Operating System 2)는 로봇 어플리케이션을 개발하기 위한 오픈 소스 프레임워크이며 이에 기반한 로봇 시스템은 실생활 및 많은 산업 현장에서 다양한 형태로 활용되고 있다. 특히, 최근에는 보안 및 안전 문제, 비용 절감 및 효율성 문제를 해결하기 위하여 자율 순찰 로봇에 관한 연구 개발이 활발히 이루어지고 있다. 그러나, 최근 개발되고 있는 자율 순찰 로봇의 기능은 주변 환경 데이터 수집 및 수동적 대응으로 한정되어 있으며 특히 실내 순찰과 같은 현장에서는 거수자가 침입할 시 이를 탐지하는 것뿐만 아니라 능동적으로 따라가며 지속해서 거수자에 대한 정보를 수집해야 할 필요가 있다. 본 논문에서는 능동적으로 거수자와 로봇과의 거리를 파악하고 일정 거리를 유지하며 추적하는 딥러닝 기반의 사람 인식 및 위치 중앙 정렬 방법을 제안한다.

### I. 서론

ROS 2(Robot Operating System 2)는 로봇 어플리케이션을 개발하기 위한 오픈소스 프레임워크로 이에 기반한 로봇 시스템은 실생활 및 많은 산업 현장에서 다양한 형태로 활용되고 있다. 특히, 이를 기반으로 한 보안 강화 및 비용 절감을 위한 순찰 로봇에 관한 연구 및 개발이 이루어지고 있다 [1]. 자율 주행 기반 순찰 로봇은 중요하며 필요성은 다음과 같다:

- 보안 강화: 공공장소와 기업의 건물 등 순찰을 통해 이상 행동 및 위험 요소를 탐지하며 24시간 정확한 센싱과 분석 능력을 통해 보안 위협을 사전에 탐지하고 예방한다.
- 비용 절감: 자율주행 로봇을 통한 순찰은 인적 자원을 대체하며 작업 비용을 절감하고 인간보다 더 효율적이고 정확한 결과 제공한다.
- 실시간 모니터링: 실시간으로 주변 환경을 모니터링하고 자료를 수집하며 사건 발생 후 빠른 대응을 가능하게 한다.

그러나, 최근 개발된 로봇의 경우 주변 환경 자료를 수집하거나 거수자 접근이 허용되지 않은 구역에서의 거수자 탐지 시 수동적으로 대응하는 한정된 기능을 가진다. 정확한 거수자 탐지 및 모니터링을 위해서는 거수자를 탐지하는 것뿐만 아니라 능동적으로 따라가며 지속해서 거수자에 대한 정보를 수집해야 할 필요가 있으며 이러한 추적 기능 개발을 위해 거수자와의 일정한 거리를 유지하며 따라가는 요소 기술은 필수적이다.

본 논문에서는 딥러닝 기반 거수자 탐지 결과를 바탕으로 거수자와 로봇과의 일정한 거리를 유지하는 방법을 제안한다. 본 논문은 서론에 이어 관련된 기술들을 설명하고 본론에서 제안하는 방법을 설명한다. 마지막으로, 결론과 향후 연구에 대하여 설명하고 본 논문을 마친다.

### II. 관련 연구

ROS 2는 자율 주행 로봇 및 로봇 응용 프로그램 개발에 많이 사용되는 오픈 소스 프레임워크로 실시간 처리, 보안, 분산 시스템 등의 기능을 강화하여 로봇 응용 프로그램을 효과적으로 개발할 수 있도록 지원한다 [3].

YOLO(You Only Look Once)는 실시간 객체 탐지를 위한 컴퓨터 비전 알고리즘이며 타 객체 탐지 알고리즘과 달리, YOLO는 이미지를 한번만 보고 객체의 위치와 클래스를 동시에 예측하는 특징을 가지며 본 논문에서는 YOLO를 개선한 v7 버전을 사용하였다 [2].

### III. 본론

본 장에서는 로봇의 사람 인식 및 위치 중앙 정렬을 위한 학습 구조, 추론 방법에 대하여 설명한다.

로봇의 사람 인식 및 위치 중앙 정렬을 위한 학습 개요는 그림 1과 같다.

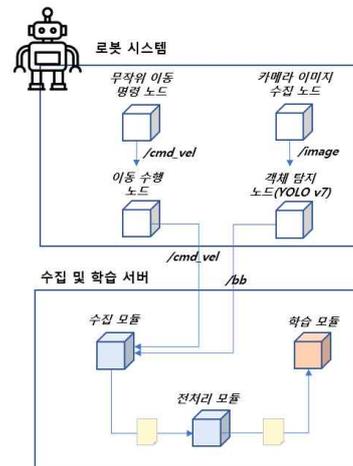


그림 1. 학습 개요

본 과제(결과물)는 2020년도 산업통상자원부 및 한국산업기술평가관리원(KEIT) 연구비 지원에 의한 연구(No. 20008902, IT비용 최소화를 위한 5채널 탐지기술 기반 SaaS SW Management Platform(SMP) 개발)이며 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과 (2021RIS-004)입니다.

로봇 시스템의 무작위 이동 명령 노드는 이동에 대한 명령을 생성하는 노드로 이동 명령의 토픽 이름은 /cmd\_vel 이며 로봇의 전진 또는 후진에 대한 값 linear.x과 angular.z 를 생성한다. linear.x, angular.z 값은 이동 수행 노드로 전달되어 로봇은 해당 값 수치만큼 전·후진 좌우 회전을 수행함과 동시에 /cmd\_vel 이라는 토픽 이름으로 다시 토픽을 발행한다. 카메라 이미지 수집 노드는 로봇의 카메라 센서로부터 이미지를 수집하여 /image라는 토픽을 발행한다. 객체 탐지 모델(YOLO v7)이 탑재된 객체 탐지 노드는 구독한 이미지를 분석하여 사진 내 사람의 윤곽에 대하여 바운딩 박스를 생성하고 이를 /bb 라는 토픽으로 발행하며 /bb 토픽은 바운딩 박스의 중심점 좌표(x,y)와 넓이(w), 높이(h)로 구성된다.

수집/학습 서버의 수집 모듈은 바운딩 박스 토픽 /bb 와 로봇 이동에 관한 토픽 /cmd\_vel을 1초 주기로 구독하여 그림 2와 같은 형식으로 파일 시스템에 저장하고 전처리 모듈은 해당 파일의 데이터를 재구조화 하여 그림 3과 같은 형식으로 변환한다.

time	linear.x	angular.z	bb.x	bb.y	bb.w	bb.h
1	$x_1$	$z_1$	$bx_1$	$by_1$	$bw_1$	$bh_1$
2	$x_2$	$z_2$	$bx_2$	$by_2$	$bw_2$	$bh_2$
3	$x_3$	$z_3$	$bx_3$	$by_3$	$bw_3$	$bh_3$
...	...	...	...	...	...	...
n	$x_n$	$z_n$	$bx_n$	$by_n$	$bw_n$	$bh_n$

그림 2. 수집된 원본 데이터

time	현재 바운딩 박스				다음 바운딩 박스				이동 명령	
	bb.x	bb.y	bb.w	bb.h	bb.x	bb.y	bb.w	bb.h	linear.x	angular.z
2	$bx_1$	$by_1$	$bw_1$	$bh_1$	$bx_2$	$by_2$	$bw_2$	$bh_2$	$x_1$	$z_1$
3	$bx_2$	$by_2$	$bw_2$	$bh_2$	$bx_3$	$by_3$	$bw_3$	$bh_3$	$x_2$	$z_2$
...	...	...	...	...	...	...	...	...	...	...
n-1	$bx_{n-1}$	$by_{n-1}$	$bw_{n-1}$	$bh_{n-1}$	$bx_n$	$by_n$	$bw_n$	$bh_n$	$x_{n-1}$	$z_{n-1}$

그림 3. 전처리한 데이터

전처리를 마친 데이터의 첫 행은  $t_1$  시점의 바운딩 박스 값과  $t_2$  시점의 바운딩 박스 값, 그리고 이에 대한  $t_1$  시점의 이동 명령으로 구성되어 있다. 이는 현재 바운딩 박스에서 다음 바운딩 박스 형태로 변경되려면 같은 행의 이동 명령을 발행해야한다는 의미를 가진다. 이렇게 전처리된 데이터는 3개 레이어로 구성된 DNN(Deep Neural Network) 모델에 입력되며 모델 구조는 표 1과 같다.

레이어	unit 개수
Input	8
FC	16
FC	16
FC	16
Output(FC)	2

표 1. 위치 중앙 정렬 학습 모델 구조

학습 모델은 현재 바운딩 박스 값(4)과 다음 바운딩 박스 값(4) 총 8개 값을 입력 받아 이동 명령 값 2개를 출력한다.

위치 중앙 정렬 모델이 모두 학습된 뒤 실제 추론하는 방법에 대한 개요는 그림 4와 같다. 추론 모델은 로봇으로부터 사람 탐지 결과인 바운딩 박스 값(xywh)과 사용자가 사전에 설정한 원하는 바운딩 박스 값(xywh)을 입력 받는다. 추론 모델은 두 개의 바운딩 박스 값을 입력 받아 현재 바운딩 박스로부터 다음 바운딩 박스 즉 사용자가 원하는 크기의 바운딩 박스 값으로 변환할 수 있는 이동 명령을 출력한다. 생성된 이동 명령은 로봇의 이동을 제어하는 이동 수행 노드로 입력되며 로봇은 해당 명령에 따라 이동하고 이를 통해 탐지한 사람과의 방향과 일정한

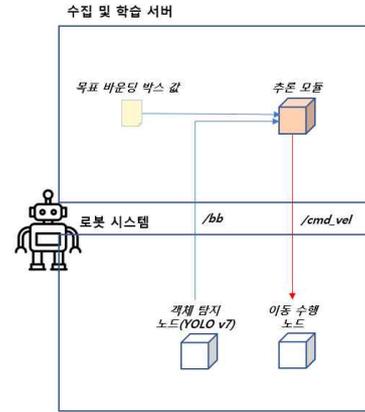


그림 4. 위치 중앙 정렬 추론 방법

거리를 유지하며 사람을 추적한다.

### III. 결론

본 논문에서는 능동적으로 거수자와 로봇과의 거리를 파악하고 일정 거리를 유지하며 추적하는 딥러닝 기반의 사람 인식 및 위치 중앙 정렬 방법을 제안하였다. 제안하는 방법은 YOLO v7를 기반으로 사람을 탐지하고 탐지 결과를 활용하여 탐지된 거수자와의 일정한 거리를 유지하는 가벼운 딥러닝 모델을 생성하였다. 실제 구현 결과, 추론 모델에서 생성된 이동 명령에 따라 로봇은 자연스럽게 이동하였다. 제안하는 방법은 이후 순찰 로봇 및 안내 로봇에 범용적으로 사용될 수 있다. 그러나, 로봇의 하드웨어 사양으로 인한 YOLO v7 모델의 탐지 속도가 느려 병목 현상이 발생하는 한계점과 거리가 너무 멀어지면 탐지 자체가 안되는 한계점이 존재한다. 따라서, 우리는 향후 연구로 로봇에서도 동작할 수 가벼운 탐지 모델의 적용 또는 높은 FPS를 위한 고성능의 탐지 서버와의 연동으로 설정하고 통해 위에서 제시된 한계점을 극복할 예정이다. 또한, 제안한 방법과 바운딩 박스값에 따른 임의의 전·후진 및 좌우 회전 값을 생성하는 휴리스틱한 기존 방법과의 비교 실험을 진행할 예정이다.

### 참고 문헌

[1] BASILICO, Nicola. Recent Trends in Robotic Patrolling. Current Robotics Reports, 2022, 3:2: 65-76.

[2] WANG, Chien-Yao; BOCHKOVSKIY, Alexey; LIAO, Hong-Yuan Mark. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696, 2022.

[3] Open robotics (Ed.). (2023, May 17). ROS 2 Documentation: Foxy. ROS 2 Documentation. <https://docs.ros.org/en/foxy/index.html>