

시각-언어 모델(VLM) 및 영상 기반 상호작용 계획(VIP)을 활용한 로봇 시뮬레이션 지능 시스템 개발

오픈소스 URL :

https://github.com/jaehyeonlim99/my_robot_lim_last.git



2025학년도 2학기 SW 캡스톤디자인 경진대회

팀 명 아자아자

지도교수 김윤경

팀 원

산업체

이호선(it지능정보공학과, 4학년), 김새얀(it지능정보공학과, 4학년),
임재현(화학과, 4학년), 최현수(it정보공학과, 4학년)
텔로스

개발 동기 및 목적

[개발 동기]

- 고창 고인돌 문화재 훼손 증가
- "그 많던 고인돌은 어디로?" 뉴스 보도
- 소중한 문화유산 보호 필요성 대두
- 야외 CCTV의 한계
- 사각지대 존재
- 사후 대응 (실시간 감시 불가)
- cctv 감시 인력 부담

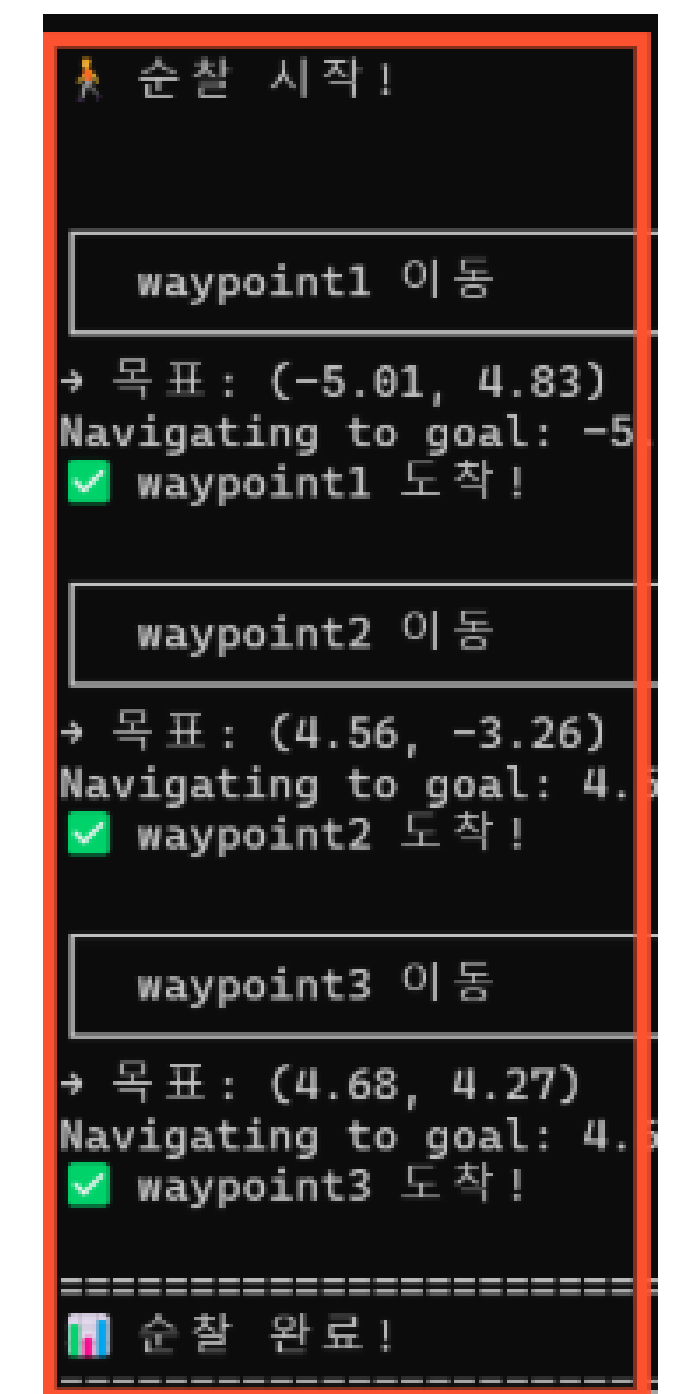
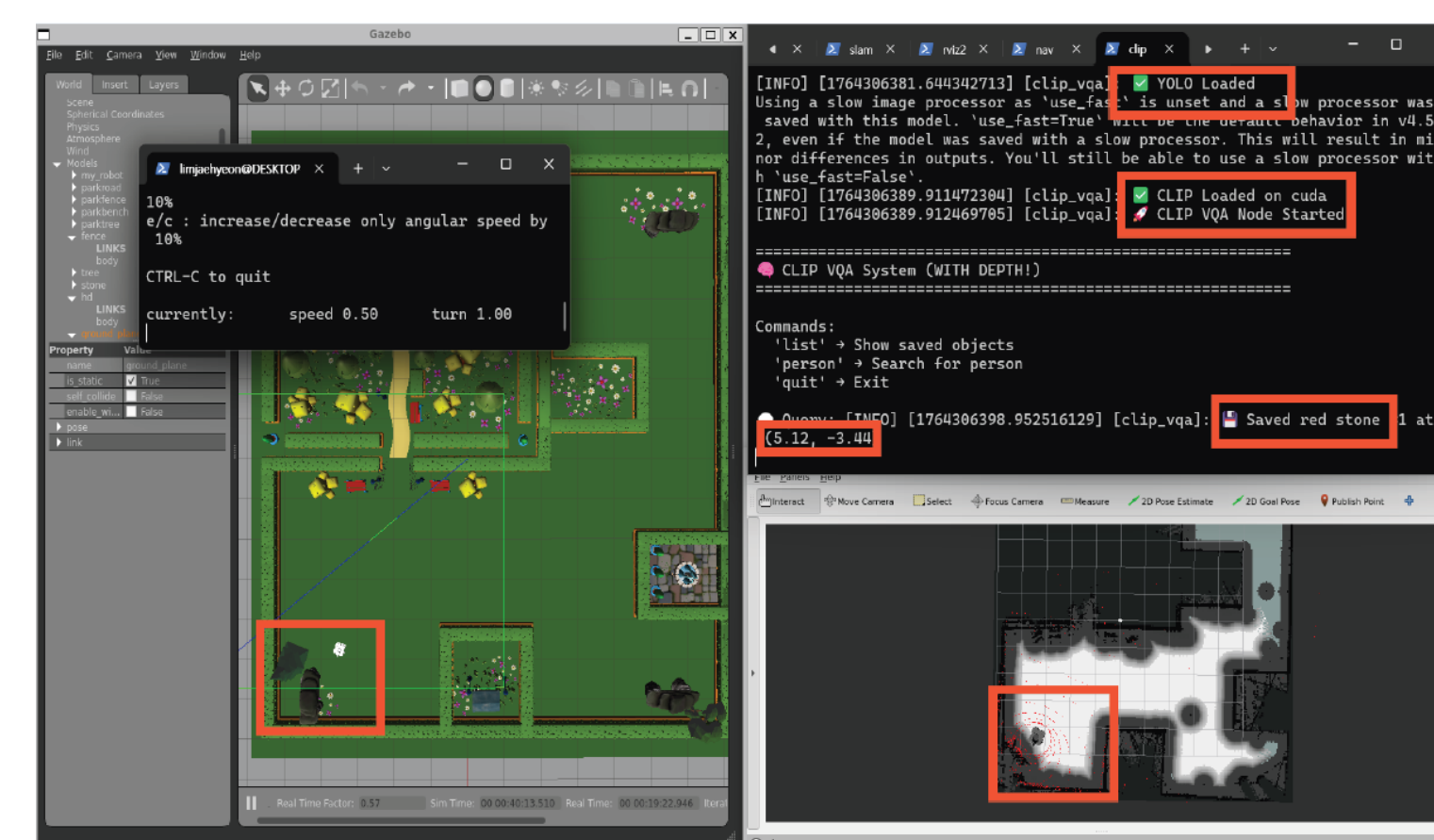
[개발 목적]

- VLM 기반 지능형 순찰 로봇 개발
- 시간 제약 없이 전 구역 순찰(로봇 이동 기반)
- VLM 기반 객체 검색 후 자율 이동
- Waypoint 기반 순찰



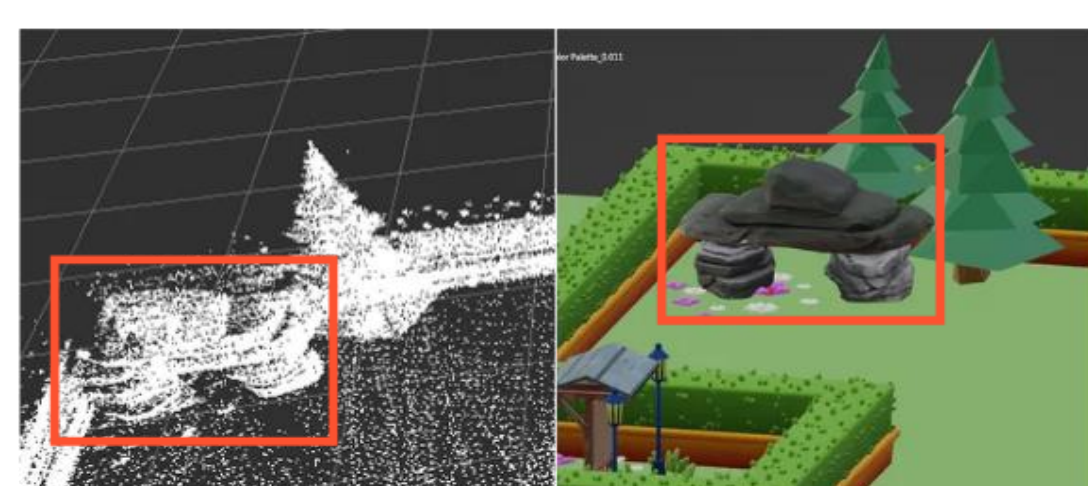
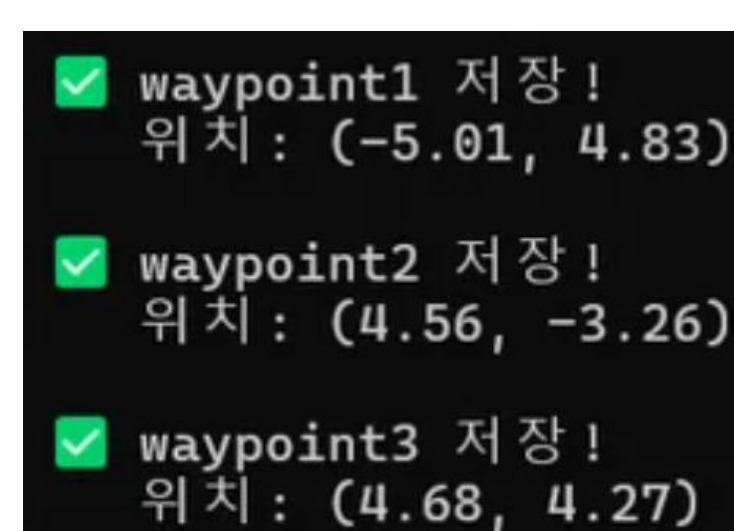
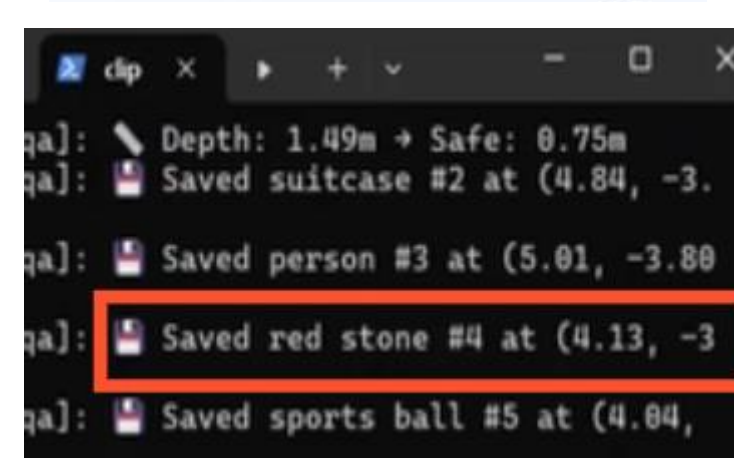
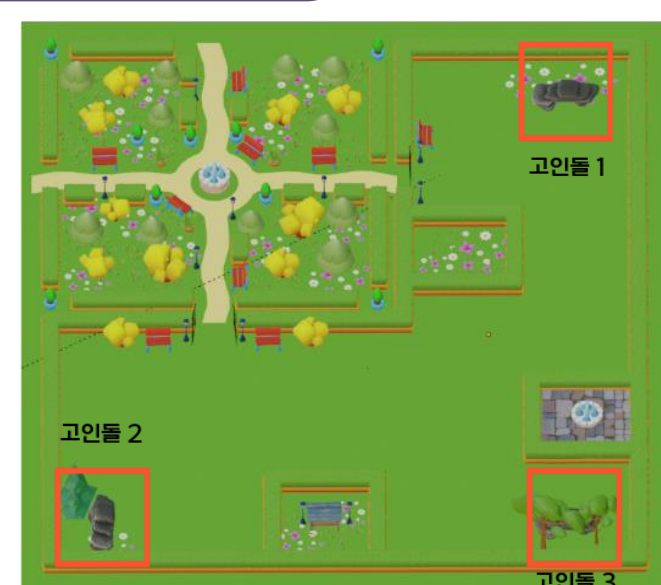
주요 기술

- VLM_VQA 텍스트 명령으로 객체 찾기
- "stone" 입력
→ NAVIGATION 통해 고인돌로 이동
→ 안전한 장애물 회피 + 최단 거리 주행
- YOLO + CLIP VQA 통합
- Waypoint 기반 자동 순찰
- 중요 위치 저장 후 자동 순회
- 야간 무인 순찰/감시
- 센서 융합 로봇 설계
- RGB + Depth 카메라
- 2D + 3D LiDAR
- SLAM 실시간 맵 생성
- 실제 로봇 배포 고려
- Jetson 기반 경량 모델 선택
- ROS2 모듈형 구조
- Telos 로봇사와 협력



개발 내용

- Gazebo + ROS2 시뮬레이션 환경 구축
- 고창 고인돌 공원 3D 맵 제작
- URDF 로봇 모델링 (센서 통합)
- SLAM Toolbox 활용
- 실시간 맵 생성
- 수동 로봇 조종으로 초기 탐색
- Waypoint 저장 시스템
- YOLO + CLIP VQA 통합
- YOLOv8n: 객체 탐지
- CLIP: 텍스트-이미지 유사도 검색
- Depth 카메라로 정확한 거리 측정
- /goal_pose 자동 발행
- Navigation 파라미터 최적화
- CLIP VQA가 발행한 목표 위치 수신
- 위험 지역 : 0.30m
- 장애물 회피 민감도 : 1.5
→ 20회 테스트로 파라미터 최적값 발견
- WAY POINT
- (X, Y, Z)좌표 WAYPOINTS.JSON에 저장
- 순차적 자동 순회
→ 시간 제약 없이 자동 감시 가능
- 3D LiDAR 포인트 클라우드 시각화
- Velodyne 센서 데이터
- 로봇이 보는 세상 직관적으로 표현



결과 및 분석

- 개발 목표 달성
- 시간 제약 없이 전 구역 순찰(로봇 이동 기반)
- VLM 기반 객체 검색 후 자율 이동
- Waypoint 기반 순찰
- Navigation 성공률: 95%
- 다양한 장애물 설치 후 40회 테스트 → 38회 성공
- 실패 원인: 위험 지역으로 인한 좁은 통로
- 확장성 검증
- WAY POINT. JSON 수정 → 다른 경로 순찰 적용 가능
- 다른 유적지 환경에도 적용 가능
- Jetson 보드 배포 준비 완료
- 향후 계획
- Unitree Go2 실제 로봇 연동
- 활용 방안
- 문화재·관광지 야외 감시/순찰 자동화
- 놀이터·공원 내 어린이 안전 모니터링 + 위험요소 접근 차단
- 지자체·공공시설 스마트 순찰 운영 모델(관제 연동형)

