creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

博士學位 請求論文
指導敎授 朴 承 熙

건설현장 위험상황 이해 및 대응을 위한 인공지능 안전인지 프레임워크 연구

成均館大學校 一般大學院
글로벌스마트시티融合專攻
災難管理專攻
金 兌 憲

SUNG KYUN KWAN UNIVERSITY

博士學位請求論文

건설현장 위험상황 이해 및 대응을 위한 인공지능 안전인지 프레임워크 연구

2026

金兌憲

SUNG KYUN KWAN UNIVERSITY

博士學位 請求論文
指導敎授 朴 承 熙

건설현장 위험상황 이해 및 대응을 위한 인공지능 안전인지 프레임워크 연구

A Study on a Artificial Intelligence Framework for Safety Perception in Understanding and Responding to Risk Situations in Construction Sites

成均館大學校 一般大學院
글로벌스마트시티融合專攻
災難管理專攻
金 兌 憲

SUNG KYUN KWAN UNIVERSITY

博士學位 請求論文
指導敎授 朴 承 熙

건설현장 위험상황 이해 및 대응을 위한 인공지능 안전인지 프레임워크 연구

A Study on a Artificial Intelligence Framework for Safety Perception in Understanding and Responding to Risk Situations in Construction Sites

이 論文을 工學 博士學位請求論文으로 提出합니다.

2025 年 10 月 日

成均館大學校 一般大學院
글로벌스마트시티融合專攻
災難管理專攻
金 兌 憲

SUNG KYUN KWAN UNIVERSITY

이 論文을 金 兌 憲의 工學

博士學位 論文으로 認定함.

2025 年 12 月 日

審査委員長

審査委員

審査委員

審査委員

審査委員

SUNG KYUN KWAN UNIVERSITY

목 차

제1장 서 론 ····· 1
1. 연구배경 ····· 1
2. 연구의 목적 ····· 2
3. 연구 방법 및 범위 ····· 4

제2장 이론적 배경 및 적용기술 ····· 5
1. 관련 연구 동향 ····· 5
1.1 센서 및 컴퓨터 비전 기반의 객체 탐지 기술 ····· 5
1.2 딥러닝 기반 객체 탐지 모델의 고도화 및 현장 적용 ····· 6
1.3 개인보호구(PPE) 착용 감지 및 중장비 협착 방지 기술 ····· 7
1.4 행동 인식 및 인체공학적 위험 평가 ····· 8
1.5 VLM을 이용한 시각적 문맥 이해 ····· 10
1.6 온프레미스 환경을 위한 sLLM 및 경량화 기술 ····· 11
1.7 RAG 기반의 할루시네이션 억제 및 법규 준수 ····· 12
2. 이론적 배경 ····· 15
2.1 건설 안전 관리 이론 ····· 15
2.2 인공지능 기초 이론 ····· 16
2.3 실시간 객체 탐지 및 추적 기술 ····· 16

SUNG KYUN KWAN UNIVERSITY

2.4 시각적 문맥 이해와 관계 추론 ····· 18
2.5 시각언어 모델(VLM) 및 최적화 기술 ····· 18
2.6 검색 증강 생성 (RAG) 및 벡터 데이터베이스 ····· 20
2.7 엣지 AI 시스템 아키텍처 및 서빙 최적화 ····· 21
3. 소결 ····· 22

제3장 실험 및 성능 평가 ····· 23
1. 실험 개요 및 목표 ····· 23
2. 실험 1: 도메인 적합성 평가 및 최적 모델 선정 ····· 25
2.1 배경 및 문제 정의 ····· 25
2.2 실험 목적 ····· 25
2.3 실험 설계 및 방법 ····· 25
2.4 실험 결과 요약 ····· 27
2.5 상세 분석 및 토의 ····· 54
3. 실험 2: 논리 추론용 sLLM 성능 평가 및 최적 엔진 선정 ····· 56
3.1 배경 및 문제 정의 ····· 56
3.2 실험 목적 ····· 56
3.3 실험 설계 및 방법 ····· 56
3.4 실험 결과 요약 ····· 57

SUNG KYUN KWAN UNIVERSITY

3.5 상세 분석 및 토의 ····· 59
4. 실험 3: 운용 경제성 분석 ····· 61
4.1 배경 및 문제 정의 ····· 61
4.2 실험 목적 ····· 61
4.3 실험 모델 설계 ····· 61
4.4 실험 결과 및 토의 ····· 62
5. 소결 ····· 63

제4장 인공지능 안전인지 프레임워크 구현 ····· 64
1. 시스템 아키텍처 및 구현 환경 ····· 64
1.1 아키텍처 설계 원칙 및 데이터 흐름 ····· 64
1.2 하드웨어 및 소프트웨어 구현 환경 ····· 66
2. 실시간 이벤트 감지 및 데이터 큐레이션 파이프라인[주간] · 68
2.1 저지연 비디오 획득 및 적응형 연결 제어 ····· 68
2.2 순환 버퍼링을 통한 문맥 데이터(Context Data) 확보 ····· 68
3. 온프레미스 멀티모달 심층 분석 파이프라인[야간] ····· 69
3.1 분산 작업 스케줄링 및 자원 할당 최적화 ····· 69
3.2 로컬 LLM 서빙 프레임워크 및 양자화 ····· 69
3.3 다단계 심층 분석 파이프라인 ····· 69

SUNG KYUN KWAN UNIVERSITY

4. 시스템 구현 상세 ····· 71
4.1 GStreamer 기반 고속 파이프라인 구현 ····· 71
4.2 Redis-Celery 기반 비동기 작업 오케스트레이션 ····· 71
4.3 오프라인 분석 엔진 구현 ····· 72
4.4 RAG 기반 문서 탐지 및 법규 매칭 ····· 72
5. 현장 적용 및 운영 결과 ····· 73
5.1 실증 현장 특성 및 관제 시스템 구축 ····· 73
5.2 온프레미스 프레임워크 심층 분석 적용결과[야간] ····· 76
5.3 정성적 평가 및 논의 ····· 79

제5장 결론 ····· 81

참 고 문 헌 ····· 84

SUNG KYUN KWAN UNIVERSITY

<표 차례>

표 1 건설 안전 모니터링 기술의 패러다임 비교 / 연구 차별성 · 3
표 2 플랫폼 핵심 요소 결정을 위한 설계 근거 및 수행 실험 … 23
표 3 하드웨어 및 소프트웨어 구성 …… 24
표 4 실험1에서 사용한 시스템 프롬프트 …… 27
표 5 실험1(도메인적합성)의 결과 …… 27
표 6 실험1의 화재이미지 결과 …… 28
표 7 실험1의 안전보호구 이미지 추론결과 …… 32
표 8 실험1의 중장비 이미지 추론결과 …… 36
표 9 실험1의 표지판 이미지 추론결과 …… 41
표 10 실험1의 화재 이미지 추론결과 …… 46
표 11 실험1의 추락 이미지 추론결과 …… 50
표 12 실험2(모델별 벤치마크)의 결과 …… 57
표 13 워크스테이션 기반 실시간 CCTV 영상감지 환경 …… 67

SUNG KYUN KWAN UNIVERSITY

<그림 차례>

그림 1 Faster R-CNN 기반 객체탐지 파이프라인 ····· 6
그림 2 실시간 PPE 감지 시스템 ····· 7
그림 3 융합데이터를 이용한 행동인식 절차 ····· 9
그림 4 Construction-GPT ····· 11
그림 5 LLM과 sLLM의 성능 비교 ····· 12
그림 6 BiFrostRAG 수행결과 ····· 13
그림 7 YOLO의 그리드(Grid) 처리 방식 개념도 ····· 17
그림 8 LLaVA 아키텍처 ····· 19
그림 9 RAG 파이프라인 흐름도 ····· 21
그림 10 시나리오별 평가 이미지(예시) ····· 26
그림 11 실험 3의 결과 ····· 62
그림 12 구현된 시스템 아키텍쳐 ····· 65
그림 13 플랫폼내 AI 탐지 태스크 ····· 74
그림 14 실시간 감시 기능 ····· 74
그림 15 저장된 감지 내역 목록 ····· 75
그림 16 감지내역 상세 확인 ····· 75
그림 17 VLM 심층 분석을 통해 발견한 오알람 사례 ····· 77

SUNG KYUN KWAN UNIVERSITY

그림 18 VLM 심층 분석을 통해 발견한 유효 위험 사례 ····· 77
그림 19 VLM 심층 분석을 통해 발견한 위험탐지 실패 사례 ··· 78
그림 20 자동 생성된 종합안전 리포트 ····· 79

SUNG KYUN KWAN UNIVERSITY

논문요약

건설현장 위험상황 이해 및 대응을 위한 인공지능 안전인지 프레임워크 연구

건설 현장에서 발생하는 재해와 사고는 작업자의 행동과 환경 요인이 복합적으로 작용하는 인과적 사건임에도 불구하고, 기존 센서나 컴퓨터 비전 기술은 단순한 객체 위치 추적이나 안전모 미착용 식별 등 현상적 데이터 수집에만 머물러 사고 위험의 심층적 맥락을 해소하지 못 하였고 한계를 극복할 수 있는 대규모 인공지능(VLM, LLM 등)은 특성상 높은 연산 비용과 클라우드 의존성으로 인하여 통신 인프라가 열악하고 데이터 보안이 필수적인 건설 현장에 직접 적용하기에는 현실적인 제약이 뚜렷하므로, 본 연구는 엣지 환경의 제약을 오히려 연산 효율화의 기회로 전환하는 '온프레미스 하이브리드 안전 인지 프레임워크'를 제안한다.

본 연구의 핵심은 연산 리소스의 시공간적 분리로, 즉시 대응이 필수적인 주간에는 경량 객체 탐지 모델로 위험 이벤트를 실시간 포착하고 유휴 자원이 발생하는 야간에 VLM과 sLLM을 가동해 축적된 데이터를 심층 분석하는 이원화된 아키텍처를 구현했다. 아울러 양자화 모델을 적용해 고사양 서버 없이 단일 워크스테이션만으로 운용 가능한 환경을 구축했으며, 검색 증강 생성(RAG)을 이용하여 생성형 AI의 환각 현상을 억제하고 산업안전보건법에 부합하는 안전 리포트를 생성하도록 구현하였고 실제 현장에서 실증하였는데, 객체 탐지 시스템의 고질적인 문제였던 오알람을 93.4%의 정확도로 필터링하여 안전관리자의 알림 피로도를 유의미하게 경감시킬 수 있음을 확인하였다.

본 연구는 최신 생성형 AI 기술을 건설 도메인의 물리적·환경적 제약 조건에 맞춰

SUNG KYUN KWAN UNIVERSITY

최적화한 해법을 제시함으로써, 스마트 건설 안전 기술의 새로운 패러다임을 정립하는 데 기여하고자 한다.

주제어 : 건설안전, 인공지능 , 소형 거대언어모델, 오알람 저감

SUNG KYUN KWAN UNIVERSITY

제1장 서 론

1. 연구배경

건설 현장의 중대재해는 작업 환경의 불안전한 상태뿐만 아니라 작업자의 불안전한 행동에서 기인하는 '휴먼 에러(Human Error)'가 주요 원인으로 지목되어왔고 “CSI자료를 활용한 국내 건설업 사망사고 심층분석(2023)”에 따르면, 자주 발생하는 사고의 주요 원인은 ‘작업자의 단순과실’으로 지목하고 있으며, 관리적 요인이 사고 예방에 결정적인 영향을 미침을 시사한다. 휴먼 에러를 선제적으로 통제하기 위해 다년간 여러 노력이 이뤄지고 있고 특히, 컴퓨터 비전 기반의 인공지능 기술 도입이 가속화되고 있으나, 여전히 사고 발생 시점의 단순 감지 및 경보에만 치중되어 있어 근본적인 개별 작업자의 안전관리 강화를 통한 재해 예방에는 한계가 있다. 즉, 개발 작업자의 안전의식을 고취하거나 재해의 근본 원인을 선제적으로 차단하는 예방적 차원에서는 여전히 공백이 존재한다. 더욱이 건설 현장은 보안상의 이유로 외부 네트워크 접속이 물리적으로 제한되거나, 통신 인프라가 아예 설치되어 있지 않는 현장도 빈번하여 클라우드 기반의 대규모 언어 모델(LLM)이나 시각 언어 모델(VLM)은 광범위한 지식과 상식 그리고 뛰어난 추론 능력 등을 보유하고 있지만 결국은 데이터 보안 문제와 막대한 통신/연산 비용으로 인해 실제 현장에 적용하기에는 현실적인 장벽이 존재한다.

따라서 건설 현장의 엄격한 보안 조건을 달성하는 동시에 열악한 인프라 환경을 준수하면서도 고성능 AI를 활용하여 현장에서 발생 가능한 사고의 징후를 분석하고 나아가 리포트화하여 안전 교육으로 환류시켜 휴먼 에러를 저감할 수 있는 이른바 '현장 중심의 작업자 안전 시스템'의 개발이 절실하게 요구된다.

SUNG KYUN KWAN UNIVERSITY

2. 연구의 목적

본 연구는 건설 현장이 내제하고 있는 특수한 제약 조건인 '통신 인프라의 불안정성', '데이터 보안의 필요성', 그리고 '엣지 컴퓨팅 자원의 한계'라는 삼중고를 공학적 관점에서 분석하고, 이를 극복할 수 있는 온프레미스 안전 인지 프레임워크(On-premise Safety Perception Framework)를 제안하고 실증적으로 도출 및 구현하는 데 목적이 있다. 이것은 단순히 고성능 AI 모델을 건설 현장에 적용하는 것을 넘어, 실무적 측면에서 시스템의 운용가능성을 담보할 수 있는 효율적인 모델 조합과 운용 파라미터를 규명하고, 이를 바탕으로 현장에 즉시 도입 가능한 경제적이고 강건한 시스템을 제시하고자 한다.

구체적인 세부 목표는 다음과 같다.

1) 단순 객체 탐지를 넘어선 '맥락 추론(Contextual Reasoning)'의 실효성 규명을 위하여 기존 객체 탐지(Object Detection) 기술은 "작업자가 있다"는 사실 탐지에 그쳐, "작업자가 왜 위험한가"에 대한 인과관계를 설명하지 못하는 의미적 공백의 한계를 1차 객체 탐지와 2차 시각언어 모델(VLM) 분석을 결합한 순차적 파이프라인을 통해, 미세한 위험 상황(예: 객체 간의 거리에 따른 종합적인 판단)을 인지하고 설명하는 데 있어 기존 단일 모델 대비 기술적 우위를 실증적으로 규명한다.

2) 자원 제약과 보안을 고려한 '온프레미스 운용 최적점' 도출을 수행하여 고성능 AI의 클라우드 의존성이라는 현장 도입의 주요 장벽을 '주간-실시간 탐지'와 '야간-심층 분석'을 분리 운영하는 하이브리드 아키텍처를 설계하고, 비용과 성능의 트레이드오프(Trade-off)를 분석하여 엣지 환경에서 운용 가능한 최적의 공학적 최적점을 도출하여 초기 구축비용 대비 장기적 운용 경제성을 입증한다.

3) 생성형 AI의 환각(Hallucination) 현상을 억제하고 법적 신뢰성을 확보하기 위한 방안으로 검색 증강 생성(RAG, Retrieval Augmented Generation) 기술을 도입하여

SUNG KYUN KWAN UNIVERSITY

산업안전보건법에 근거한 논리적이고 구체적인 안전 리포트를 생성하는 체계를 구현한다.

기술 단계	핵심 기술	주요 기능	한계점	본 연구의 차별점
1세대 (Sensing)	IoT 센서 (RFID, GPS, UWB)	• 위치 추적 • 접근 경보	• 태그 미소지 시 감지 불가 • 상황 맥락 인지 불가	• 비접촉식 • 별도 장비 없이 영상만으로 분석
2세대 (Vision AI)	객체 탐지 (YOLO, R-CNN)	• 객체(사람, 장비) 식별 • 안전모 미착용 감지	• 의미적 공백 • "왜 위험한가"에 대한 추론 부재	• 맥락적 이해 • 객체 간 상호작용 및 위험성 추론
3세대 (Gen AI)	클라우드 LLM (GPT, Gemini)	• 고차원 상황 묘사 • 자연어 질의응답	• 현장 적용성 낮음 (보안 위협, 통신 비용, 지연 시간)	• 온프레미스 • 보안/비용 문제를 해결한 엣지 운용
본 연구 (Proposed)	하이브리드 (Hybrid Framework)	• 객체 식별+심층 추론 • 법적 근거 기반 리포팅	-	• 실무형 기술 구현 • RAG 기반의 신뢰성

표 1 건설 안전 모니터링 기술의 패러다임 비교 및 연구 차별성

SUNG KYUN KWAN UNIVERSITY

3. 연구 방법 및 범위

본 연구는 건설 현장의 복합적인 제약 사항을 해결하기 위하여 '선행 분석에 기반한 최적 설계' 방법론을 채택하였으며 연구의 논리적 절차는 다음의 3단계로 구성하였다.

첫째, 온프레미스 적용을 위한 요소 기술 분석 및 최적화: 시스템 설계에 앞서, AI 모델과 운용 플랫폼에 대한 다각적인 벤치마크 실험을 기초로 건설 도메인의 시각적·언어적 특성을 반영한 AI모델의 적합성 평가와 클라우드 대비 온프레미스 운용의 경제성 분석을 통해 시스템 구축을 위한 최적의 기술적 요구사항을 도출한다.

둘째, 하이브리드 안전인지 프레임워크 설계 및 구현: 선행 분석에서 도출된 최적의 모델 조합(VLM + sLLM)를 반영하여, 실제 현장에서 가동할 수 있는 하이브리드 아키텍처를 설계하는데 주간의 실시간 객체탐지 파이프라인과 야간의 시각언어 모델 기반의 심층 분석 파이프라인으로 구성되며, RAG 모듈을 통합하여 법적 신뢰성을 확보한 최종 시스템으로 구현된다.

셋째, 시스템 구현 결과 및 기능 검증: 구현된 프레임워크를 실제 운영중인 건설 현장에 적용하여 기능적 작동 여부를 확인하고, AI가 추론한 결과를 바탕으로 현재 시스템의 한계점과 앞으로 발전방향에 대하여 토의한다.

SUNG KYUN KWAN UNIVERSITY

제2장 이론적 배경 및 적용기술

1. 관련 연구 동향

1.1 센서 및 컴퓨터 비전 기반의 객체 탐지 기술

일반적인 건설 현장의 안전 관리 체계는 RFID(Radio Frequency Identification), IMU(Inertial Measurement Unit), GPS(Global Positioning System) 등의 부착형 센서를 활용하여 작업자의 위치나 상태를 파악하는 연구가 주로 다뤄져 왔었는데 RFID 태그가 적용된 건설 자재와 작업자의 위치를 실시간으로 추적함으로써 충돌 사고를 예방하는 시스템을 제안되는 형태를 띄고 있었고, 이러한 센서 기반 방식은 작업자가 태그를 항시 소지해야 한다는 불편함과 배터리와 시스템 유지보수 문제, 그리고 태그가 없는 사각지대나 예기치 못한 장애물에 대해서는 감지가 불가능하다는 본질적인 제약이 존재하였고 그 후 머신러닝 기술의 하나인 딥러닝의 발전과 함께 컴퓨터 비전을 활용하는 원격 안전 모니터링 연구가 급격하게 증가하였다.

딥러닝(Deep Learning)은 인공 신경망(Artificial Neural Network)을 여러 층(layer)으로 쌓아 만든 심층 신경망(Deep Neural Network)을 의미한다. 각 층은 이전 층의 출력을 받아 더 높은 수준의 특징을 학습하며, 이를 통해 복잡하고 추상적인 패턴을 인식하고 수 있는데, 2012년 발표된 AlexNet이 이미지 인식 분야에서 혁신적인 성능을 보여주면서 딥러닝의 가능성을 사실로 입증하였으며 CNN연구의 활성화를 촉진하였다. 딥러닝의 하나인 합성곱 신경망(CNN, Convolutional Neural Network)은 컨볼루션, 풀링 등의 연산을 통해 이미지의 공간적 특징을 효과적으로 추출하고, 이를 기반으로 이미지 데이터에서 의미 있는 정보를 학습할 수 있도록 설계되어 있어 이미지 인식, 객체 탐지, 이미지 분할 등 다양한 컴퓨터 비전 문제 해결에 널리 활용되고 있다.

SUNG KYUN KWAN UNIVERSITY

1.2 딥러닝 기반 객체 탐지 모델의 고도화 및 현장 적용

2010년대 중반 이후 합성곱 신경망(CNN)의 비약적인 발전은 건설 현장 영상 분석의 정확도를 획기적으로 개선하였다. 초기 연구인 Girshick 등(2013)의 R-CNN은 객체 후보 영역을 추출하는 방식의 높은 정확도를 보였으나 연산 속도가 느려 실시간 적용에는 어려움이 있었고 추론 속도를 개선하려는 여러 연구가 지속되었는데, Ren 등(2015)이 제안한 Faster R-CNN은 Region Proposal Network(RPN)를 도입하여 탐지 속도를 높였으며, Fang 등(2015)은 이를 건설 현장에 적용하여 비계(Scaffolding) 위에서 작업하는 인원과 구조적 결함을 동시에 탐지하는 연구를 수행, 평균 정밀도(mAP) 90% 이상의 성능을 달성하였다. RPN은 이미지 전체를 스캔하면서, 잠재적으로 객체가 있을 만한 영역(region)을 제안하는데 다양한 크기와 비율로 구분한다. 이후 제안된 영역을 클래스별 분류하여 최종적인 객체 탐지를 하게 된다. 즉, 2단계(2-Stage)의 구조를 통해 객체를 탐지하는 방법이다.

그러나 24시간 모니터링이 필요한 현장 특성상, 1단계(One-stage) 탐지 모델인 YOLO(You Only Look Once) 시리즈가 주류로 자리 잡았다. Redmon 등(2016)이 제안한 YOLO는 특유의 알고리즘과 구조에 기반한 빠른 처리 속도를 바탕으로 이미지 전체를 한 번에 추론하여 실시간 처리가 가능했으며, 건설 현장의 실시간 객체 탐지에 혁신을 가져오게 했다. 또한, Park 등(2023)은 건설 현장의 복잡한 배경과 작업자 간의 겹침(Occlusion) 문제를 해결하기 위하여 작은 객체 탐지에 특화된 개선된 YOLO 모델을 제안하여 밀집된 작업 환경에서도 높은 탐지율을 달성하였다.

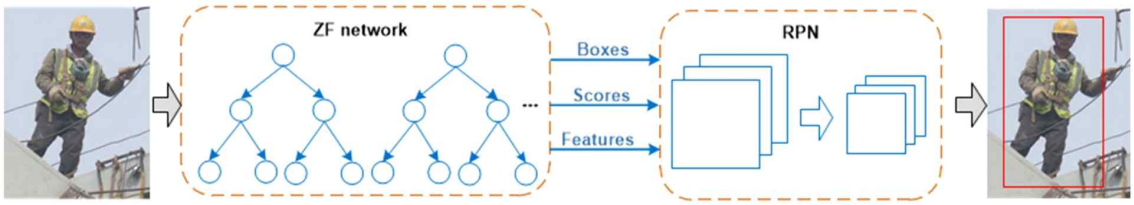


그림 1 Faster R-CNN 기반 객체탐지 파이프라인, Fang 등(2018)

SUNG KYUN KWAN UNIVERSITY

1.3 개인보호구(PPE) 착용 감지 및 중장비 협착 방지 기술

건설 사고의 가장 큰 비중을 차지하는 추락 및 낙하 사고를 예방하기 위해 개인보호구(PPE, Personal Protective Equipment) 착용 여부를 감지하는 연구가 집중적으로 수행되었다. Mneymneh 등(2017)은 현장 영상에서 작업자의 상반신을 검출한 후 색상 히스토그램을 분석하여 안전모 착용 여부를 판별하였으나, 복잡한 배경에서는 오탐율이 높은 한계가 있었고, 이를 개선하기 위해 Wu 등(2019)은 Single Shot Detector(SSD) 기반의 모델을 통해 안전모뿐만 아니라 안전조끼의 색상(적색, 황색 등)까지 분류하여 작업자의 역할군을 식별하는 연구를 진행하였다. 더 나아가 Nath 등(2020)은 YOLOv3 아키텍처를 기반으로 '작업자', '안전모', '조끼'를 계층적으로 탐지하고, 베이지안 추론을 결합하여 작업자가 자재를 들고 있거나 뒤를 돌아선 상태에서도 보호구 착용 여부를 정확히 추정하는 알고리즘을 제안하는 한편, Park 등(2024)은 개구부로 인한 추락 사고를 방지하기 위해, 사각형 검출 알고리즘과 작업자 자세 추정 기술을 결합하여, 작업자가 개구부 주변 위험 반경에 접근하거나 불안정한 자세를 취할 때 즉각 경보를 울리는 시스템을 제안하였다.

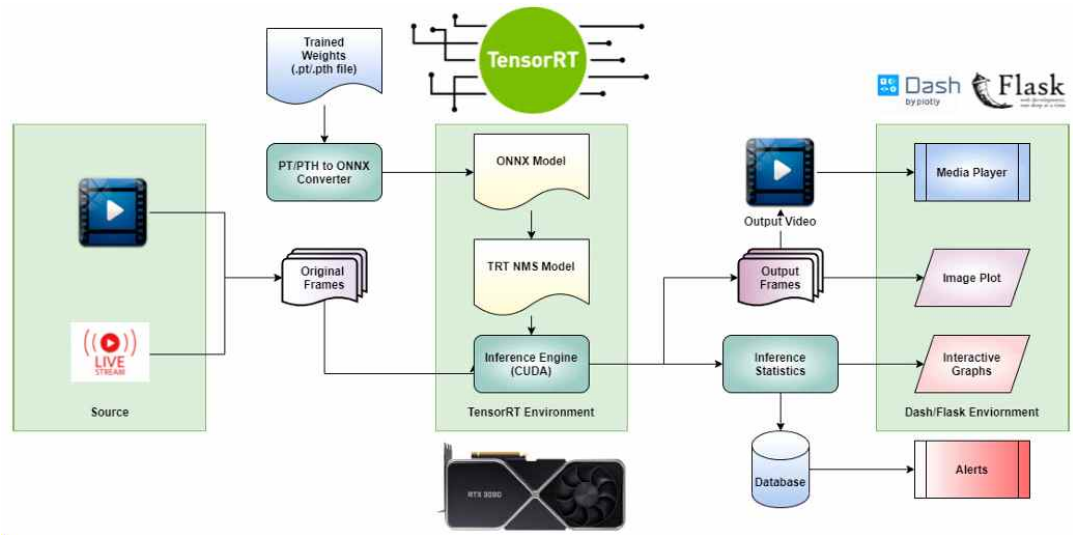


그림 2 실시간 PPE 감지 시스템, Al-Azani 등(2024)

SUNG KYUN KWAN UNIVERSITY

Al-Azani 등(2024)은 CCTV 영상을 기반으로 작업자의 개인보호구(PPE) 착용 여부를 실시간으로 감지하는 시스템을 제안한 바 있다. 이들은 CenterNet, DETR, YOLOv7 등 다양한 최신 객체 탐지 모델의 성능을 비교 분석하였으며, 그 중 가장 우수한 성능을 보인 YOLOv7 모델을 적용하여 기존 데이터셋 환경에서는 92.36%라는 유의미한 정확도를 달성하였다. 다만, 학습되지 않은 새로운 현장 환경에서는 탐지 정확도가 66.70%로 다소 하락하는 경향을 보였는데, 이는 단일 객체 탐지 기술만으로는 다양한 현장 변수에 강건하게 대응하는 데 한계가 있음을 시사한다.

1.4 행동 인식 및 인체공학적 위험 평가

단순히 객체의 존재 여부를 탐지하는 차원을 넘어, 작업자의 구체적인 행동 패턴을 분석함으로써 잠재적 위험을 선제적으로 예측하려는 시도 또한 활발히 진행되고 있다. Ding 등(2018)은 CNN과 LSTM(Long Short-Term Memory)을 결합한 하이브리드 아키텍처를 통해, 영상 시퀀스 내에서 작업자가 사다리를 오르거나 난간에 매달리는 등의 시계열적 동작을 학습시켰다. 이 연구는 불안전 행동을 92%의 정확도로 분류해냄으로써, 기존 정지 영상(Static Image) 분석이 갖는 시간적 맥락(Temporal Context)의 부재라는 한계를 기술적으로 보완했다는 점에서 학술적 의의를 찾을 수 있다.

한편, 근골격계 질환 예방을 위한 인체공학적 접근으로는 Seo 등(2015)이 키넥트(Kinect) 센서와 컴퓨터 비전을 융합하여 작업자의 관절 포인트(Skeleton)를 추출하고, 이를 OWAS(Ovako Working Posture Analysis System) 기법에 대입하여 허리나 어깨에 과도한 부하가 걸리는 자세를 실시간으로 평가하는 시스템을 제안하였으며, Park 등(2023)은 이동식 사다리 작업 중 발생할 수 있는 불안전 행동을 실시간으로 감지하기 위해 비전 데이터 기반의 딥러닝 모델을 활용하였으며, 작업자의 자세와 사다리의 상태를 동시에 분석하여 전도 사고를 예방하는 기술을 실증하였고, 전소연 등(2020) 또한 딥러닝 기반 영상 분석 알고리즘을 통해 작업자의 쓰러짐이

SUNG KYUN KWAN UNIVERSITY

나 이상 행동을 실시간으로 감지하고 관리자에게 알리는 통합 안전관리 시스템을 개발하여 국내 현장 적용성을 높였다.

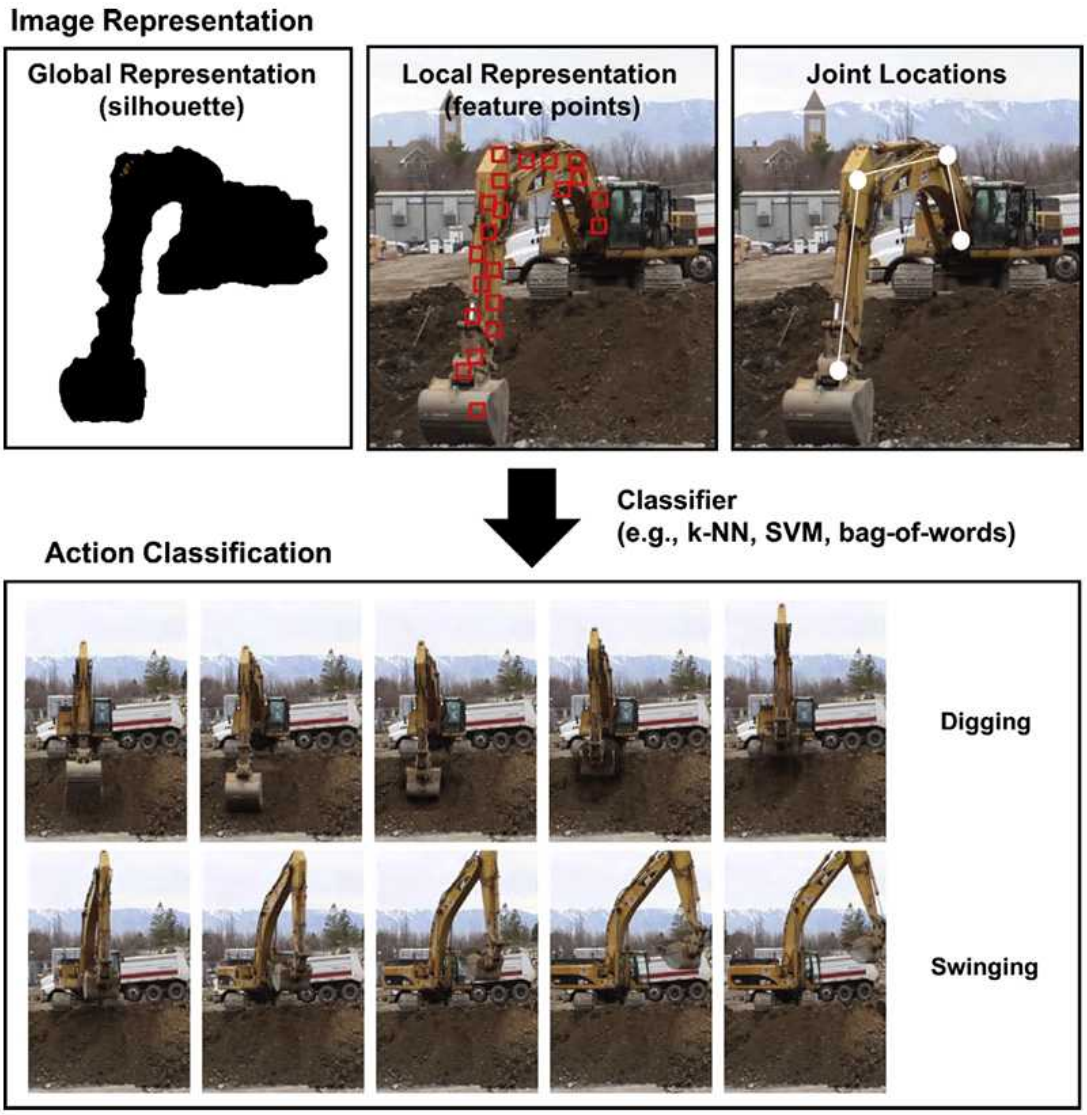


그림 3 융합데이터를 이용한 행동인식 절차, Seo 등(2015)

SUNG KYUN KWAN UNIVERSITY

1.5 VLM을 이용한 시각적 문맥 이해

2023년을 기점으로 컴퓨터 비전 분야는 단순 객체 탐지를 넘어, 이미지와 자연어를 통합적으로 이해하는 대규모 시각-언어 모델(Large Vision-Language Model, VLM) 중심의 연구로 급격히 전환되었는데, 그 중심에서 OpenAI의 GPT-4V와 Google의 Gemini 등 파운데이션 모델의 등장은 비정형화된 건설 현장 이미지에서 '안전모 미착용'과 같은 단순 사실뿐만 아니라, '작업자가 불안정한 발판 위에서 균형을 잃고 있다'는 인과적 맥락'을 추론할 수 있는 가능성을 열었으며, Liu 등(2024)은 시각적 인스트럭션 튜닝(Visual Instruction Tuning)을 적용한 오픈소스 모델 LLaVA(Large Language and Vision Assistant)를 제안하였으며, 이를 통해 상용 API를 사용하지 않고도 이미지에 대한 복합적인 질문에 답변할 수 있는 온프레미스 환경 구축의 기틀을 마련하였다.

건설 도메인 적용과 관련한 다양한 시도들도 많았는데 Adil 등(2025)은 건설 현장에서 나타날 수 있는 위험 시나리오에 기반한 이미지와 텍스트 쌍으로 VLM을 미세 조정한 'Construction-GPT' 개념을 제안하였는데, 이 연구에서 VLM은 기존 객체 탐지 모델은 식별이 불가능한 '정돈되지 않은 자재로 인한 전도 위험'과 같은 잠재적 위험 요소를 텍스트로 상세히 설명하는 데 성공하였으며 위험도 평가 작업에서 인간 전문가와 85% 이상의 일치율을 보이는 한편, Wang 등(2024)은 텍스트와 이미지의 의미적 유사성을 분석하여 잠재적 위험 요소를 식별하는 프레임워크를 발표하여 기존의 사전 정의된 클래스만 탐지하던 한계를 극복하고, 생성형 AI가 현장의 다양한 문맥을 해석하여 예기치 못한 위험까지 사전에 경고할 수 있음을 실증하였다.

Kim(2023)은 텍스트-이미지 합성기술을 활용하여 건설 현장의 다양한 위험 상황 이미지를 생성하고, 이를 DNN 모델 학습에 활용함으로써 아직 발생하지 않은 상황의 데이터 부족 문제를 해결하고 모델의 일반화 성능을 향상시키는 연구를 수행하였으며, 김현우 등(2025)은 영상뿐만 아니라 소리 데이터까지 분석하는 멀티모달 신

SUNG KYUN KWAN UNIVERSITY

경망을 적용하여, 영상에서는 감지하기 불가능한 장비의 이상 소음 또는 작업자의 비명 소리 등을 종합적으로 인지하는 지능형 안전 시스템을 제안하였다.

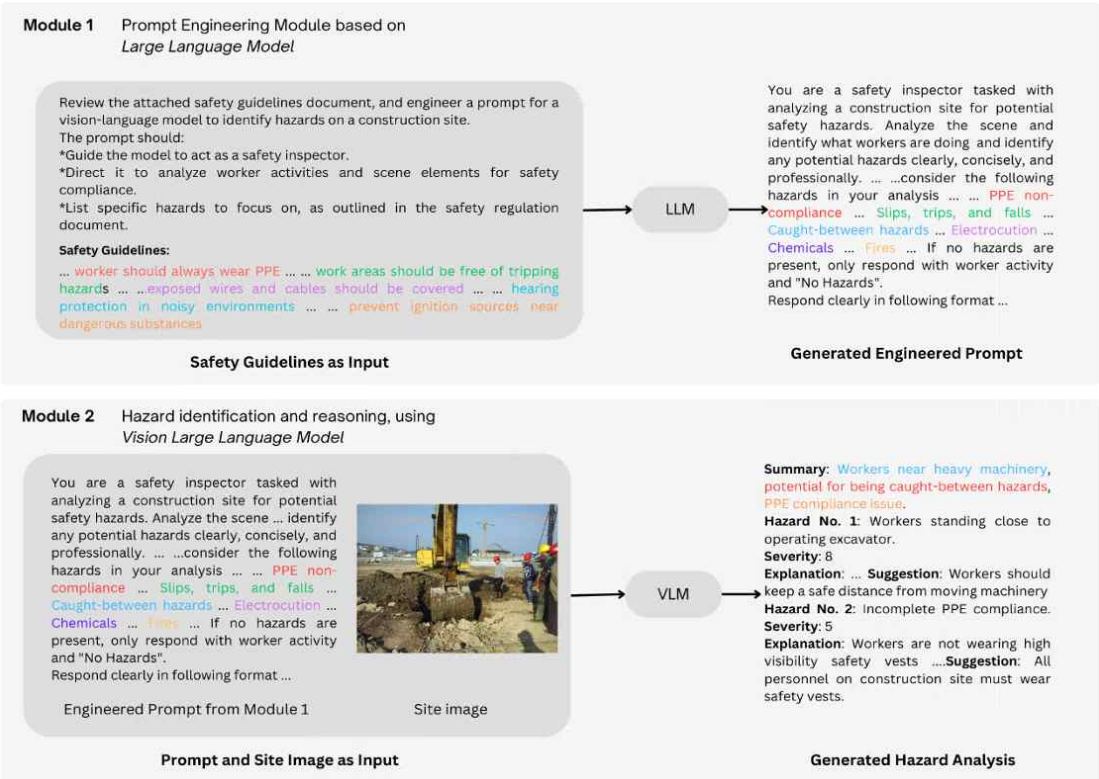


그림 4 Construction-GPT, Adil 등(2025)

1.6 온프레미스 환경을 위한 sLLM 및 경량화 기술

거대 모델의 성능을 유지하면서도 연산 비용을 획기적으로 낮춘 sLLM(Small Large Language Model) 연구가 2023년 하반기부터 활발히 진행되고 있는데 대표적으로 Touvron 등(2023)이 공개한 Llama 2와 이후의 Llama 3 모델은 7B(70억) 파라미터 수준의 경량 모델로도 적절한 튜닝을 거치면 수천억 파라미터 모델에 준하는 추론 성능을 낼 수 있음을 입증하였고, Jiang 등(2023)은 슬라이딩 윈도우 어

SUNG KYUN KWAN UNIVERSITY

텐션(Sliding Window Attention) 기술을 적용한 Mistral 7B 모델을 공개하였고 제한된 GPU 메모리를 가진 엣지 서버에서도 긴 문맥(Long Context)을 처리할 수 있는 효율성을 증명하였다.

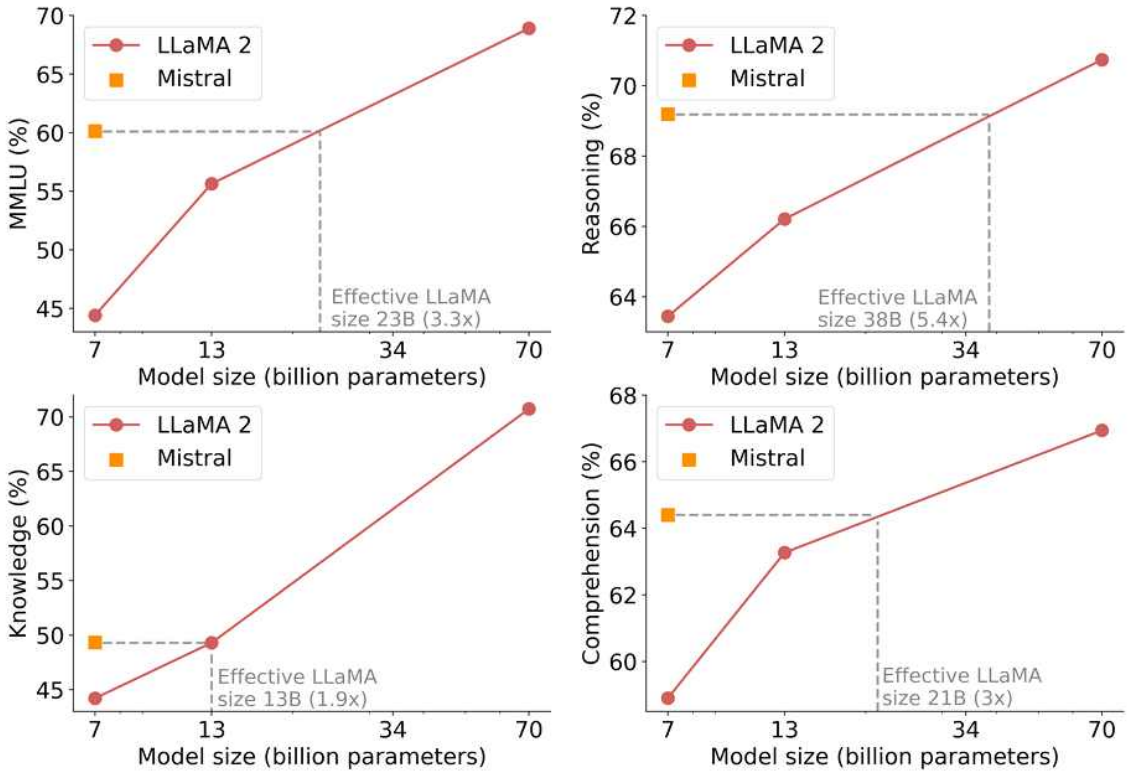


그림 5 LLM과 sLLM의 성능 비교, Jiang 등(2023)

1.7 RAG 기반의 할루시네이션 억제 및 법규 준수

생성형 AI의 치명적 단점인 환각(Hallucination) 현상을 억제하고, 전문적인 건설안전 법규에 근거한 신뢰성 있는 답변을 생성하기 위해 검색 증강 생성(Retrieval-Augmented Generation, RAG) 기술이 필수적으로 도입되고 있는데, Lewis 등(2020)이 정립한 RAG 프레임워크는 건설 도메인에 특화한 지식베이스

SUNG KYUN KWAN UNIVERSITY

(Knowledge Base)와 결합하며 고도화된 시스템을 선보였고. Zhang 등(2024)은 비정형 텍스트뿐만 아니라, 현장의 도면과 시공 계획서정보를 지식 그래프형태로 구조화하여 RAG에 연동하는 'BifrostRAG' 접근법을 시도하며 AI가 "이 작업이 왜 위험한가?"라는 질문에 대해, 단순히 일반적인 안전 상식을 나열하는 것이 아니라, "현재 3층 슬래브 작업은 표준 시방서 제5조에 따른 추락 방지망이 설치되지 않았으므로 위험하다"와 같이 구체적인 근거와 조항을 명시하는 설명 가능한 AI(XAI) 수준의 답변을 생성할 수 있음을 보여주었다.

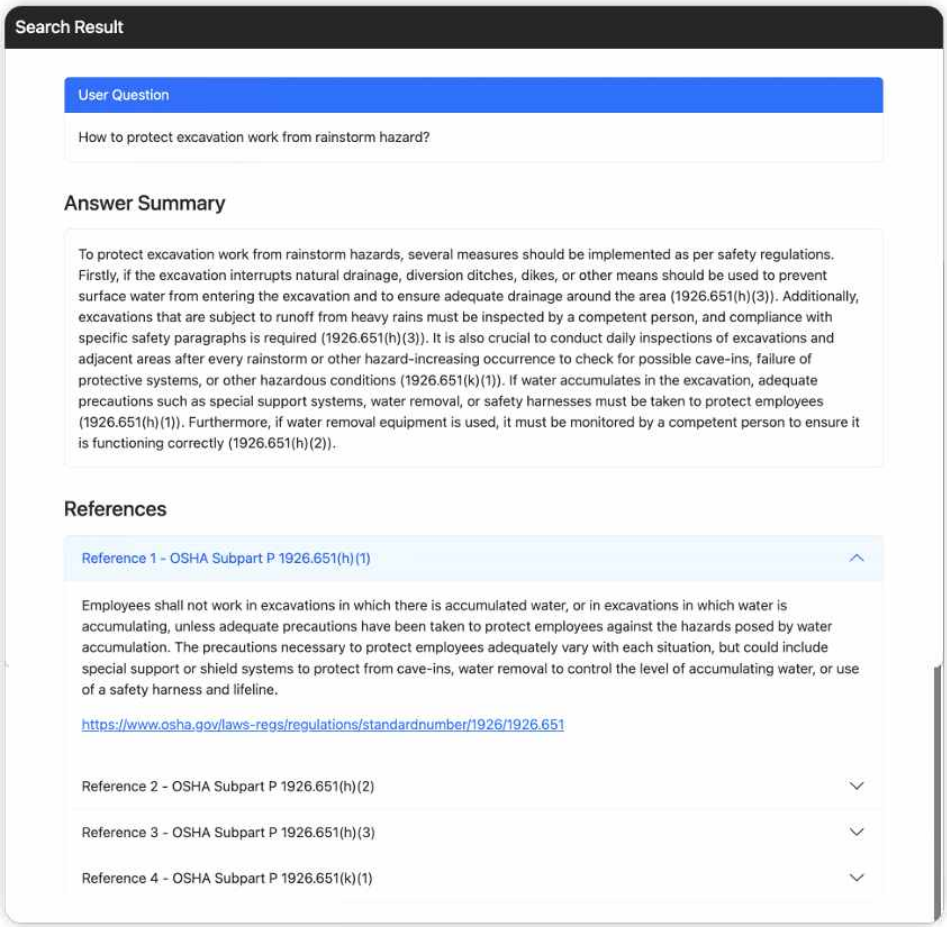


그림 6 BiFrostRAG 수행결과, Zhang 등(2024)

SUNG KYUN KWAN UNIVERSITY

Uhm 등(2025)은 RAG 기능을 차용한 LLM 모델이 일반적인 LLM에 대비하여 안전 정보 생성의 정확도와 구체성을 향상시킴을 정량적으로 입증하기 위하여 안전 전문가와 현장 작업자를 대상으로 한 블라인드 테스트를 수행하였고 RAG를 포함한 모델이 제시한 안전 가이드라인이 실제 규정과의 부합도(Compliance) 면에서 월등히 우수함을 확인하였다.

SUNG KYUN KWAN UNIVERSITY

2. 이론적 배경

2.1 건설 안전 관리 이론

먼저, 하인리히 법칙은 1건의 대형 사고(Major Injury)가 발생하기 전에는 29건의 경미한 사고(Minor Injury)와 300건의 아차 사고(Near Miss)가 선행한다는 경험적 통계 법칙인데 대형 사고가 우연히 발생하는 단발성 사건이 아니라, 현장에 만연한 사소한 위반과 불안전한 상태(Unsafe Condition) 및 행동(Unsafe Act)이 누적된 결과임을 시사한다.

1990년 리즌이 제안한 스위스 치즈 모델은 사고가 단일 원인이 아닌, 조직적·기술적·개인적 결함들이 동시에 정렬될 때 발생한다고 설명하였고 안전 관리 체계의 각 요소(규정, 관리자 감독, 물리적 안전장치)를 치즈의 슬라이스로 비유할 때, 각 슬라이스에 뚫린 구멍(결함)들이 일직선상에 놓이면 사고가 통과하게 되는 것이며 건설 현장에서는 관리자의 감독 공백(Supervision Gap)이나 개인의 부주의(Human Error)가 이러한 구멍에 해당한다.

마지막으로 국내 건설 안전 관리의 법적 근거는 크게 '재난 및 안전관리 기본법', '산업안전보건법', 그리고 '중대재해처벌법'으로 체계화되었으며 '재난 및 안전관리 기본법'이 국가 차원의 재난 예방 및 대응을 위한 기본 이념을 제시한다면, '산업안전보건법'은 사업주에게 구체적인 안전 조치 의무를 부과하여 산업 재해를 예방하는 데 중점을 두었고 특히 2022년 시행된 '중대재해처벌법'은 경영 책임자에게 안전에 관한 확보 의무를 강력히 부여하여, 이를 위반하여 중대 재해가 발생하는 경우에는 형사 처벌을 강화하였는데 이것은 기업으로 하여금 형식적인 안전 점검을 넘어, 실질적인 안전 관리 체계를 구축하도록 강제하는 강력한 제약으로 작용하도록 하고 있다.

SUNG KYUN KWAN UNIVERSITY

2.2 인공지능 기초 이론

텐서(Tensor)는 딥러닝에서 데이터를 표현하는 다차원 배열(Multi- dimensional Array)이다. 스칼라(0차원), 벡터(1차원), 행렬(2차원)을 일반화한 개념으로, 컬러 이미지는 (높이 × 너비 × 채널) 형태의 3차원 텐서로 표현된다. CNN은 이러한 이미지 텐서에서 유의미한 시각적 특징을 추출하는 핵심 알고리즘으로 CNN의 합성곱 층은 필터가 이미지를 슬라이딩하며 국소적인 특징(엣지, 텍스처 등)을 감지하고, 풀링 층은 중요 특징을 유지하면서 데이터의 차원을 축소하여 연산 효율성을 높이며 이 과정을 반복함으로써 모델은 단순한 픽셀 정보를 넘어 객체의 고차원적인 형태 정보를 학습하게 된다.

한편, 자연어 처리(NLP, Natural Language Processing) 모델은 텍스트를 기계가 이해할 수 있는 숫자로 변환하는 과정을 거치는데 먼저 문장을 의미 있는 최소 단위인 토큰(Token)으로 분리하고, 각 토큰을 고차원 벡터 공간상의 좌표인 임베딩(Embedding)으로 변환하고 이때 '왕'과 '남자'의 관계가 '여왕'과 '여자'의 관계와 같도록, 의미적으로 유사한 단어들은 벡터 공간에서 가까운 거리에 위치하게 된다. 최신 LLM의 기반이 되는 트랜스포머(Transformer) 아키텍처는 어텐션(Attention) 메커니즘을 통해 문장 내 단어들 간의 상관관계를 계산하는데 어텐션은 문맥상 중요한 단어에 더 높은 가중치를 부여함으로써, "그는 강가에 있는 은행에 갔다"라는 문장에서 '은행'이 금융기관이 아닌 둑(Bank)을 의미함을 정확히 파악하게 해준다.

2.3 실시간 객체 탐지 및 추적 기술

건설 현장의 모니터링 시스템은 다수의 CCTV 채널을 지연 없이 처리하는 것이 가장 우선되므로 모델의 연산 기능이 다른 무엇보다 중요하게 되었는데 YOLO는 이미지를 그리드(Grid)로 분할하고, 각 그리드에서 객체의 위치(Bounding Box)와 클래스를 동시에 예측하는 방식을 사용하였고 이는 후보 영역 추출과 분류가 분리되어 있는 기존 방식(2-Stage)에 비해 연산 단계를 획기적으로 줄여주며, 이미지 전

SUNG KYUN KWAN UNIVERSITY

체를 한 번에 조망하여 학습하기 때문에 배경을 객체로 오인하는 오류가 적다는 강점을 가진다. 특히 최신 YOLO 아키텍처에서는 앵커 프리(Anchor-free) 방식이 도입되었는데 사전에 정의된 비율의 박스에 의존하지 않고 객체의 중심점과 크기를 직접 예측하면서 쓰러진 작업자나 형태가 일그러진 자재와 같이 비정형적인 객체가 빈번한 건설 현장에서 탐지의 유연성과 강건성을 확보하는 데 결정적인 역할을 할 것으로 기대된다.

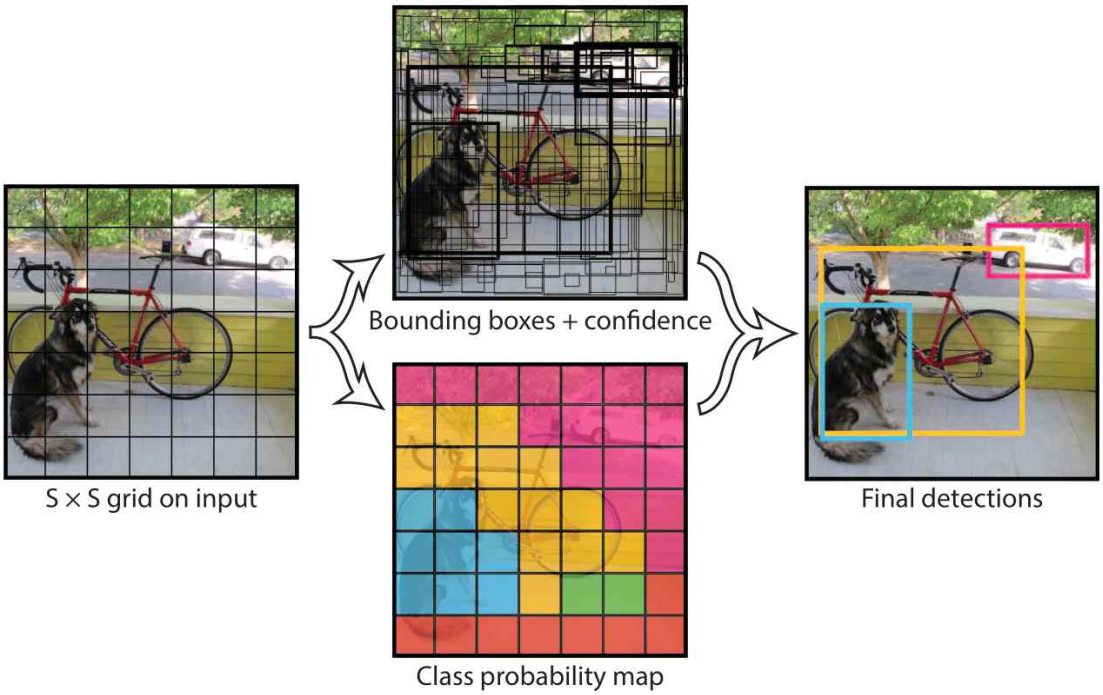


그림 7 YOLO의 그리드(Grid) 처리 방식 개념도, Redmon 등(2016)

단순히 객체를 탐지하는 것을 넘어서 작업자의 연속적인 동선을 파악하고 위험 구역에서 체류한 시간을 산출하기 위해서는 다중 객체 추적(Multi-Object Tracking, MOT) 기술이 필수적인데, 탐지된 객체의 프레임 간 이동 경로를 연결하는 대표적인 알고리즘인 SORT(Simple Online and Realtime Tracking)는 칼만 필터(Kalman

SUNG KYUN KWAN UNIVERSITY

Filter)를 사용하여 이전 프레임에서의 객체 속도와 방향을 기반으로 현재 위치를 예측하고, 실제 탐지된 위치와의 오차를 보정하여 추적을 유지하게 하는데 이를 통하여 작업자가 자재 뒤로 잠시 가려지거나(Occlusion) 카메라 앵글이 겹치는 상황에서도 개별 ID를 유지하며 추적할 수 있어, 연속적으로 각 객체의 안전 위반 행위를 분석한 데이터를 제공할 수 있게 한다.

2.4 시각적 문맥 이해와 관계 추론

컴퓨터 비전 분야에서 객체 간의 상호작용(Interaction)을 이해하기 위한 전통적인 접근법 중 하나인 씬 그래프(Scene Graph)는 이미지 내의 객체를 노드(Node)로, 객체 간의 관계를 엣지(Edge)로 정의한다. 이는 비정형적인 이미지 데이터를 <주어-서술어-목적어> 형태의 구조화된 데이터(Structured Data)로 변환하는 기술적 토대를 제공한다. 엔지니어링 관점에서 볼 때, 씬 그래프는 단순 탐지 모델이 식별하기 어려운 '작업자'와 '망치' 사이의 <들고 있다(Holding)>와 같은 관계적 맥락을 명시적으로 표현할 수 있어, 위험 상황의 인과관계를 파악하는 데 유용한 이론적 기반이 된다.

최근에는 별도의 그래프 생성 모듈을 두는 대신, 대규모 데이터로 학습된 시각언어 모델(VLM)을 통해 이미지 내의 관계 정보를 암시적으로 추론하는 연구가 주류를 이루고 있는데, 이는 명시적인 그래프 구조를 생성하지 않더라도, 모델 내부의 어텐션(Attention) 메커니즘이 객체 간의 시공간적 연관성을 학습하여 "작업자가 사다리에서 떨어지고 있다"와 같은 복합적인 상황을 자연어로 묘사할 수 있게 함을 의미한다.

2.5 시각언어 모델(VLM) 및 최적화 기술

본 연구의 핵심 추론 엔진인 VLM(Vision-Language Model)은 시각 정보와 언어 정보라는 이질적인 두 개의 모달리티를 하나의 통합된 임베딩 공간에서 처리하는

SUNG KYUN KWAN UNIVERSITY

모델로 이미지를 압축하여 추상적인 특징을 추출하는 비전 인코더(Vision Encoder)와 텍스트 생성을 처리하는 언어 디코더(Language Decoder) 그리고 이 둘을 연결하는 프로젝션 레이어(Projection Layer)로 구성된다. 프로젝션 레이어는 비전 인코더가 출력한 2차원 또는 1차원 벡터를 LLM이 이해할 수 있는 텍스트 임베딩 벡터의 차원으로 선형 변환(Linear Transformation)하는 역할을 수행하게 되는데 엔지니어링 측면에서 이는 별도의 규칙 정의(Hard-coding) 없이도, 모델이 "이 작업자가 왜 위험한가?"와 같은 개방형 질문에 대해 이미지 내의 다양한 단서들을 조합하여 논리적으로 답변할 수 있는 유연성을 제공한다.

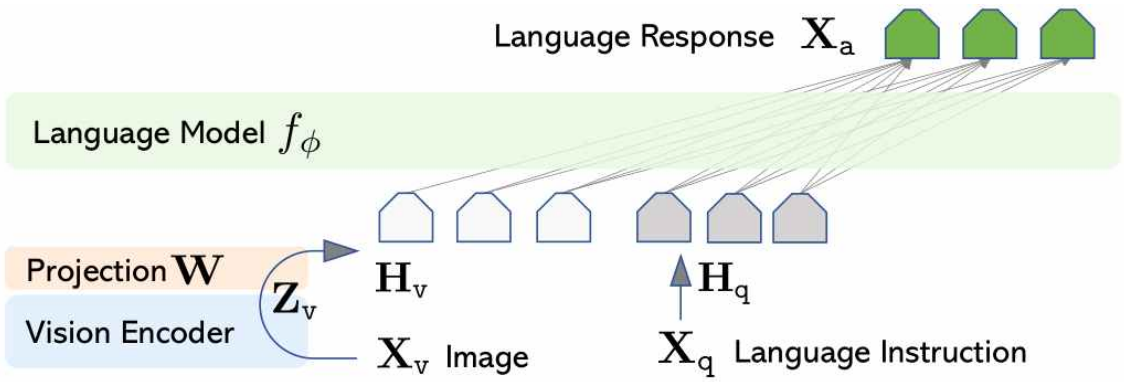


그림 8 LLaVA 아키텍처, Liu 등(2023)

한편 현장 사무실의 일반적인 워크스테이션 환경에서 수십억 파라미터 규모의 거대 모델을 구동하기 위해서는 모델 경량화 기술이 필수적으로 요구되고 있는 상황으로 그중 양자화(Quantization)는 모델의 가중치 데이터를 32비트 고정밀도 실수형(FP32)에서 4비트 또는 8비트 정수형(INT4/8)으로 압축하여 메모리 대역폭 소모를 줄이고 추론 속도를 가속화하는 핵심 기술이다. 예를 들어, FP32 모델을 INT4로 양자화할 경우 모델 사이즈와 메모리 대역폭 요구량은 이론적으로 1/8로 감소하여 추론 속도가 비약적으로 향상되는 한편 정밀도 하락으로 인한 정보 손실과 추론 정확도 저하(Accuracy Degradation)는 불가피하나, 최신 연구들은 대표적인 데이터셋

SUNG KYUN KWAN UNIVERSITY

(Calibration Set)을 이용하여 가중치 분포의 다이나믹 레인지(Dynamic Range)를 분석하고, 이상치(Outlier)를 보존하는 정교한 보정(Calibration) 과정을 통해 성능 저하를 1% 미만으로 억제하며 엣지 디바이스에서의 운용 효율을 극대화하고 있다.

2.6 검색 증강 생성 (RAG) 및 벡터 데이터베이스

일반적인 AI 모델은 학습데이터에 존재하지 않았던 건설 분야의 전문적인 안전 수칙을 알지 못하거나, 사실과 다른 정보를 생성하는 환각(Hallucination) 현상을 보일 수 있는데 이를 해결하기 위해 일반적으로 모델을 파인튜닝하기 보다는 검색 증강 생성(Retrieval-Augmented Generation, RAG) 기술을 도입한다.

RAG의 핵심은 '임베딩(Embedding)'과 '유사도 측정(Similarity Measure)'으로 볼 수 있는데 먼저 산업안전보건법, 표준시방서 등 텍스트 형식의 지식 데이터를 임베딩 모델을 통해 고차원의 벡터(Vector)로 변환하여 벡터 데이터베이스에 저장하게 되는데 벡터 공간 상에서 의미적으로 유사한 텍스트들은 서로 가까운 거리에 위치하게 되므로 비정형 텍스트 데이터베이스를 '의미 검색(Semantic Search)'이 가능한 형태로 구조화하는 엔지니어링 과정이라고 할 수 있다.

즉, VLM이 분석한 현장의 위험 상황 텍스트(Query)가 입력되면, 시스템은 벡터 데이터베이스에서 이와 가장 의미적으로 유사한 법규 조항을 검색하는데 키워드 매칭 방식과 달리, 벡터 유사도 검색은 단어가 정확히 일치하지 않아도 맥락이 유사하면 관련 정보를 찾아낼 수 있게 되고 검색된 법적 근거는 프롬프트 엔지니어링을 통해 AI 모델에 입력 정보로 함께 제공(Augment)되어 결과적으로 AI는 자신의 사전 지식이 아닌, 검색된 신뢰할 수 있는 법령 정보를 바탕으로 답변을 생성하게 되며, 이는 안전 리포트의 법적 정합성과 신뢰도를 보장하는 기술적 안전장치 역할을 한다.

SUNG KYUN KWAN UNIVERSITY

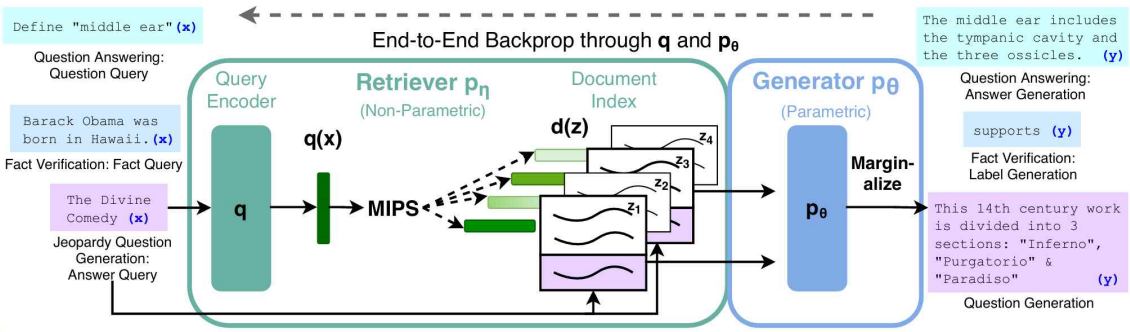


그림 9 RAG 파이프라인 흐름도, Lewis 등(2020)

2.7 엣지 AI 시스템 아키텍처 및 서빙 최적화

현대의 지능형 영상 분석 시스템은 단순히 높은 정확도의 AI 모델을 개발하는 것을 넘어, 이를 제한된 컴퓨팅 자원 하에서 효율적으로 운용하는 시스템 엔지니어링 능력이 필수적으로 요구되는데 특히 네트워크 대역폭이 불안정하고 하드웨어 확장이 어려운 엣지 환경에서는 고성능 추론 서빙(Inference Serving) 기술이 시스템의 가용성을 결정짓는 핵심 요소로 작용한다. 본 연구로 개발된 AI 모델을 실제 서비스에 배포하기 위해서는 모델의 가중치를 메모리에 적재하고, 클라이언트의 요청(Request)을 효율적으로 처리하는 서빙 프레임워크가 필요한데, NVIDIA Triton Inference Server는 다중 모델을 단일 GPU에서 동시에 구동할 수 있도록 지원하며, 동적 배칭(Dynamic Batching) 기술을 통해 개별적으로 유입되는 추론 요청을 설정된 시간 윈도우(Time Window) 동안 모아 하나의 배치(Batch)로 처리함으로써 GPU 활용률(Utilization)과 처리량(Throughput)을 극대화할 수 있었다. 한편, LLM과 같은 거대 모델의 서빙을 위해서는 Ollama나 vLLM과 같은 특화된 런타임이 활용되는데 모델의 레이어(Layer)를 GPU 메모리에 최적화하여 매핑하고, KV 캐시(Key-Value Cache) 관리 및 양자화(Quantization) 커널을 내장하여, 범용 딥러닝 프레임워크(PyTorch 등) 대비 현격히 낮은 지연 시간과 적은 메모리 점유율을 달성한다.

SUNG KYUN KWAN UNIVERSITY

3. 소결

본 장에서는 건설 현장의 안전 모니터링 고도화를 위해 필수적인 핵심 요소 기술들의 이론적 배경과 최신 연구 동향을 고찰을 통하여 기존의 단순 객체 탐지 기술은 실시간성은 우수하나 위험 상황의 인과관계를 설명하는 데에는 한계가 있었으며, 이를 보완할 수 있는 시각언어 모델(VLM)은 높은 연산 비용과 환각 현상으로 인해 현장 도입에 제약이 따름을 확인하였다. 따라서 본 연구는 이러한 기술적 트레이드오프를 해결하기 위해, 1) 실시간 탐지(YOLO)와 심층 추론(VLM)을 결합한 하이브리드 파이프라인, 2) 엣지 환경에서의 운용을 위한 모델 양자화 및 최적화, 3) 법적 신뢰성 확보를 위한 RAG 기술을 통합적으로 적용하고자 하며 이어지는 제3장에서는 이러한 이론적 토대를 바탕으로, 실제 건설 도메인 데이터에 기반한 다각적인 벤치마크 실험을 수행하고 이를 통해 최적의 모델 조합과 운용 파라미터를 도출하여 제안하는 온프레미스 시스템의 기술적 타당성과 경제적 실효성을 실증적으로 검증할 것이다.

SUNG KYUN KWAN UNIVERSITY

제3장 실험 및 성능 평가

1. 실험 개요 및 목표

본 연구가 제안하는 '현장 중심형 안전 인지 프레임워크'가 실제 건설 현장에서 실효성을 갖기 위해서는, 본격적인 시스템 설계에 앞서 핵심 요소 기술에 대한 면밀한 타당성 검토와 최적화 과정이 선행되어야 한다. 특히, 건설 현장이 내포한 열악한 통신 인프라와 제한적인 엣지(Edge) 컴퓨팅 자원, 그리고 엄격한 데이터 보안 규정은 고성능 모델의 도입을 제약하는 실질적인 장벽으로 작용한다. 이에 본 장에서는 도메인에 적합한 모델을 선별하고, 실제 운용 시나리오를 가정한 경제성 시뮬레이션을 수행하고자 한다. 이러한 실험적 데이터는 향후 제안될 프레임워크의 기술적 사양과 운용 전략을 결정하는 객관적인 지표로 활용될 것이며, 각 실험 항목이 플랫폼 설계 과정에서 갖는 당위성과 검증 목표는 아래 표와 같다.

검증 요소	플랫폼 설계 근거 및 필요성	수행 실험
핵심 AI 엔진	● 범용 모델은 복잡한 가설 구조물 식별과 국내법(산업안전보건법) 기반의 리포팅에 한계가 있음 ● 현장 특수성(한국어, 법규, 객체)을 충족하는 **최적의 모델 조합(VLM+sLLM)** 선정이 선행되어야 함	도메인 적합성 평가 및최적 모델 선정
운용 전략	● 24시간 분석에 따른 클라우드(API) 비용 부담을 해소해야 함 ● 초기 구축비 대비 운영비 절감 효과를 분석하여, **온프레미스 도입의 경제적 타당성**을 입증해야 함	운용 경제성 분석

표 2 플랫폼 핵심 요소 결정을 위한 설계 근거 및 수행 실험

SUNG KYUN KWAN UNIVERSITY

실험을 위한 하드웨어 환경은 현장의 예산 제약과 설치 공간의 한계를 고려하여 단일 워크스테이션으로 제한하였는데 특히 24GB의 VRAM은 최신 대규모 모델(LLM/VLM)을 구동함에 있어 가장 중요한 물리적 제약 조건으로 작용할 것이며 상세한 실험 환경은 다음과 같다.

구분	세부 사양
CPU	24 Core
GPU	24GB VRAM (CUDA 12.8)
RAM	64GB
OS	Ubuntu 24.04 LTS
추론엔진	Ollama (v0.13.0), NVIDIA Triton Inference Server (25.05)
구동환경	Docker Container

표 3 하드웨어 및 소프트웨어 구성

SUNG KYUN KWAN UNIVERSITY

2. 실험 1: 도메인 적합성 평가 및 최적 모델 선정

2.1 배경 및 문제 정의

오픈소스 커뮤니티를 통해 공개된 최신 멀티모달 모델(VLM)과 소형 언어 모델(sLLM)들은 비약적인 성능 향상을 이루었으나, 이들을 건설 안전 도메인에 그대로 적용하기에는 '도메인 불일치(Domain Mismatch)' 문제가 존재한다. 건설 현장의 비정형적인 가설 구조물이나 전문적인 법률 용어는 범용 모델이 학습하지 못한 데이터 분포이기 때문에 가용 가능한 후보 모델들을 대상으로 건설 도메인 특화 성능을 정량적으로 벤치마킹하여, 시스템의 '눈'과 '뇌' 역할을 수행할 최적의 모델 조합을 실증적으로 규명할 필요가 있다.

2.2 실험 목적

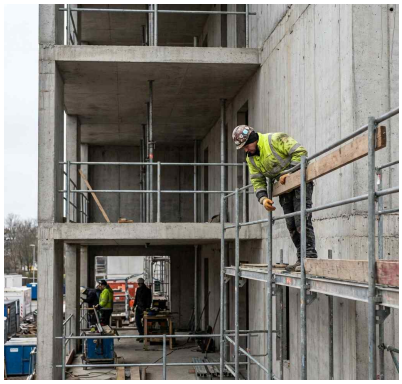
본 실험은 파라미터 30B(Billion) 이하의 경량 모델들을 대상으로 비교 평가를 수행하여, 시각 정보를 텍스트로 변환하는 시각적 인지 모델(VLM)과 이를 바탕으로 법적 리포트를 생성하는 논리적 추론 모델(sLLM)의 최적 조합을 도출하는 것을 목적으로 한다.

2.3 실험 설계 및 방법

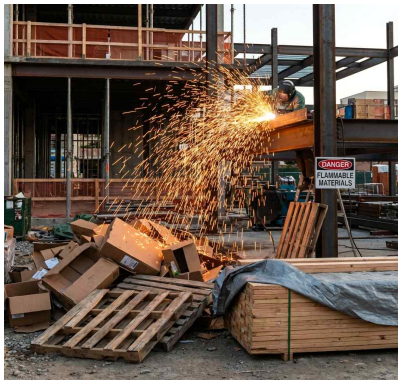
+ 비교 모델: Gemma-3 (27B, 12B, 4B), Llama-3.2-Vision (11B), LLaVA-13B
+ 평가 데이터: 건설 현장의 5대 핵심 재해(추락, 협착, 전도, 화재, 낙하)를 포함하는 시나리오 이미지 각 5장에 대하여 3회씩 반복 추론 수행
+ 평가 기준:
  - 객체 식별력: 복잡한 배경 속에서 작업자, 안전고리(Hook), 비계 파이프 등 미세 객체의 식별 정확도

SUNG KYUN KWAN UNIVERSITY

- OCR(Optical Character Recognition) 성능: 이미지 내 포함된 안전 표지판(예: '추락주의', '출입금지')의 한글 텍스트 판독 능력



추락



화재

중장비

안전보호구

한글판독

그림 10 시나리오별 평가 이미지(예시)

SUNG KYUN KWAN UNIVERSITY

시스템 프롬프트는 아래와 같이 작성하여 총 3가지의 핵심 과업(객체 확인, 위험 요소, OCR)을 정의하고 반드시 한국어로 출력하는 시스템 페르소나를 설정하였다.

```
SYSTEM_PROMPT = """
Analyze this construction site image for safety issues.
Describe:
1. Objects (Workers, Equipment)
2. Safety violations (helmet, harness)
3. Text on signs
Output in Korean.
"""
```

표 4 실험1에서 사용한 시스템 프롬프트

2.4 실험 결과 요약

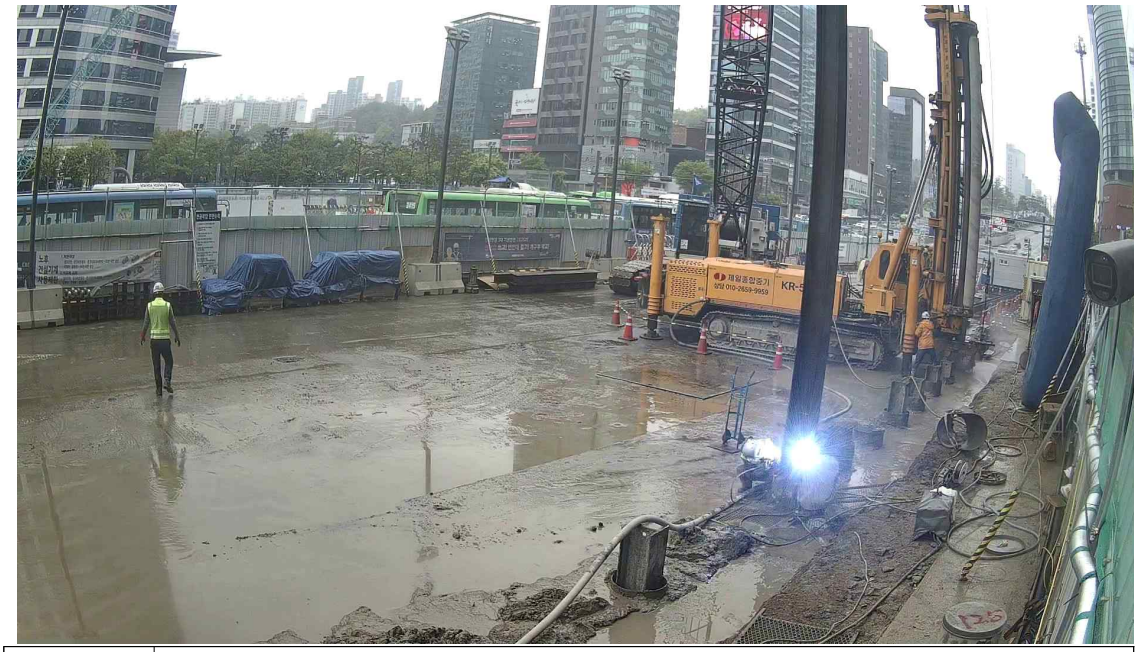
본 실험에서는 Vision-Language Model(VLM)의 건설 도메인 적합성을 평가하기 위해 Gemma-3 시리즈(27B, 12B, 4B), Llama-3.2-Vision, LLaVA-13B를 대상으로 5가지 핵심 재해 시나리오(추락, PPE, 중장비, 화재, 표지판)를 테스트하였고 주요 정량적 지표와 정성적 평가결과는 아래의 표와 같다.

모델	평균 추론 시간 (단위: 초)	한국어 유창성	위험 식별 정확도	종합 평가
Gemma-3 (27B)	31.48s	최상	높음	**채택**
Gemma-3 (12B)	12.49s	중 (English Mixed)	중간	후보군
Gemma-3 (4B)	6.81s	하 (Broken/Translit)	낮음	부적합
Llama-3.2 (11B)	10.39s	중하 (English Dominant)	중간	언어 장벽 존재
LLaVA (13B)	8.98s	하 (Hallucination)	낮음	부적합

표 5 실험1(도메인적합성)의 결과

SUNG KYUN KWAN UNIVERSITY

각각의 이미지를 추론한 결과는 아래와 같다. 모델명 옆의 숫자는 추론시간을 의미한다.



gemma3:27b (36.25)	## 건설 현장 안전 분석 (건설 현장 안전 분석) **1. 객체 (객체)** * **작업자 (작업자):** 현장 내에 최소 2명의 작업자가 보입니다. 한 명은 용접 작업을 하고 있고, 다른 한 명은 주변에서 지켜보고 있습니다. * **장비 (장비):** * **파일 드라이버 (파일 드라이버):** 큰 노란색 파일 드라이버 장비가 보입니다. * **용접기 (용접기):** 용접 작업자가 사용하고 있는 용접기. * **호스 (호스):** 파일 드라이버 주변에 여러 개의 호스가 놓여 있습니다. * **컨테이너 (컨테이너):** 현장 주변에 컨테이너 박스가 여러 개 있습니다. * **안전 펜스 (안전 펜스):** 현장 주변을 둘러싼 녹색 안전 펜스. * **기타 (기타):** 진흙이 고여있는 현장 바닥, 건물, 도로 등.

SUNG KYUN KWAN UNIVERSITY

	2. 안전 위반 사항 (안전 위반 사항) * **안전모 미착용 (안전모 미착용):** 용접 작업을 하고 있는 작업자는 안전모를 착용하고 있지만, 주변에서 지켜보고 있는 작업자는 안전모를 착용하지 않고 있습니다. 이는 안전 규정 위반입니다. * **안전 장비 부족 (안전 장비 부족):** 용접 작업자는 용접 마스크를 착용하고 있지만, 다른 안전 장비 (보안경, 장갑 등) 착용 여부는 확인하기 어렵습니다. * **미끄러운 바닥 (미끄러운 바닥):** 진흙이 고여있는 바닥은 미끄러워 작업자에게 넘어짐 사고의 위험을 초래할 수 있습니다. * **정리 정돈 부족 (정리 정돈 부족):** 호스 등이 현장 바닥에 흩어져 있어 작업 공간이 정리되지 않은 상태입니다. **3. 표지판 텍스트 (표지판 텍스트)** * **파일 드라이버 측면 (파일 드라이버 측면):** "제일종합건설 (제일종합건설)"과 전화번호 "010-3639-8559 (010-3639-8559)"가 적혀 있습니다. * **현장 입구 표지판 (현장 입구 표지판):** 표지판 내용이 흐릿하여 정확히 읽기 어렵지만, 안전 관련 경고 문구 또는 현장 정보가 포함되어 있을 가능성이 높습니다. **종합 의견 (종합 의견)** 전반적으로 현장 안전 관리에 개선이 필요한 부분이 있습니다. 특히 안전모 미착용, 미끄러운 바닥, 정리 정돈 부족 등은 사고 발생 가능성을 높이는 요인이므로 즉시 개선해야 합니다. 또한, 모든 작업자가 적절한 안전 장비를 착용하고 안전 수칙을 준수하도록 교육해야 합니다.
gemma3:12b (12.81)	**분석 결과 (Analysis Result)** **1. 객체 설명 (Object Description)** * **작업자 (Workers):** 이미지에서 몇몇 작업자가 보입니다. 그들은 주로 장비 주변에 서 있거나 작업 중인 것으로 보입니다. * **장비 (Equipment):** 이미지에는 드릴링 장비 (드릴링 리그)가 주요 장비로 보입니다. 또한, 트럭과 다른 건설 장비도 배경에 보입니다. * **배경 (Background):** 건설 현장 주변에는 고층 건물과 도로가 보입니다. **2. 안전 위반 사항 (Safety Violations)** * **헬멧 미착용 (Helmet Non-Compliance):** 이미지에서 일부 작업자가 헬멧을 착용하지 않은 것으로 보입니다. 이는 매우 심각한 안전 위반입니다. 건설 현장에서는 모든 작업자가 항상 헬멧을 착용해야 합니다.

SUNG KYUN KWAN UNIVERSITY

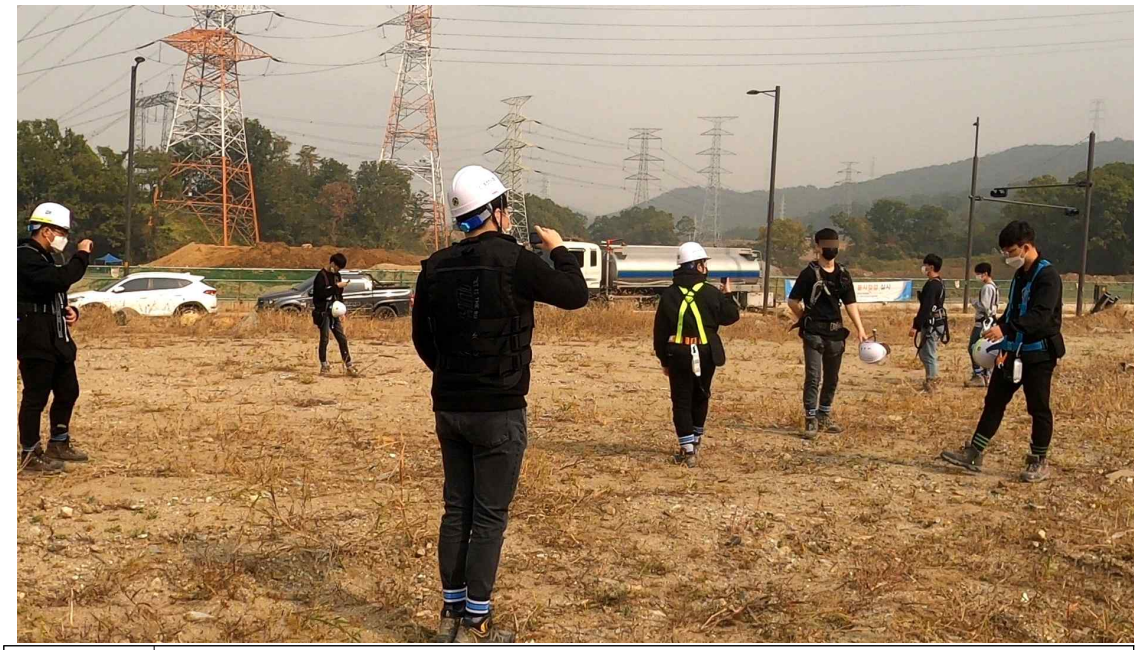
	* **안전벨트 미착용 (Harness Non-Compliance):** 작업자가 높은 곳에서 작업하는 경우 안전벨트 착용이 필수적입니다. 이미지에서 안전벨트 착용 여부를 정확히 판단하기는 어렵지만, 높은 곳에서 작업하는 작업자가 있다면 안전벨트 착용 여부를 확인해야 합니다. * **보호구 미착용 (Lack of PPE):** 작업 환경에 따라 보안경, 장갑, 안전화 등 다른 보호 장비가 필요할 수 있습니다. 이미지에서 이러한 보호 장비 착용 여부를 확인하기는 어렵습니다. **3. 표지판 텍스트 (Sign Text)** 이미지에서 보이는 표지판의 텍스트는 다음과 같습니다. * **"매질콘설팅" (Maejil Consulting):** 이것은 건설 컨설팅 회사 이름으로 보입니다. 전화번호도 함께 표시되어 있습니다. * **"KR-5"**: 장비의 식별 번호로 보입니다. * **다른 표지판**: 다른 표지판들은 가독성이 너무 낮아 정확한 내용을 파악하기 어렵습니다. **결론 (Conclusion)** 이 건설 현장에서는 헬멧 미착용과 같은 심각한 안전 위반 사항이 발생했습니다. 모든 작업자가 안전 규정을 준수하고, 적절한 보호 장비를 착용하는 것이 중요합니다. 또한, 건설 현장 관리자는 안전 점검을 강화하고, 작업자들에게 안전 교육을 제공해야 합니다.
gemma3:4b (22.15)	**분석 결과 (Bunseok Gyeolhwa - Analysis Result)** **1. 대상 (Daedang - Objects)** * **근무자 (Geonmuba):** 작업복을 입은 작업자 1명이 보입니다. (Gyageobeul ibun geonmuba 1myae go bominida.) - One worker wearing work clothes is visible. * **장비 (Jabe):** 크레인 (Keurein - Crane)이 주요 장비입니다. 덤프트럭 (Dumeutreok - Dump Truck)도 보입니다. (Gyodae bominida.) - A crane is the main equipment. A dump truck is also visible. * **작업 환경 (Jageob Hwan-yeong):** 작업 현장에는 물이 고여 있습니다. (Gyageob-e ne mui goyeo itseumnida.) - The construction site has water accumulated. **2. 안전 위반 사항 (Anjeong Wiban Sangeop - Safety Violations)** * **안전모 미착용 (Anjeongmo Migatchyong):** 작업자는 안전모를 착용하지 않고 있습니다. (Gyageobeul chatyonghaji ango itseumnida.) - The

SUNG KYUN KWAN UNIVERSITY

	worker is not wearing a hard hat. * **안전벨트 미착용 (Anjeongbeulte Migatchyong):** 작업자의 안전을 위한 안전벨트 착용이 확인되지 않았습니다. (Gyageobeul anjeong-e watseun anjeongbeulte chatyong-e watneun
llama3.2-vision:11b (3.54)	1. **Objects (Workers, Equipment)** * 2 workers are visible in the image, one in the foreground and one in the background. * A yellow excavator is visible in the background, and a crane is visible in the foreground. * A green bus is parked in the background, and a blue tarp is visible in the foreground. 2. **Safety violations (helmet, harness)** * One worker is not wearing a helmet, and another worker is not wearing a harness. * The worker in the foreground is not wearing a helmet, and the worker in the background is not wearing a harness. 3. **Text on signs** * The sign in the background reads "경
llava:13b (9.82)	안녕하세요! 이 사진은 건설 현장을 보여주는 것으로 보입니다. 1. 물체 (직원, 기계) * 직원: 이미지에서는 직원을 확인하지 않았습니다. 그러나 일반적으로 건설 현장에서는 직원들이 작업을 수행하기 위해 있어야합니다. * 기계: 이미지에서는 기계를 확인하지 않았습니다. 그러나 일반적으로 건설 현장에서는 다양한 기계들이 사용됩니다. 2. 안전 위반 (헬멧, 하르너스) * 헬멧: 이미지에서는 헬멧을 확인하지 않았습니다. 건설 현장에서는 안전상의 헬멧을 착용해야합니다. * 하르너스: 이미지에서는 하르너스를 확인하지 않았습니다. 건설 현장에서는 안전상의 하르너스를 착용해야합니다. 3. 문자(표지) * 이미지에서는 표지나 문자가 보이지 않았습니다. 건설 현장에서는 안전상의 표지나 문자를 따라야합니다. 위의 내용을 바탕으로 이미지에서 발견되는 안전 위반과 관련된 문제들을 설명하였습니다.

표 6 실험1의 화재이미지 결과

SUNG KYUN KWAN UNIVERSITY



gemma3:27b (31.02)	## 건설 현장 안전 분석 (Korean) **1. 객체 (Objects):** * **작업자 (Workers):** 약 8명의 작업자가 보입니다. 대부분 안전모를 착용하고 있지만, 일부 작업자는 안전모 착용 상태가 불분명하거나, 안전 장비 착용 여부가 확인되지 않습니다. * **차량 (Vehicles):** 흰색 승용차, 덤프트럭, 검은색 픽업 트럭이 보입니다. * **전력선 타워 (Power Line Towers):** 현장 주변에 고압 전력선 타워가 여러 개 있습니다. * **기타 (Others):** 안전 표지판, 조명 기둥, 흙바닥, 풀 등이 보입니다. **2. 안전 위반 사항 (Safety Violations):** * **안전모 미착용 또는 부적절한 착용 (Non-compliance with helmet use):** 일부 작업자는 안전모를 착용하지 않았거나, 안전모 착용 상태가 제대로 확인되지 않습니다. * **안전대 미착용 (Lack of harness use):** 현장 상황을 고려할 때, 일부 작업자는 추락 위험에 대비한 안전대 착용이 필요할 수 있습니다. 특히, 고압 전력선 근처에서 작업하는 경우 안전대 착용은 필수적입니다. * **작업복 미착용 (Lack of work clothes):** 일부 작업자는 작업복이 아

SUNG KYUN KWAN UNIVERSITY

	닌 일반 복장을 착용하고 있습니다. 작업복은 작업 환경에 적합한 보호 기능을 제공해야 합니다. * **고압 전력선 근접 작업 (Working near high-voltage power lines):** 고압 전력선 근처에서 작업하는 경우 감전 위험이 매우 높습니다. 안전 거리를 확보하고, 절연 장비를 사용해야 합니다. **3. 표지판 텍스트 (Text on Signs):** * 표지판의 텍스트는 흐릿하게 보여 정확히 읽기 어렵습니다. 하지만, 안전 관련 경고 또는 지시 사항이 적혀 있을 가능성이 높습니다. (표지판의 내용을 명확히 파악하기 위해서는 고해상도 이미지가 필요합니다.) **종합 의견 (Overall Opinion):** 전반적으로 건설 현장의 안전 관리가 미흡한 것으로 보입니다. 작업자들의 안전 장비 착용 상태를 점검하고, 고압 전력선 근접 작업에 대한 안전 대책을 강화해야 합니다. 또한, 안전 교육을 통해 작업자들의 안전 의식을 높이는 것이 중요합니다.
gemma3:12b (12.09)	**1. 객체 설명 (Object Description)** * **작업자 (Workers):** 사진에는 안전모를 착용한 작업자들이 여러 명 보입니다. 대부분 마스크를 착용하고 있습니다. * **장비 (Equipment):** 사진에는 지게차, 트럭, 승용차 등이 보입니다. 또한, 고압 전력탑이 배경에 위치하고 있습니다. * **지형 (Terrain):** 사진은 흙먼지가 많은 공사 현장으로 보입니다. 경사가 있는 지형입니다. **2. 안전 위반 사항 (Safety Violations)** * **안전모 미착용 (Helmet Non-Compliance):** 몇몇 작업자들이 안전모를 착용하지 않았습니다. (안전모를 착용하지 않은 작업자가 있는지 확인해야 합니다.) * **안전 조끼 미착용 (Safety Vest Non-Compliance):** 몇몇 작업자들이 고위험 구역에서 안전 조끼를 착용하지 않았습니다. (안전 조끼를 착용하지 않은 작업자가 있는지 확인해야 합니다.) * **개인 보호 장비 (PPE) 부족 가능성 (Potential Lack of PPE):** 작업 환경을 고려할 때, 먼지 마스크 외 다른 개인 보호 장비 (예: 보안경, 장갑) 착용이 필요할 수 있습니다. * **교통 안전 (Traffic Safety):** 작업자와 차량의 분리 문제가 발생할 수 있습니다. 작업 구역 내 차량 통행 시 작업자의 안전 확보가 중요합니다. **3. 표지판 텍스트 (Sign Text)**

SUNG KYUN KWAN UNIVERSITY

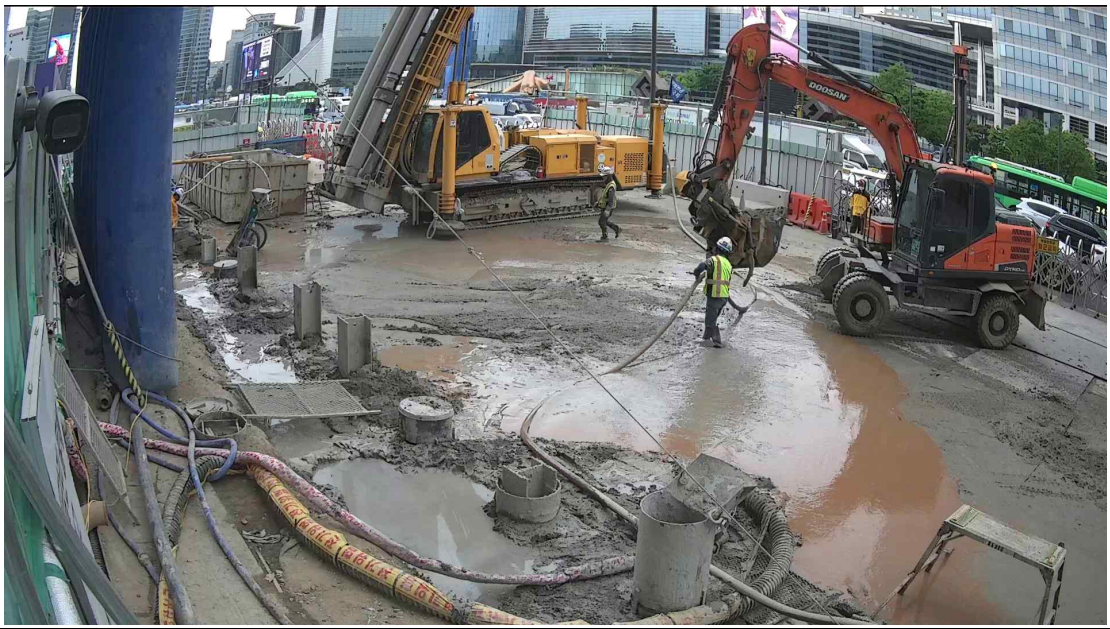
	* 사진에 보이는 표지판의 텍스트는 가독성이 너무 낮아 정확히 파악하기 어렵습니다. (The text on the sign is too difficult to read clearly.) **추가 고려 사항 (Additional Considerations)** * **작업 구역 설정 (Work Area Designation):** 작업 구역과 위험 구역을 명확하게 구분해야 합니다. * **안전 교육 (Safety Education):** 모든 작업자가 안전 교육을 받았는지 확인해야 합니다. * **감독 (Supervision):** 안전 관리자가 현장을 감독하고 안전 수칙 준수를 확인해야 합니다.
gemma3:4b (7.83)	**분석 결과 (Bunseok Gyeolkwang - Analysis Result)** **1. 대상 (Daedang - Objects)** * **근로자 (Geunroja):** 6명의 근로자가 현장(site)을 걸어 이동하고 있습니다. (6myeongui geunroja ga hyeonjang-eul georeo iyonghago itseumnida.) - There are 6 workers walking around the site. * **장비 (Jabi):** * 흰색 덤프트럭 (hwinsa dumpeul-teuk) - 흰색 덤프트럭 (hwinsa dumpeul-teuk) - White dump truck * 검은색 SUV (geomseok SUV) - 검은색 SUV (geomseok SUV) - Black SUV * 작은 흙을 옮기는 장비 (jakuneu heugeul mwoningneun jabi) - Small equipment for moving dirt. * **배경 (Baegyeong):** 산 (san) - 산 (san) - Mountains in the background. **2. 안전 위반 사항 (Anjeong Wiban Sajang - Safety Violations)** * **보안모 미착용 (Boaenmo Mijeogyong):** 일부 근로자가 보안모를 착용하지 않았습니다. (Ilbu geunroja ga boaenmo-reul chatyonghaji anseumnida.) - Some workers are not wearing helmets. * **안전벨트 미착용 (Anjeonbeulleu Mijeogyong):** 근로자 중 일부가 안전벨트를 착용하지 않았습니다. (Geunroja jung ilbu ga anjeonbeulleu-reul chatyonghaji anseumnida.) - Some workers are not wearing safety harnesses. * **안전 교육 부족 (Anjeon Gyoyuk Buchuk):** 전반적으로 안전 교육에 대한 정보가 부족해 보입니다. (Jeonbanryeokjeo anjeon gyoyuk-e daehan jibo-ga bujogehae boimnida.) - Overall, it appears there is a lack of information regarding safety training.

SUNG KYUN KWAN UNIVERSITY

	3. 표지판 텍스트 (Pyooji-pan Teukseu - Text on Signs) * **사진에 보이는 표지판에는 텍스트가 명확하게 보이지 않습니다.** (Sajange bonae-neun pyooji-pan-eun tekseuga myeonghwageo boijanhaetseumnida.) - The text on the signs in the picture is not clearly visible. * **가능성 있는 텍스트 (Ganeungseong inneun tekseu):** "안전" (anjeon) - "Safety" (likely on some signs).
llama3.2-vision:11b (10.08)	1. **Objects (Workers, Equipment)** * Workers: 여러 명의 작업자들이 보도에 서서 작업을 수행하는 것을 볼 수 있다. * Equipment: 작업자들이 사용하는 다양한 장비를 볼 수 가 이른다. 예로, 작업자들은 보도에 가로로 위치한 장비를 사용하는 것을 볼 수 가 이른다. 2. **안전 위안데서 * 헌트: 작업자들은 모두 헌트를 사용하는 것을 볼 수 가 이른다. 헌트는 작업자들의 안전을 보장하는 데 도서 * 보도: 보도는 안전을 유지하는 데 도서 * 장비: 작업자들은 다양한 장비를 사용하는 * 안전 * 3. **문 * “안
llava:13b (6.57)	이미지는 건설 현장에서 일하는 사람들을 보여주고 있습니다. 1. 개인 (직원, 장비) * 직원: 이미지에서는 총 8명의 직원이 등장합니다. 그들은 모두 안전 헬멧을 착용하고 있습니다. * 장비: 이미지에서는 건설 기계나 장비가 보이지 않습니다. 2. 안전 위반 (헬멧, 하르너스) * 헬멧: 모든 직원들은 안전 헬멧을 착용하고 있습니다. * 하르너스: 이미지에서는 하르너스를 사용하는 장면이 보이지 않습니다. 3. 표시된 문자 (게이트, 신호) * 게이트: 이미지에서는 게이트를 사용하는 장면이 보이지 않습니다. * 신호: 이미지에서는 신호를 사용하는 장면이 보이지 않습니다.

표 7 실험1의 안전보호구 이미지 추론결과

SUNG KYUN KWAN UNIVERSITY



gemma3:27b (32.93)	## 건설 현장 안전 분석 (Korean) **1. 객체 (Objects):** * **작업자 (작업원):** 현장 내에 최소 1명의 작업자가 보입니다. 작업자는 진흙탕 안에서 작업하고 있습니다. * **굴착기 (Excavator):** 빨간색 두산 굴착기가 현장 오른쪽에 위치하고 있습니다. * **파일잭 (Pile Driver):** 노란색 파일잭이 현장 중앙에 위치하고 있습니다. * **콘크리트 방호벽 (Concrete Barriers):** 현장 주변에 콘크리트 방호벽이 설치되어 있습니다. * **파이프 및 호스 (Pipes and Hoses):** 현장 곳곳에 파이프와 호스가 흩어져 있습니다. * **맨홀 덮개 (Manhole Covers):** 진흙탕 안에 여러 개의 맨홀 덮개가 보입니다. * **공사 자재 (Construction Materials):** 현장 곳곳에 다양한 공사 자재가 흩어져 있습니다. **2. 안전 위반 사항 (Safety Violations):** * **개인 보호 장비 (PPE):** 작업자는 헬멧을 착용하고 있지만, 진흙탕

SUNG KYUN KWAN UNIVERSITY

	안에서 작업하는 상황을 고려할 때 안전 조끼, 안전화, 방수 장비 등 추가적인 개인 보호 장비가 필요해 보입니다. 특히 미끄러운 환경에서 작업하는 경우, 미끄럼 방지 기능이 있는 신발이 필수적입니다. * **넘어짐 위험 (Fall Hazard):** 진흙탕은 미끄러워 넘어질 위험이 매우 높습니다. 작업자는 넘어지지 않도록 주의해야 하며, 안전 난간이나 작업 발판 등의 안전 시설이 필요합니다. * **장비 주변 안전 (Equipment Safety):** 굴착기와 파일잭 주변에 작업자가 가까이 있는 것은 위험합니다. 장비 작동 시 작업자와의 안전 거리를 확보해야 합니다. * **정리 정돈 (Housekeeping):** 현장 내에 파이프, 호스, 공사 자재 등이 흩어져 있어 작업 공간이 혼잡합니다. 정리 정돈을 통해 작업 공간을 확보하고, 넘어짐 사고를 예방해야 합니다. **3. 표지판 텍스트 (Sign Text):** 이미지에서 명확하게 판독 가능한 표지판 텍스트는 보이지 않습니다. 하지만 현장 주변에 안전 관련 표지판이 설치되어 있을 가능성이 높습니다. **종합 의견 (Overall Opinion):** 전반적으로 현장의 안전 관리가 미흡해 보입니다. 특히 진흙탕 안에서 작업하는 상황은 매우 위험하며, 작업자의 안전을 위해 추가적인 안전 조치가 필요합니다. 작업 공간 정리 정돈, 개인 보호 장비 착용, 장비 주변 안전 확보 등을 통해 사고 예방에 힘써야 합니다.
gemma3:12b (15.79)	**분석 결과 (Analysis Result)** **1. 현장 물체 (Objects on Site)** * **작업자 (Workers):** 사진에 한 명의 작업자가 보입니다. (One worker is visible in the photo.) * **장비 (Equipment):** * 드릴링 장비 (Drilling equipment): 큰 드릴링 장비가 보입니다. (A large drilling machine is visible.) * 굴삭기 (Excavator): 굴삭기가 옆에 주차되어 있습니다. (An excavator is parked beside.) * 파이프 (Pipes): 여러 개의 파이프가 현장에 놓여 있습니다. (Several pipes are lying on the site.) * 콘크리트 블록 (Concrete blocks): 콘크리트 블록들이 쌓여 있습니다. (Concrete blocks are stacked.) * 안전 펜스 (Safety fence): 현장 주변에 안전 펜스가 설치되어 있습니다. (A safety fence is installed around the site.)

SUNG KYUN KWAN UNIVERSITY

	2. 안전 위반 사항 (Safety Violations) * **헬멧 미착용 (Helmet Not Worn):** 작업자가 헬멧을 착용하지 않았습니다. 이는 매우 심각한 안전 위반입니다. (The worker is not wearing a helmet. This is a serious safety violation.) * **안전 장비 부족 (Lack of Safety Gear):** 작업자가 안전 조끼나 기타 개인 보호 장비를 착용하고 있는지 확인하기 어렵습니다. (It's difficult to confirm if the worker is wearing a safety vest or other personal protective equipment.) * **미끄러운 바닥 (Slippery Ground):** 바닥이 젖어 있어 미끄러울 수 있습니다. 작업자는 미끄럼 방지 신발을 착용해야 합니다. (The ground is wet and slippery. Workers should wear slip-resistant shoes.) * **구조물 안전 (Structure Safety):** 주변 구조물에 대한 안전 점검이 필요합니다. (A safety inspection of surrounding structures is needed.) **3. 표지판 텍스트 (Sign Text)** * 사진에서 표지판을 명확하게 읽을 수 없습니다. (The sign text cannot be clearly read from the photo.) 표지판이 있다면, 그 내용이 안전 지침을 따르는지 확인해야 합니다. (If there are signs, you need to verify that they follow safety guidelines.) **결론 (Conclusion)** 이 현장은 심각한 안전 문제를 가지고 있습니다. 작업자가 헬멧을 착용하지 않은 것은 매우 위험하며, 즉시 시정 조치가 필요합니다. 또한, 바닥의 미끄러움, 안전 장비 부족, 주변 구조물 안전 점검 등 다른 안전 문제도 해결해야 합니다. (This site has serious safety issues. The worker not wearing a helmet is very dangerous and requires immediate corrective action. Other safety issues such as the slippery floor, lack of safety gear, and surrounding structure safety inspection also need to be addressed.) **추가 조치 (Additional Actions)** * 작업자 안전 교육 실시 (Conduct worker safety training) * 안전 점검 강화 (Strengthen safety inspections) * 안전 규정 준수 감독 (Supervise compliance with safety regulations) * 작업 환경 개선 (Improve working environment)
gemma3:4b	**이미지 분석 보고서** **1. 대상 (대상물)**

SUNG KYUN KWAN UNIVERSITY

(6.82)	* **노동자:** 한 명의 작업자가 흙으로 덮인 공간에서 작업 중입니다.
	* **장비:**
	  * Doosan 브랜드의 굴착기 (Excavator) 한 대가 눈에 띄게 보입니다.
	  * 콘크리트 파이프 (Concrete Pipe) 여러 개가 땅에 묻혀 있습니다.
	  * 다양한 파이프와 케이블이 땅에 흩어져 있습니다.
	  * 작업 현장 주변에 안전 표지판 및 경고 표시가 있습니다.
	
	2. 안전 위반 사항
	* **헬멧 미착용:** 작업자는 헬멧을 착용하지 않았습니다. (헬멧 착용은 필수 안전 장비입니다.)
	* **안전벨트 미착용:** 작업자는 안전벨트를 착용하지 않았습니다. (낙하물로부터 보호하기 위해 필수입니다.)
	* **안전 표지판 미확인:** 작업자는 현장 안전 표지판을 확인하지 않은 것으로 보입니다. (각 현장의 안전 규정을 준수해야 합니다.)
	* **흙으로 인한 위험:** 흙으로 덮인 공간은 낙하물로 인한 위험이 있습니다.
	
	3. 표지판 내용 (Signs)
	* **Doosan:** 장비 브랜드명입니다.
	* **경고 표지판:** (이미지에서 정확한 문구는 식별하기 어렵지만) “낙하물 주의” 또는 “안전 거리 확보”와 같은 내용의 경고 표지판이 있을 것으로 예상됩니다.
	* **기타 표지판:** (이미지에서 식별 가능한 내용이 제한적입니다.)
	
	전반적인 의견:
	이 현장에서는 헬멧 및 안전벨트 미착용이 심각한 안전 위반 사항입니다. 작업자는 항상 안전 장비를 착용하고, 현장 안전 규정을 준수해야 합니다. 또한, 흙으로 덮인 공간은 낙하물로 인한 위험이 있으므로 주의해야 합니다.

llama3.2-vision:11b (36.27)	1. **Objects (Workers, Equipment)** * Workers: 1 * Equipment: 1 (crane) * Other objects: 1 (concrete mixer) 2. **Safety violations (helmet, harness)** * No safety violations are visible in the image. 3. **Text on signs** * No text on signs is visible in the image.

SUNG KYUN KWAN UNIVERSITY

	Output in Korean 1. **Objects (Workers, Equipment)** * **노작자**: 1 * **기기**: 1 (크라리) * **다음 객체**: 1 (콘크리트 마이서) 2. **안전 위험 (헌트, 하르니) * **안전 위험**은 이미지에 나타이되지 3. **표시의 * **표시의 은 이미지에 나타이되지
llava:13b (4.55)	1. 객체(작업자, 장비) * 작업자: 한 명의 작업자가 포인트에서 일하고 있습니다. * 장비: 반경이 깨진 구멍을 벌리기 위해 사용된 건축 기계가 있습니다. 1. 안전 위반(헬멧, 하르네스) * 안전 위반: 작업자가 헬멧을 쓰지 않고 구멍에서 일하고 있으며, 하르네스도 사용되지 않았습니다. 1. 표지의 문자 * 표지에는 "안전 주의"라는 문구가 있습니다.

표 8 실험1의 중장비 이미지 추론결과

SUNG KYUN KWAN UNIVERSITY

gemma3:27b (32.58)	## 건설 현장 안전 분석 (건설 현장 안전 분석) **1. 객체 (객체)** * **장비 (장비):** 크레인 (크레인), 건설 자재 (건설 자재 - 콘크리트 패널, 벽돌 등), 안전 펜스 (안전 펜스), 팔레트 (팔레트). * **인력 (인력):** 사진에는 작업자가 보이지 않습니다. **2. 안전 위반 사항 (안전 위반 사항)** * **개인 보호 장비 (개인 보호 장비):** 사진에 작업자가 없으므로 헬멧이나 안전대 착용 여부를 확인할 수 없습니다. 하지만, 건설 현장에서는 항상 헬멧, 안전대, 안전화 등 개인 보호 장비를 착용해야 합니다. * **안전 펜스 (안전 펜스):** 안전 펜스가 설치되어 있지만, 펜스 주변에 자재가 쌓여 있어 이동 통로를 막고 있을 가능성이 있습니다. * **자재 적재 (자재 적재):** 팔레트에 쌓인 자재가 불안정하게 쌓여 있을 수 있습니다. 자재가 떨어져 사고를 유발할 수 있습니다. * **통로 확보 (통로 확보):** 안전 펜스 주변과 자재 적재 공간으로 인해 작업자 이동 통로가 충분히 확보되지 않았을 수 있습니다. **3. 표지판 텍스트 (표지판 텍스트)** * **파란색 표지판 (파란색 표지판):** "안전 제일 (안전 제일)" - 안전을

SUNG KYUN KWAN UNIVERSITY

	최우선으로 하라는 의미입니다. * **크레인 주변 표지판 (크레인 주변 표지판):** 크레인 작동 관련 경고 및 안전 수칙이 적혀 있을 것으로 예상됩니다. (자세한 내용은 사진 해상도 때문에 확인하기 어렵습니다.) * **기타 표지판 (기타 표지판):** 크레인 주변에 여러 표지판이 설치되어 있으며, 크레인 작동 시 주의사항, 접근 금지 구역 등을 안내하고 있을 것으로 예상됩니다. **전반적인 의견 (전반적인 의견)** 사진만으로는 정확한 안전 상태를 판단하기 어렵지만, 몇 가지 잠재적인 안전 문제를 발견할 수 있습니다. 작업자 안전을 위해 개인 보호 장비 착용을 철저히 하고, 작업 공간을 정리 정돈하여 통로를 확보하며, 자재 적재 시 안전 수칙을 준수해야 합니다. 또한, 크레인 주변에는 접근을 제한하고, 크레인 작동 시 안전 수칙을 철저히 준수해야 합니다.
gemma3:12b (18.36)	**분석 결과 (Analysis Result)** **1. 객체 (Objects)** * **작업 공간 (Work Area):** 넓은 콘크리트 바닥이 있는 작업 공간입니다. (It's a work area with a large concrete floor.) * **지게차 (Forklift):** 사진 왼쪽 하단에 지게차가 보입니다. (A forklift is visible in the lower left corner of the photo.) * **크레인 (Crane):** 사진 상단에 크레인이 보입니다. (A crane is visible at the top of the photo.) * **합판 (Plywood):** 바닥에 놓여있는 합판 더미가 있습니다. (There is a pile of plywood on the floor.) * **안전 펜스 (Safety Fence):** 작업 공간을 둘러싼 안전 펜스가 있습니다. (There is a safety fence surrounding the work area.) * **안전 표지판 (Safety Signs):** 여러 개의 안전 표지판이 보입니다. (Several safety signs are visible.) * **작업자 (Workers):** 사진에는 작업자가 직접적으로 보이지 않습니다. (Workers are not directly visible in the photo.) **2. 안전 위반 (Safety Violations)** * **작업자 안전 장비 미착용 (Lack of Personal Protective Equipment - PPE):** 사진에 작업자가 보이지 않으므로, 안전모, 안전벨트 착용 여부를 확인할 수 없습니다. 작업자가 있다면, 안전 장비 착용 여부를 확인해야 합니다. (Since workers are not visible in the photo, we cannot confirm whether they are wearing helmets or harnesses. If workers were present,

SUNG KYUN KWAN UNIVERSITY

	their PPE usage would need to be checked.) * **합판 적재 상태 (Plywood Stacking):** 합판 더미가 불안정하게 쌓여 있는 것처럼 보입니다. 넘어질 위험이 있으므로, 안전하게 적재되어 있는지 확인해야 합니다. (The pile of plywood appears to be stacked unstably. The risk of falling needs to be checked.) * **바닥 정리 상태 (Floor Condition):** 바닥에 합판 조각이나 잔해물이 있으면 미끄러짐 사고의 위험이 있습니다. (Debris on the floor poses a slip hazard.) **3. 표지판 텍스트 (Sign Text - Approximate Translations)** * **왼쪽 상단 표지판 (Top Left Sign):** "안전 제일" (Safety First) - 일반적인 안전 경고 표지판입니다. (A general safety warning sign.) * **오른쪽 상단 표지판 (Top Right Sign):** "작업 중" (Under Construction/Work in Progress) - 작업 중임을 알리는 표지판입니다. (A sign indicating work in progress.) * **안전 펜스 표지판 (Safety Fence Sign):** "안전 관리자" (Safety Manager) - 안전 관리자의 연락처 정보가 있을 가능성이 있습니다. (Likely contains contact information for the safety manager.) * **왼쪽 펜스 표지판 (Left Fence Sign):** "안전 관리자" (Safety Manager) - 안전 관리자의 연락처 정보가 있을 가능성이 있습니다. (Likely contains contact information for the safety manager.) **추가적인 안전 점검 사항 (Additional Safety Check Points)** * 지게차 운전자의 안전 교육 이수 여부 (Verification of forklift operator's safety training) * 크레인 작업자의 자격증 보유 여부 (Verification of crane operator's license) * 작업 공간의 조명 상태 (Adequacy of lighting in the work area) * 비상 상황 발생 시 대피 경로 확보 여부 (Ensuring escape routes are clear in case of emergency)

gemma3:4b
 	
(7.66)
	사진 분석 결과
	
	1. 현장 요소 (작업자, 장비)
	* **작업자:** 사진에서 작업자는 보이지 않지만, 작업 공간에 작업 자재가 놓여 있어 작업이 진행 중임을 알 수 있습니다.
	* **장비:**
	  * 크레인: 사진 오른쪽에 보이는 크레인이 현장의 주요 장비입니다.
	  * 콘크리트 믹서: 크레인 옆에 콘크리트 믹서가 있습니다.

SUNG KYUN KWAN UNIVERSITY

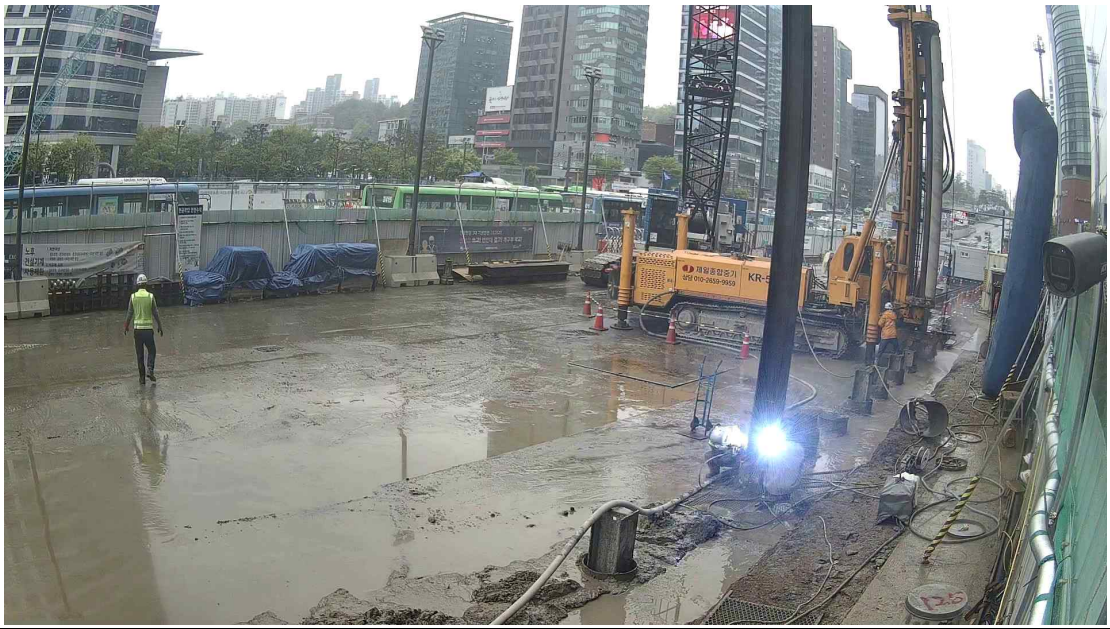
	* 콘크리트 블록: 바닥에 놓여 있는 콘크리트 블록은 작업 자재입니다. **2. 안전 위반 사항** * **헬멧 미착용:** 작업자들의 머리에 헬멧을 착용하지 않았습니다. 헬멧 착용은 필수 안전 장비입니다. * **안전벨트 미착용:** 작업자들의 몸에 안전벨트를 착용하지 않았습니다. 안전벨트 착용은 낙하물로부터 보호하기 위한 필수 장비입니다. * **보호 장비 부족:** 전반적으로 작업자들의 안전 장비 착용이 부족한 것으로 보입니다. **3. 표지판 내용** * **표지판 1:** "위험! 작업 중" (위험! 작업 중) - 작업 중임을 알리는 표지판입니다. * **표지판 2:** "안전모 착용 필수" (안전모 착용 필수) - 안전모 착용의 중요성을 강조하는 표지판입니다. * **표지판 3:** "안전벨트 착용 필수" (안전벨트 착용 필수) - 안전벨트 착용의 중요성을 강조하는 표지판입니다. * **표지판 4:** "주의! 낙하물 주의" (주의! 낙하물 주의) - 낙하물로부터 보호하기 위한 경고 표지판입니다. **전반적인 평가:** 현장에는 안전 표지판이 설치되어 있지만, 작업자들의 안전 장비 착용이 미흡하여 안전사고 발생 위험이 있습니다. 즉시 작업자들에게 안전 수칙을 강조하고, 모든 작업자들이 안전 장비를 착용하도록 해야 합니다.
llama3.2-vision:11b (35.66)	Based on the image, here are the safety issues I've identified: **1. Objects (Workers, Equipment)** * There are no workers or equipment visible in the image. The site appears to be empty. **2. Safety Violations (Helmet, Harness)** * There are no workers on the site, so there are no safety violations to report. * However, I did not see any workers or equipment, so I couldn't check for any safety equipment like helmets or harnesses. **3. Text on Signs** * The signs on the site are in Korean, but I can't read them. The text

SUNG KYUN KWAN UNIVERSITY

	on the signs appears to be: + "C + "C I couldn't read I
llava:13b (8.06)	안녕하세요! 이 건설 현장의 사진을 분석해보겠습니다. 1. 물체 (직원, 장비) * 현장에는 인공지능 기계와 인부가 있습니다. 이들은 건설 과정에서 필요한 역할을 수행하고 있습니다. * 인부는 노동자로, 현장에서 일을 수행하고 있습니다. 1. 안전위반 (헬멧, 해럴라이스) * 현장에서는 헬멧과 해럴라이스를 적절하게 사용하지 않고 있습니다. 이는 안전위반으로, 실제로 사용하면 충분한 보호를 제공할 수 없습니다. 1. 표지에 있는 문자 * 현장에서는 안전과 관련된 정보를 제공하는 표지가 있습니다. 그러나 이 표지의 내용은 알 수 없습니다. 위의 분석을 바탕으로, 현장에서는 안전과 관련된 문제가 있어 보이고, 추가적인 안전 조치를 취하여야 합니다.

표 9 실험1의 표지판 이미지 추론결과

SUNG KYUN KWAN UNIVERSITY



gemma3:27b (31.02)	## 건설 현장 안전 분석 (건설 현장 안전 분석) **1. 객체 (객체)** * **작업자 (작업자):** 현장 내에 최소 2명의 작업자가 보입니다. 한 명은 용접 작업을 하고 있고, 다른 한 명은 주변에서 지켜보고 있습니다. * **장비 (장비):** * **파일 드라이버 (파일 드라이버):** 큰 노란색 파일 드라이버 장비가 보입니다. * **용접기 (용접기):** 용접 작업자가 사용하고 있는 용접기. * **호스 (호스):** 파일 드라이버 주변에 여러 개의 호스가 놓여 있습니다. * **컨테이너 (컨테이너):** 현장 주변에 컨테이너 박스가 여러 개 있습니다. * **안전 펜스 (안전 펜스):** 현장 주변을 둘러싼 녹색 안전 펜스. * **기타 (기타):** 진흙이 고여있는 현장 바닥, 건물, 도로 등. **2. 안전 위반 사항 (안전 위반 사항)** * **안전모 미착용 (안전모 미착용):** 용접 작업을 하고 있는 작업자는 안전모를 착용하고 있지만, 주변에서 지켜보고 있는 작업자는 안전모를 착용하지 않고 있습니다. 이는 안전 규정 위반입니다.

SUNG KYUN KWAN UNIVERSITY

	* **안전 장비 부족 (안전 장비 부족):** 용접 작업자는 용접 마스크를 착용하고 있지만, 다른 안전 장비 (보안경, 장갑 등) 착용 여부는 확인하기 어렵습니다. * **미끄러운 바닥 (미끄러운 바닥):** 진흙이 고여있는 바닥은 미끄러워 작업자에게 넘어짐 사고의 위험을 초래할 수 있습니다. * **정리 정돈 부족 (정리 정돈 부족):** 호스 등이 현장 바닥에 흩어져 있어 작업 공간이 정리되지 않은 상태입니다. **3. 표지판 텍스트 (표지판 텍스트)** * **파일 드라이버 측면 (파일 드라이버 측면):** "제일종합건설 (제일종합건설)"과 전화번호 "010-3639-8559 (010-3639-8559)"가 적혀 있습니다. * **현장 입구 표지판 (현장 입구 표지판):** 표지판 내용이 흐릿하여 정확히 읽기 어렵지만, 안전 관련 경고 문구 또는 현장 정보가 포함되어 있을 가능성이 높습니다. **종합 의견 (종합 의견)** 전반적으로 현장 안전 관리에 개선이 필요한 부분이 있습니다. 특히 안전모 미착용, 미끄러운 바닥, 정리 정돈 부족 등은 사고 발생 가능성을 높이는 요인이므로 즉시 개선해야 합니다. 또한, 모든 작업자가 적절한 안전 장비를 착용하고 안전 수칙을 준수하도록 교육해야 합니다.
gemma3:12b (12.09)	**분석 결과 (Analysis Result)** **1. 객체 설명 (Object Description)** * **작업자 (Workers):** 이미지에서 몇몇 작업자가 보입니다. 그들은 주로 장비 주변에 서 있거나 작업 중인 것으로 보입니다. * **장비 (Equipment):** 이미지에는 드릴링 장비 (드릴링 리그)가 주요 장비로 보입니다. 또한, 트럭과 다른 건설 장비도 배경에 보입니다. * **배경 (Background):** 건설 현장 주변에는 고층 건물과 도로가 보입니다. **2. 안전 위반 사항 (Safety Violations)** * **헬멧 미착용 (Helmet Non-Compliance):** 이미지에서 일부 작업자가 헬멧을 착용하지 않은 것으로 보입니다. 이는 매우 심각한 안전 위반입니다. 건설 현장에서는 모든 작업자가 항상 헬멧을 착용해야 합니다. * **안전벨트 미착용 (Harness Non-Compliance):** 작업자가 높은 곳에서 작업하는 경우 안전벨트 착용이 필수적입니다. 이미지에서 안전벨트 착용 여부를 정확히 판단하기는 어렵지만, 높은 곳에서 작업하는 작업자가 있다면 안전벨트 착용 여부를 확인해야 합니다.

SUNG KYUN KWAN UNIVERSITY

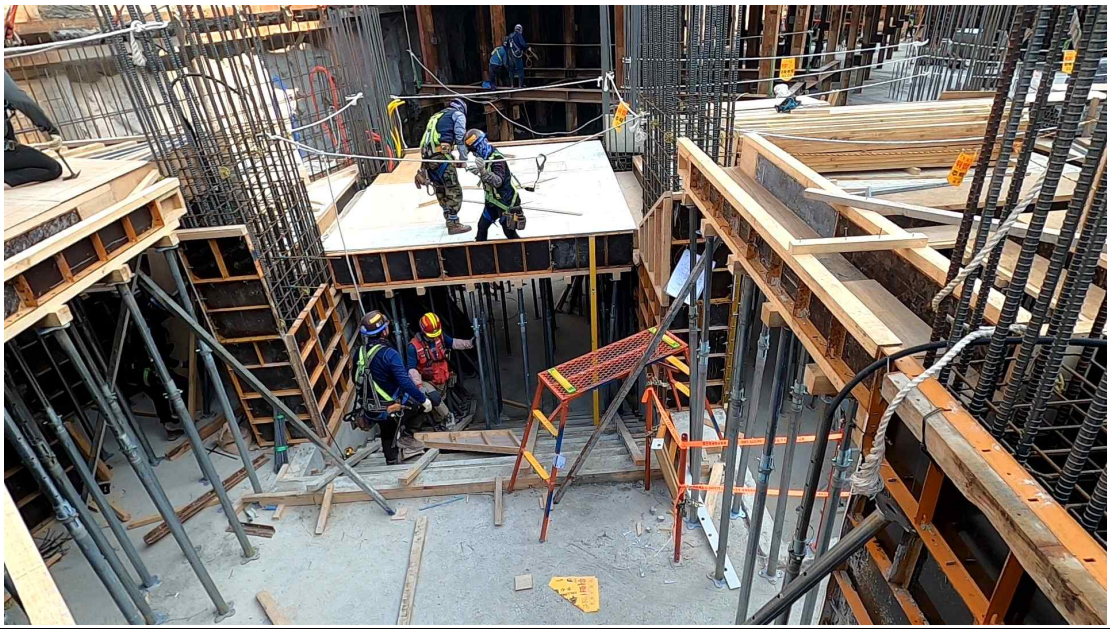
	* **보호구 미착용 (Lack of PPE):** 작업 환경에 따라 보안경, 장갑, 안전화 등 다른 보호 장비가 필요할 수 있습니다. 이미지에서 이러한 보호 장비 착용 여부를 확인하기는 어렵습니다. **3. 표지판 텍스트 (Sign Text)** 이미지에서 보이는 표지판의 텍스트는 다음과 같습니다. * **"매질콘설팅" (Maejil Consulting):** 이것은 건설 컨설팅 회사 이름으로 보입니다. 전화번호도 함께 표시되어 있습니다. * **"KR-5"**: 장비의 식별 번호로 보입니다. * **다른 표지판**: 다른 표지판들은 가독성이 너무 낮아 정확한 내용을 파악하기 어렵습니다.
gemma3:4b (7.83)	**분석 결과 (Bunseok Gyeolhwa - Analysis Result)** **1. 대상 (Daedang - Objects)** * **근무자 (Geonmuba):** 작업복을 입은 작업자 1명이 보입니다. (Gyageobeul ibun geonmuba 1myae go bominida.) - One worker wearing work clothes is visible. * **장비 (Jabe):** 크레인 (Keurein - Crane)이 주요 장비입니다. 덤프트럭 (Dumeutreok - Dump Truck)도 보입니다. (Gyodae bominida.) - A crane is the main equipment. A dump truck is also visible. * **작업 환경 (Jageob Hwan-yeong):** 작업 현장에는 물이 고여 있습니다. (Gyageob-e ne mui goyeo itseumnida.) - The construction site has water accumulated. **2. 안전 위반 사항 (Anjeong Wiban Sangeop - Safety Violations)** * **안전모 미착용 (Anjeongmo Migatchyong):** 작업자는 안전모를 착용하지 않고 있습니다. (Gyageobeul chatyonghaji ango itseumnida.) - The worker is not wearing a hard hat. * **안전벨트 미착용 (Anjeongbeulte Migatchyong):** 작업자의 안전을 위한 안전벨트 착용이 확인되지 않았습니다. (Gyageobeul anjeong-e watseun anjeongbeulte chatyong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun gyeong-e watneun
llama3.2-vision:11b	1. **Objects (Workers, Equipment)** * 2 workers are visible in the image, one in the foreground and one in the background. * A yellow excavator is visible in the background, and a crane is visible

SUNG KYUN KWAN UNIVERSITY

(10.08)	in the foreground. * A green bus is parked in the background, and a blue tarp is visible in the foreground. 2. **Safety violations (helmet, harness)** * One worker is not wearing a helmet, and another worker is not wearing a harness. * The worker in the foreground is not wearing a helmet, and the worker in the background is not wearing a harness. 3. **Text on signs** * The sign in the background reads "경
llava:13b (6.57)	안녕하세요! 이 사진은 건설 현장을 보여주는 것으로 보입니다. 1. 물체 (직원, 기계) * 직원: 이미지에서는 직원을 확인하지 않았습니다. 그러나 일반적으로 건설 현장에서는 직원들이 작업을 수행하기 위해 있어야합니다. * 기계: 이미지에서는 기계를 확인하지 않았습니다. 그러나 일반적으로 건설 현장에서는 다양한 기계들이 사용됩니다. 2. 안전 위반 (헬멧, 하르너스) * 헬멧: 이미지에서는 헬멧을 확인하지 않았습니다. 건설 현장에서는 안전상의 헬멧을 착용해야합니다. * 하르너스: 이미지에서는 하르너스를 확인하지 않았습니다. 건설 현장에서는 안전상의 하르너스를 착용해야합니다. 3. 문자(표지) * 이미지에서는 표지나 문자가 보이지 않았습니다. 건설 현장에서는 안전상의 표지나 문자를 따라야합니다. 위의 내용을 바탕으로 이미지에서 발견되는 안전 위반과 관련된 문제들을 설명하였습니다.

표 10 실험1의 화재 이미지 추론결과

SUNG KYUN KWAN UNIVERSITY



gemma3:27b (31.02)	## 건설 현장 안전 분석 (Korean) **1. 객체 (Objects):** * **작업자 (Workers):** 최소 3명의 작업자가 보입니다. 한 명은 아래쪽 계단에서, 다른 두 명은 위쪽 목재 플랫폼에서 작업하고 있습니다. * **장비 (Equipment):** * 목재 틀 (Wooden formwork): 콘크리트 타설을 위한 틀로 보입니다. * 철근 (Rebar): 콘크리트 보강을 위한 철근이 많이 보입니다. * 계단 (Stairs): 작업자들이 오르내리는 임시 계단이 설치되어 있습니다. * 안전망 (Safety Net): 일부 구역에 안전망이 설치되어 있습니다. * 경고 테이프 (Caution Tape): 일부 구역에 경고 테이프가 설치되어 있습니다. * 기타: 헬멧, 안전 조끼, 안전 벨트 등 개인 보호 장비가 보입니다. **2. 안전 위반 사항 (Safety Violations):** * **안전모 (Helmets):** 모든 작업자가 안전모를 착용하고 있지만, 일부 작업자의 안전모가 제대로 고정되지 않았거나 손상된 것으로 보입니다. * **안전 벨트 (Harnesses):** 일부 작업자는 안전 벨트를 착용하고 있지만, 모든 작업자가 착용하고 있는지 확인하기 어렵습니다. 특히 위쪽 플랫폼

SUNG KYUN KWAN UNIVERSITY

	에서 작업하는 작업자들의 안전 벨트 연결 상태를 확인해야 합니다. * **작업 공간 정리 (Workplace Organization):** 작업 공간이 다소 어수선하며, 자재들이 흩어져 있어 걸려 넘어질 위험이 있습니다. * **추락 방지 (Fall Protection):** 위쪽 플랫폼의 난간이 충분히 높지 않거나 부실해 보입니다. 추락 방지 시설을 강화해야 합니다. * **계단 안전 (Stair Safety):** 계단에 안전 난간이 부족하거나 손상된 부분이 있을 수 있습니다. **3. 표지판 텍스트 (Sign Text):** * 이미지 품질이 낮아 표지판의 텍스트를 명확하게 읽을 수 없습니다. 하지만, 경고 또는 안전 관련 표지판으로 추정됩니다. 표지판의 내용을 확인하여 안전 수칙을 준수해야 합니다. **전반적인 의견:** 전반적으로 건설 현장의 안전 관리가 미흡한 부분이 있습니다. 작업자들의 개인 보호 장비 착용 상태를 확인하고, 작업 공간을 정리하며, 추락 방지 시설을 강화하는 등 안전 관리를 철저히 해야 합니다. 또한, 표지판의 내용을 확인하여 안전 수칙을 준수해야 합니다.
gemma3:12b (12.09)	**1. 객체 (Workers, Equipment) - Objects (Workers, Equipment)** * **작업자 (Workers):** 사진에는 세 명의 작업자가 보입니다. 두 명은 높은 곳에서 작업하고 있고, 한 명은 아래에서 작업하고 있습니다. * (세 명의 작업자가 보입니다. 두 명은 높은 곳에서 작업하고 있고, 한 명은 아래에서 작업하고 있습니다.) * **장비 (Equipment):** * **거푸집 (Formwork):** 콘크리트 타설을 위한 거푸집 구조물이 보입니다. * **사다리 (Ladder):** 작업자들이 올라가기 위해 사다리가 사용되고 있습니다. * **안전 로프 (Safety Rope):** 높은 곳에서 작업하는 작업자가 안전 로프를 사용하고 있습니다. * **철근 (Rebar):** 콘크리트 구조물을 짓기 위한 철근이 보입니다. * **도구 (Tools):** 작업자들이 사용하는 도구들이 보입니다. **2. 안전 위반 (Safety Violations)** * **헬멧 미착용 (Helmet Non-Compliance):** 사진에서 보이는 작업자 중 한 명이 헬멧을 착용하지 않았습니다. 이는 매우 심각한 안전 위반입니다. * (작업자 중 한 명이 헬멧을 착용하지 않았습니다. 이는 매우 심각한 안전 위반입니다.)

SUNG KYUN KWAN UNIVERSITY

	* **안전 장비 미착용 (Lack of Safety Gear):** 높은 곳에서 작업하는 작업자 중 한 명이 안전 벨트(안전 하네스)를 착용하지 않았습니다. 이는 추락 사고의 위험을 크게 증가시킵니다. * (높은 곳에서 작업하는 작업자 중 한 명이 안전 벨트(안전 하네스)를 착용하지 않았습니다. 이는 추락 사고의 위험을 크게 증가시킵니다.) * **사다리 안전 (Ladder Safety):** 사다리 주변에 장애물이 있는지, 사다리가 안정적으로 설치되었는지 확인해야 합니다. 사진만으로는 판단하기 어렵지만, 주의가 필요합니다. * (사다리 주변에 장애물이 있는지, 사다리가 안정적으로 설치되었는지 확인해야 합니다. 사진만으로는 판단하기 어렵지만, 주의가 필요합니다.) **3. 표지판 텍스트 (Text on Signs)** * 사진에 보이는 표지판의 텍스트는 읽기 어렵습니다. 하지만, 안전 경고 표지판으로 보입니다. * (사진에 보이는 표지판의 텍스트는 읽기 어렵습니다. 하지만, 안전 경고 표지판으로 보입니다.) **전반적인 안전 문제 (Overall Safety Concerns)** 전반적으로, 이 건설 현장은 안전에 대한 심각한 문제가 있습니다. 헬멧 미착용과 안전 하네스 미착용은 즉시 해결해야 할 문제입니다. 또한, 작업 환경 전반에 걸쳐 안전 점검을 강화하고, 작업자들에게 안전 교육을 실시해야 합니다.
gemma3:4b (7.83)	**건설 현장 사진 분석 (건설 현장 사진 분석)** **1. 객체 (Objects):** * **근로자 (Geunroja - Workers):** 3명의 근로자가 사진 속에 보입니다. 모두 작업복을 입고 있으며, 작업 중인 것으로 보입니다. * **장비 (Jabe - Equipment):** * 콘크리트 폼 (Konkeuri konpyom - Formwork): 콘크리트를 부어 넣기 위한 나무 폼으로 만들어진 계단과 벽체. * 철근 (Cheulgeun - Reinforcement Bars): 콘크리트 내부에 보강을 위해 사용된 철근. * 안전 울타리 (Jeonangeultari - Safety Barriers): 작업 구역을 보호하기 위한 울타리. * 계단 (Gyodan - Stairs): 작업자들이 이동하기 위한 계단. * 작업 플랫폼 (Jageop Pulmot - Work Platform): 작업 공간을 제공하는 플랫폼.

SUNG KYUN KWAN UNIVERSITY

	2. 안전 위반 사항 (Anjeong Wiban Sajang - Safety Violations): * **안전모 미착용 (Anjeonmo Migatyong - Helmet Not Worn):** 두 명의 근로자가 안전모를 착용하지 않고 있습니다. 이는 가장 심각한 안전 위반 사항입니다. * **체인 하네스 미착용 (Chein Hanessu Migatyong - Harness Not Worn):** 한 명의 근로자가 체인 하네스를 착용하지 않고 있습니다. 높은 곳에서 작업할 때 하네스를 착용하는 것은 필수적입니다. * **안전 울타리 미사용 (Anjeon Ultaeri Miswamyong - Safety Barriers Not Used):** 계단 측면에 안전 울타리가 설치되어 있지만, 일부 구간에서는 안전 울타리가 설치되어 있지 않아 위험합니다. * **작업 플랫폼 안전 (Jageop Pulmot Anjeong - Work Platform Safety):** 작업 플랫폼의 안전 상태가 명확하지 않습니다. 미끄럼 방지 매트가 있는지 확인해야 합니다. **3. 표지판 텍스트 (Jeobjipan Text - Sign Text):** * **주황색 표지판 (Juhwangsaek Pyeojidan - Orange Sign):** "안전" (Anjeon - Safety) 이라고 쓰여 있습니다. * **노란색 표지판 (Nurangwaek Pyeojidan - Yellow Sign):** "위험" (Wiheom - Danger) 이라고 쓰여 있습니다. * **기타 표지판 (Gita Pyeojidan - Other Signs):** 다른 표지판의 내용은 사진에서 명확하게 읽을 수 없습니다. **전반적인 평가 (Jeonbanjeogeo Pyeongga - Overall Assessment):** 이 건설 현장은 여러 가지 심각한 안전 위반 사항을 가지고 있습니다. 안전모와 체인 하네스 착용은 필수이며, 모든 작업 구역에 안전 울타리가 설치되어 있어야 합니다. 작업 환경의 안전을 확보하기 위해 즉각적인 조치가 필요합니다.
llama3.2-vision:11b (10.08)	**1. Objects (Workers, Equipment)** * A worker is seen in the center of the image, wearing a blue long-sleeve shirt, a yellow and green harness, and a hard hat. He is crouing on a wooden platform, working on a metal structure. * There are several metal beams and steel on the platform. * A yellow is seen on the right side of the platform. **2. Safety violations (helmet, harness)** * The worker is wearing a hard hat, which is a good for .

SUNG KYUN KWAN UNIVERSITY

	* The worker is wearing a , which is a for . * The worker is not wearing , which is for . **3. Text on signs** * The on the reads , which means .
llava:13b (6.57)	1. 객체(직원, 장비): 이미지에서는 두 명의 직원과 다양한 장비가 확인됩니다. 직원들은 건축 현장에서 일하고 있으며, 장비로는 드롭 프레이, 패널트 척 테일, 그리고 다양한 글라스 부품들이 보인합니다. 2. 안전 위반(헬멧, 하르네스): 이미지에서는 안전 위반으로 건축 현장에서 일하는 사람들의 헬멧과 하르네스가 확인되지 않습니다. 이는 안전 규정을 위반한 상태입니다. 3. 표시된 문자: 이미지에서는 표시된 문자가 없습니다.

표 11 실험1의 추락 이미지 추론결과

2.5 상세 분석 및 토의

모델 크기(Size)와 추론 품질의 트레이드오프 실험 결과, 모델의 파라미터 크기와 안전 분석의 품질 사이에는 뚜렷한 정비례 관계가 확인되었다. Gemma-3:27B 모델은 평균 31.48초의 긴 추론 시간이 소요되었으나, "추락 방지 장비 미착용"과 같은 전문 용어를 정확한 한국어로 구사하며 가장 상세한 리포트를 생성하였다. 반면, 경량화 모델인 Gemma-3:4B는 6.81초대의 빠른 속도를 보였으나, "Anjeongmo Mijiitnyong(안전모 미착용)"과 같이 한국어를 발음대로 영문 표기하거나 문맥을 단순화하는 경향을 보여, 법적 효력이 필요한 리포트 작성에는 부적합한 것으로 판명되었다.

언어적 장벽과 도메인 특화 성능 측면에서, Llama-3.2-Vision은 객체 인식 속도(10.39초)는 준수하였으나, 답변의 주 언어가 영어로 고정되거나 한국어 답변 시 부자연스러운 번역투가 나타나는 한계를 보였는데 건설 현장의 안전 리포트는 작업자와 관리자가 직관적으로 이해할 수 있는 한국어로 작성하는 것이 반드시 필요한 기

SUNG KYUN KWAN UNIVERSITY

능이므로 한국어 데이터 처리 능력이 강화된 Gemma-3 계열의 모델이 Llama 계열보다 본 연구의 목적에 부합한다고 판단된다.

결론적으로 실시간성이 중요한 '감지' 단계가 아닌, 정확성이 요구되는 '심층 분석' 단계에서는 추론 속도가 다소 느리더라도(약 31초/건) Gemma-3:27B와 같은 고성능 모델을 채택하는 것이 시스템의 전체적인 신뢰도(Reliability)를 담보하는 필수 조건임을 확인하였다.

SUNG KYUN KWAN UNIVERSITY

3. 실험 2: 논리 추론용 sLLM 성능 평가 및 최적 엔진 선정

3.1 배경 및 문제 정의

시각 언어 모델(VLM)이 현장 상황을 묘사(Captioning)한 이후, 이를 바탕으로 산업안전보건법에 근거한 구체적인 위험성 평가와 안전 수칙을 도출하기 위해서는 고도의 논리적 추론 능력을 갖춘 별도의 언어 모델이 요구되지만 수천억 파라미터 규모의 LLM을 엣지 디바이스에서 구동하는 것은 메모리와 연산 속도 측면에서 불가능하므로 제한된 GPU 자원 하에서도 '실시간성(Throughput)'과 '추론의 깊이(Reasoning Capability)'라는 두 가지 상충되는 요구사항을 동시에 만족시킬 수 있는, 최적화된 소형 언어 모델(sLLM)을 선별하는 과정이 필수적이다.

3.2 실험 목적

본 실험은 대표적인 로컬 소형 언어 모델(sLLM)들의 추론 속도와 효율성을 비교 분석하여, 단일 GPU 환경에서 멀티모달 파이프라인을 안정적으로 구동할 수 있는 최적의 추론 엔진을 선정하는 것을 목적으로 한다.

3.3 실험 설계 및 방법

+ 비교 대상:
  - Granite 3.3 (8B): IBM의 엔터프라이즈급 경량 모델
  - GPT-OSS (20B): 추론(Thinking) 능력이 강화된 중형 모델
  - Llama 3.1 (8B): 범용적인 성능을 갖춘 표준 베이스라인
  - Gemma-3n (2B) / Phi-4 / Qwen 3: 기타 비교군
+ 실험 조건: 동일한 하드웨어(RTX A6000) 및 Ollama 플랫폼 환경에서 구동

SUNG KYUN KWAN UNIVERSITY

+ 평가 지표:

- 처리량(Throughput): 초당 토큰 처리 수(Tokens Per Second, TPS), 5회 평균
- PROMPT: "건설 현장에서 작업자가 안전모를 착용하지 않았을 때 발생할 수 있는 위험에 대해 전문가적 입장에서 논리적으로 3줄 요약을 해줘."

3.4 실험 결과 요약

텍스트 기반의 논리적 추론을 담당할 sLLM과 서빙 플랫폼의 효율성을 검증하기 위해, Ollama 환경에서 다양한 모델의 초당 토큰 처리량(TPS)을 측정하였다.

모델	평균 TPS	응답 결과	비고
Granite 3.3 (8B)	94.66	건설 현장에서는 작업자가 안전모를 착용하지 않으면 심각한 부상 및 사망 위험이 증가합니다. 이는 떨어지는 물체, 전기적 충격, 또는 극단적인 경우 환경 요인(예: 태양 빛)로 인한 손상을 받을 수 있기 때문입니다. 안전모 착용은 이러한 위험을 최소화하고 작업자의 생명과 건강을 보호하는 기본적인 안전 조치입니다.	최고 속도
GPT-OSS (20B)	92.13	1. 안전모 미착용은 낙하물·파편·낙석에 의한 두개골 골절, 뇌진탕 등 심각한 외상 위험을 크게 증가시킨다. 2. 전기 설비 근처에서 작업 시 접지되지 않은 금속 부품에 충격을 받아 감전 사고가 발생할 수 있으며, 이는 사망에 이를 수 있다. 3. 또한, 안전모 부재는 현장 안전 규정 위반으로 인한 법적 책임과 보험 청구 제한, 기업 이미지 손상 등 조직 차원의 재정적·법적 위험을 초래한다.	CoT 가능

SUNG KYUN KWAN UNIVERSITY

모델	평균 TPS	응답 결과	비고
Llama 3.1 (8B)	80.00	건설 현장에서 작업자가 안전모를 착용하지 않은 경우, 다음과 같은 위험이 발생할 수 있습니다. 1. **신체 부상**: 작업자의 머리를 보호하지 못한 채로 작업하면 충돌, 떨어짐, 또는 물리적 충격에 의한 머리 부상이 발생할 수 있습니다. 2. **안전장비 미착용으로 인한 사고**: 안전모를 착용하지 않으면 작업자가 다른 안전장비도 착용하지 않을 가능성이 높아져, 더 많은 위험을 초래할 수 있습니다. 3. **법적 책임**: 작업자가 안전모를 착용하지 않은 경우, 부상이나 사고가 발생한 경우 법적으로 책임이 지워질 수 있으며, 이는 개인의 재산과 평판에 큰 영향을 미칠 수 있습니다.	
Gemma-3n (2B)	82.04	## 건설 현장 작업자 안전모 미착용 시 발생 가능한 위험 (전문가적 요약) 1. **심각한 머리 부상 위험 증가:** 건설 현장은 고소 작업, 중장비 충돌, 낙하물 등 머리에 직접적인 충격 위험이 높습니다. 안전모 미착용 시, 이러한 위험은 극대화되어 심각한 뇌진탕, 외상성 뇌출혈, 심지어 사망까지 이어질 수 있습니다. 2. **작업 효율 및 안전 관리 비용 증대:** 안전모 미착용은 작업 중 집중력 저하, 판단력 흐림을 유발하여 사고 발생 가능성을 높입니다. 이는 작업 지연, 추가적인 안전 조치 필요, 사고 발생 시 치료비 증가 등 직접적인 비용 부담으로 이어지며, 장기적으로는 생산성 저하를 초래합니다. 3. **법적 책임 및 기업 이미지 실추:** 안전모 미착용은 산업안전보건법 위반에 해당하며, 법적 처벌을 받을 수 있습니다. 또한, 안전 불감증으로 인식될 경우 기업의 안전 관리 시스템에 대한 신뢰를 잃게 하고, 소비자 및 투자자에게 부정적인 이미지를 심어줄 수 있습니다. **참고:** 안전모는 머리 부상으로부터 보호하는 필수적인 안전 장비이며, 작업자의 안전과 기업의 책임을 위해 반드시 착용해야 합니다.	

SUNG KYUN KWAN UNIVERSITY

모델	평균 TPS	응답 결과	비고
Phi-4 (Reasoning)	49.89	Line 1: ""안전모를 착용하지 않으면 떨어지는 물체나 갑작스러운 충격으로 인해 심각한 두부손상의 위험이 큽니다."" Line 2: ""추락이나 부딪침 사고로 생명을 위협받으며, 안전관리 규정 위반으로 법적 처벌도 받을 수 있습니다."" Line 3: ""따라서, 모든 작업자는 필수적인 안전장비를 착용하여 위험을 최소화해야 합니다.""	CoT 반복, 많은 토큰 사용
Qwen 3 (32B)	18.01	1. 안전모 미착용 시 추락물, 장비 충돌, 낙하 등 현장 내 다양한 외부 충격으로 인해 두부 외상 및 치명적 손상 위험이 급격히 증가합니다. 2. 뇌진탕, 척추 손상, 의식 상실 등 즉각적 생명 위협은 물론, 장기적인 신경학적 후유증으로 작업자의 삶의 질과 경제적 안정에 심각한 영향을 미칠 수 있습니다. 3. 산업안전보건법 제14조에 따라 개인 보호구 착용은 의무사항으로, 미준수 시 사업장에 과태료 부과 및 법적 책임 소재가 발생하며, 안전 문화 약화로 사고 확산 리스크가 높아집니다.	과도한 리소스 점유

표 12 실험2(모델별 벤치마크)의 결과

3.5 상세 분석 및 토의

일반적으로 언어 모델은 파라미터 규모가 증가할수록 추론 성능은 향상되는 반면 처리 속도는 저하되는 경향을 나타내는데 본 실험에서 GPT-OSS 모델은 약 20B 규모의 중형 추론 특화 모델임에도 불구하고 92.13 TPS(Token Per Second)라는 매우 높은 처리 속도를 기록하였는데 이것은 8B급 경량 모델인 Llama 3.1의 처리 속도(80.00 TPS)를 상회하는 수치로 최신 아키텍처와 추론 최적화 기법이 하드웨어 가속 효율을 극대화한 결과로 해석되며 이러한 특성은 복잡한 사고 과정을 요구하는 안전 법규 해석과 같은 논리 추론이 중심이 되는 작업에서 GPT-OSS가 매우 효율적인 추론 엔진으로 활용될 수 있음을 시사한다.

반면, Qwen3 32B 모델은 18.01 TPS로 측정되어 높은 추론 성능을 가지고 있음에

SUNG KYUN KWAN UNIVERSITY

도 불구하고 준실시간 처리가 요구되는 환경에는 부적합한 특성을 보였는데 대규모 파라미터와 상대적으로 무거운 추론 경로방식으로 인한 연산 부담에 기인한 것으로 판단된다.

Granite 3.3 모델은 85.08 TPS를 기록하며 실험군 내에서 가장 우수한 처리 속도를 보였으나, 직선적인 추론 구조로 인해 복잡한 다단계 논리를 전개하는 데는 표현력의 한계가 관찰되었다. Phi-4 모델의 경우 명시적인 추론 과정(Chain of Thought)을 수행하도록 설계된 탓에, 실제 응답 지연(Latency)이 증가하고 토큰 소모량이 상대적으로 높은 것으로 나타났다. 이는 연산 자원이 한정된 환경에서는 효율성 저하로 이어질 수 있음을 시사한다.

한편, Gemma-3n(2B) 모델은 경량 구조임에도 82.04 TPS라는 높은 처리 속도를 달성하여 단순 추론 작업에는 효과적일 것으로 판단되나, 2B 파라미터가 갖는 태생적인 한계로 인해 복잡한 법규 해석이나 장문 논리 추론에서는 정확도가 다소 미흡할 수 있음이 확인되었다.

종합적으로 볼 때, 본 연구에서 구성한 오프라인 분석 파이프라인의 논리 추론 엔진으로는 GPT-OSS가 추론 능력과 처리 속도 간의 균형 측면에서 가장 우수한 성능을 보이는 모델로 판단된다.

SUNG KYUN KWAN UNIVERSITY

4. 실험 3: 운용 경제성 분석

4.1 배경 및 문제 정의

통상적인 AI 시스템 구축 방식인 클라우드 API(SaaS) 모델은 초기 도입 장벽이 낮으나, 24시간 영상 데이터를 처리해야 하는 안전 모니터링 시스템의 특성상 운영 비용의 불확실성이 매우 크다. 본 연구에서 제안하고자 하는 온프레미스 구축 방식이 실무적으로 수용되기 위해서는, 초기 하드웨어 투자 비용을 상쇄하고도 남는 장기적 비용 효율성이 시뮬레이션을 통해 확인하려고 한다.

4.2 실험 목적

본 실험은 가상의 건설 현장 운용 시나리오를 수립하고, 상용 클라우드 서비스 이용 시와 제안하는 온프레미스 시스템 구축 시의 총소유비용을 비교 시뮬레이션한다. 이를 통해 초기 투자 비용이 회수되는 손익분기점(Break-Even Point, BEP)을 도출하여 온프레미스 도입의 경제적 타당성을 규명하고자 한다.

4.3 실험 모델 설계

+ 운용 시나리오: 12채널 CCTV 운영, 일일 총 500건의 이벤트 분석 발생 가정.
+ 데이터 처리량 산정: 1건당 입력 2400토큰(로그 및 프롬프트), 출력 600토큰(리포트) 등 총 3,000토큰 소요 기준.
+ 비교군 설정:
  - Case A (Public Cloud 도입시 운용 비용): OpenAI GPT-4o API 공시 요금 ($5.00/Input 1M, $15.00/Output 1M) 적용.
  - Case B (온프레미스 P사에서 운용중인 서버 비용): 워크스테이션 구축비(935만 원) 및 월 산업용 전기료(5만 원) 적용.

SUNG KYUN KWAN UNIVERSITY

4.4 실험 결과 및 토의

몬테카를로 시뮬레이션(n=5,000) 결과, 온프레미스 시스템은 초기 구축비용에도 불구하고, 통신비와 API 비용을 절감하여 아래 그림과 같이 운용 7개월 차에 BEP을 돌파하는 것으로 나타났다. 특히, PoC 현장에서 영상 데이터 업로드 비용이 포함된 시나리오에서 온프레미스 방식이 클라우드 대비 99% 이상의 확률로 경제적 우위를 점함을 확인하였다.

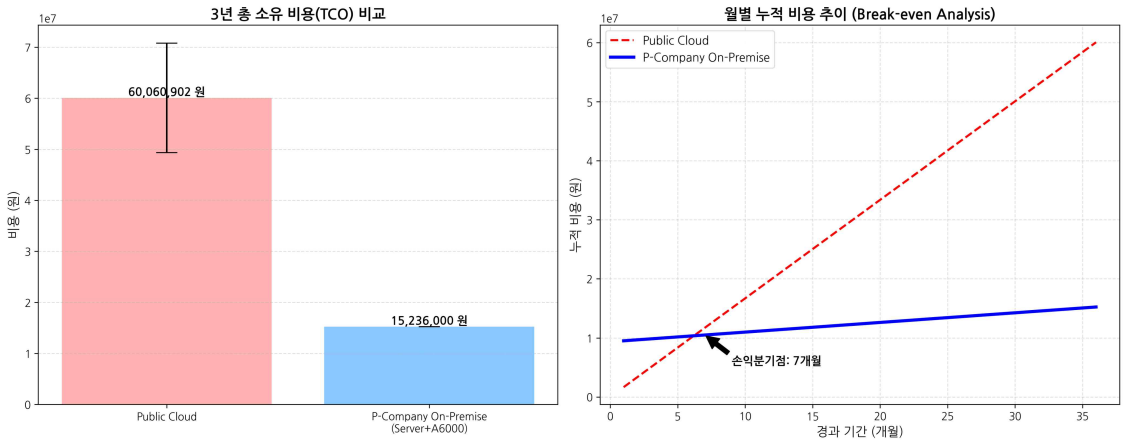


그림 11 실험 3의 결과

SUNG KYUN KWAN UNIVERSITY

5. 소결

본 장에서 수행한 실험 결과를 종합할 때, 차기 챕터에서 구현할 시스템의 설계 요구사항을 다음과 같이 구체화할 수 있다. 먼저, 시각적 정황 파악과 한국어 묘사 능력에서 강점을 보인 Gemma-3 27B와, 효율적인 논리 추론이 가능한 GPT-OSS 20B를 조합하여 AI 엔진을 구성하는 것이 타당해 보인다. 운용 플랫폼 측면에서는 VRAM 자원 관리의 효율성과 다양한 양자화 모델 지원이 가능한 Ollama 아키텍처를 기반으로 하되, 고성능 모델의 추론 지연을 상쇄하기 위해 '주간 즉시 알람(YOLO)'와 '야간 심층 분석(VLM/sLLM)'으로 이원화된 하이브리드 파이프라인 구조를 최종적으로 채택하고자 한다.

SUNG KYUN KWAN UNIVERSITY

제4장 인공지능 안전인지 프레임워크 구현

1. 시스템 아키텍처 및 구현 환경

본 연구에서 제안하는 '지능형 안전 인지 프레임워크(Smart Safety Perception Framework)'는 건설 현장이 직면한 세 가지의 내재적 제약으로 ① 통신 인프라의 불안정성, ② 데이터 보안의 특수성, ③ 엣지 컴퓨팅 자원의 물리적 한계를 극복하는 데 주안점을 두고 설계되었다. 이를 위해 마이크로서비스 아키텍처(Micro services Architecture, MSA)를 기반으로 구축하였고 실시간성이 요구되는 위험 감지 프로세스와 고비용 연산을 수반하는 맥락 추론 프로세스를 시공간적으로 분리하는 전략을 취하였다.

1.1 아키텍처 설계 원칙 및 데이터 흐름

시스템의 논리적 구조는 데이터 처리의 시의성과 분석 깊이에 따라 '주간 실시간 추론'과 '야간 심층 추론'으로 구분되며, 각 계층은 비동기 메시지 브로커를 매개로 연결된다. 첫째로 주간 실시간 추론은 다채널 CCTV로부터 유입되는 대용량 비디오 스트림을 실시간으로 처리하는 역할을 수행하는 것으로 볼 수 있다. 해당 계층에서는 FFMPEG 기반의 저지연 파이프라인이 RTSP 스트림을 디코딩하고, NVIDIA Triton Inference Server가 객체 탐지(Object Detection) 모델을 병렬로 수행하여 1차적인 위험 이벤트를 필터링하는 기능을 담당한다. 이때, 단순한 스냅샷 저장이 아닌 사고 전후의 인과관계를 보존하기 위해 순환 버퍼(Circular Buffer) 메커니즘이 적용되며, 트리거된 시점의 전후 맥락을 포함한 비디오 클립이 생성되어 저장소에 적재된다.

SUNG KYUN KWAN UNIVERSITY

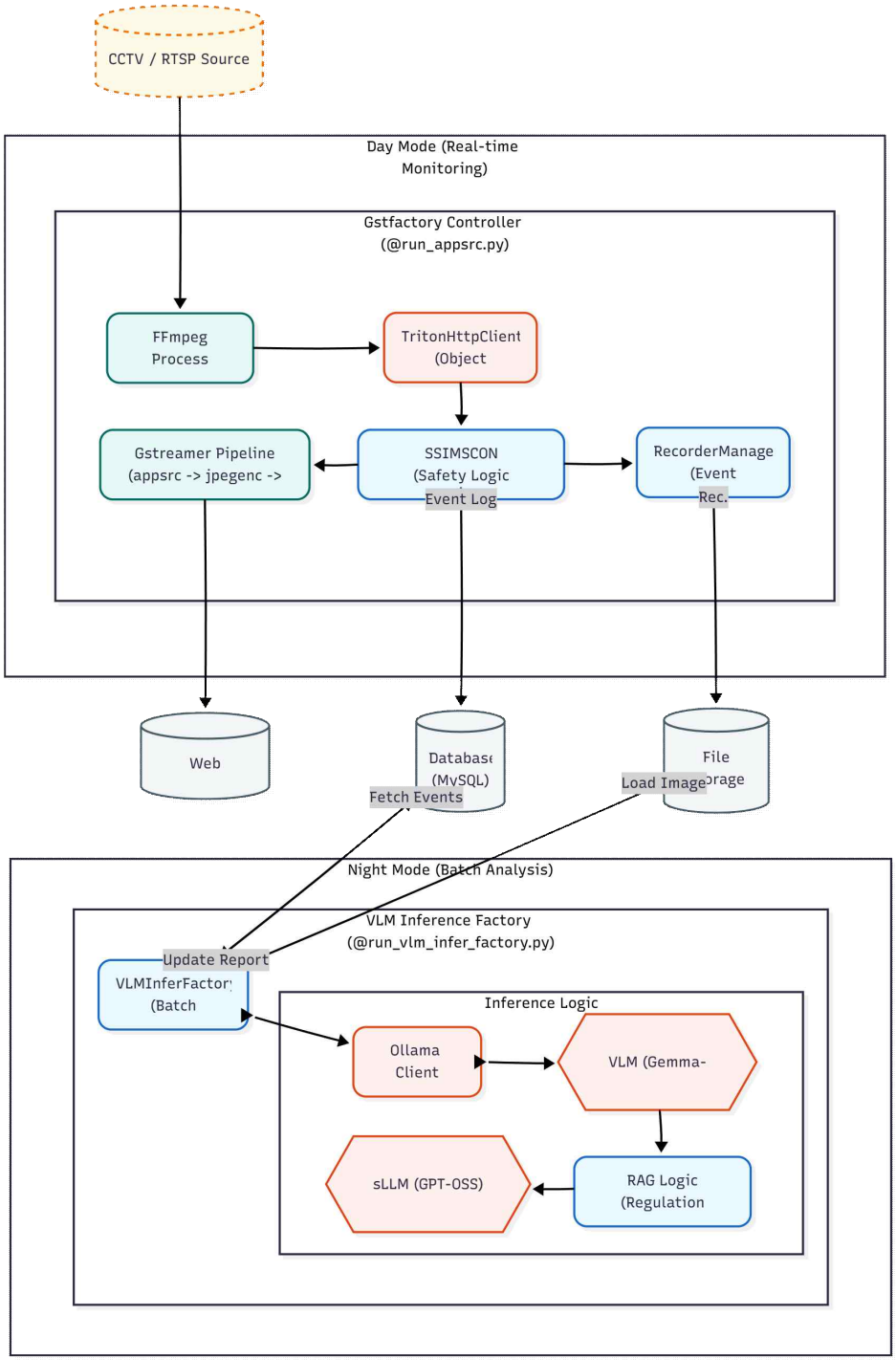


그림 12 구현된 시스템 아키텍처

SUNG KYUN KWAN UNIVERSITY

둘째, 야간 심층 분석 계층은 시스템 리소스의 여유가 확보되는 야간 또는 유휴 시간에 가동되는 특징을 갖는다. 해당 계층은 급변하는 오픈소스 AI 생태계에 대응하기 위해 특정 모델에 종속되지 않고 공개된 다양한 언어 모델군을 로딩하여 운용할 수 있는 플러그인(Plug-in) 구조를 갖는다. 로컬 LLM 런타임인 Ollama가 미들웨어로 활용하며 여기에 탑재되는 모델군은 시각 정보와 텍스트를 통합하는 비전 통합 모델군(Qwen2-VL, Gemma3 등)과 텍스트 기반의 고차원 추론을을 수행하는 소형 언어 모델군(sLLM: Qwen3, Phi-4, GPT-OSS 등)으로 분류된다. 시스템은 데이터 특성에 따라 모델을 인스턴스화하여 추론 파이프라인을 구성하며, 이는 객체 나열을 넘어 작업자의 안전수칙을 위반하게 된 경위와 잠재적 위험성을 구조화된 텍스트로 변환하는 과정이다.

셋째, 지식 통합 및 피드백 계층은 분석된 위험 시나리오와 법적 기준을 결합하는 최종 단계에 해당한다. 벡터 데이터베이스(Vector DB)에 구축된 산업안전보건법 기반의 지식베이스는 RAG 기술을 통해 추론 결과와 결합하여 환각 현상(Hallucination)이 억제된 안전 리포트의 신뢰성을 확보하기 위한 구성이며, 모든 데이터 흐름은 건설 현장 내부의 온프레미스 환경 내에서 처리되어 외부 클라우드 의존성을 배제한 형태로 운용된다.

1.2 하드웨어 및 소프트웨어 구현 환경

제안된 아키텍처의 실무적 적용 가능성을 가늠하기 위한 실험 환경은 현장 구축이 가능한 규격의 단일 워크스테이션을 기반으로 조성되었다. 해당 컴퓨팅 환경은 우분투 운영체제를 기저로 하여, 다양한 양자화 모델의 실험적 적용과 배포가 용이하도록 Ollama 및 Triton 추론 서버를 연동한 런타임 구조를 갖추고 있으며, 이에 대한 구체적인 하드웨어 및 소프트웨어 명세는 하단의 표와 같이 정리할 수 있다.

SUNG KYUN KWAN UNIVERSITY

분류	항목	명세 및 버전
Computing Resource	Processor	Intel Core i9-13900K (24 Cores)
	Accelerator	NVIDIA GeForce RTX 4090 (24GB)
	Memory	DDR5 64GB (5600MHz)
	Storage	NVMe SSD 2TB (Gen4)
Software Stack	OS	Ubuntu 22.04 LTS
	Runtime	Python 3.10, Docker 24.0.5
	Middleware	Redis 7.2.4, Celery 5.3.6
Detection AI Framework	Runtime	Tritonserver (25.05)
	Detection Engine	Ultralytics YOLO11
LLM AI Framework	Runtime	Ollama v0.13.0
	Vision Engine	Gemma-3 (27B)
	Reasoning Engine	GPT-OSS (20B)

표 13 워크스테이션 기반 실시간 CCTV 영상감지 환경

하드웨어 구성에 있어 RTX 4090 GPU를 선정한 것은, 24GB의 VRAM이 4-bit 양자화된 7B~9B급 sLLM과 실시간 탐지 모델을 동시에 메모리에 상주시킬 수 있는 최소한의 임계점이기 때문이며, 여러 개의 객체탐지 모델을 병렬로 구동하기 위한 램상주(Batch) 모델에서도 충분한 메모리가 필요하다. 또한, 소프트웨어 측면에서는 Python 기반의 FastAPI 백엔드와 Celery 워커를 통해 비동기 처리 파이프라인을 구축함으로써, 단일 GPU 자원에 대한 경합을 스케줄링 레벨에서 제어하도록 구현하였다.

SUNG KYUN KWAN UNIVERSITY

2. 실시간 이벤트 감지 및 데이터 큐레이션 파이프라인[주간]

주간 운영 단계(Online Phase)의 핵심 목표는 연속적인 비디오 스트림으로부터 객체 탐지 모델을 통하여 위험 상황으로 인지되는 여러 시점을 추출하여 큐레이션 하는 것이며 본 연구에서는 저지연 스트리밍 인터페이스 그리고 순환 버퍼링 메커니즘이 유기적으로 결합된 고속 데이터 파이프라인을 구현하였다.

2.1 저지연 비디오 획득 및 적응형 연결 제어

건설 현장의 CCTV 인프라가 사용하는 RTSP(Real-Time Streaming Protocol)에 기반을 두고 있으나 내부 네트워크 환경에서 패킷 손실(Packet Loss)과 지터(Jitter) 등의 불안정성을 보일 수 있으므로 FFmpeg 디코더 기반의 비디오 획득 인터페이스를 기반으로 서브프로세스를 통해 제어되며, '-flags low_delay' 및 '-buffer_size' 파라미터 설정을 통해 입력 지연을 억제하는 방식으로 동작한다. 또한 네트워크 단절 상황 시에는 지수 백오프(Exponential Backoff) 알고리즘 기반의 자동 재연결 로직이 작동하도록 구현하였으며 이는 재시도 간격을 조절하여 일시적인 네트워크 장애시 워크스테이션의 자원 소모를 관리하고 연결 복구 시 신속히 스트림을 재개하는 운영 전략을 구축하였다.

2.2 순환 버퍼링을 통한 문맥 데이터(Context Data) 확보

사고 발생 순간의 단편적 스냅샷은 사고의 원인과 과정을 설명하는 데 제한적이므로 순환 버퍼(Circular Buffer)를 이용하여 구현되었는데 메모리 상의 큐(Queue/Deque)는 최근 150초 분량의 원본 영상을 링 버퍼(Ring Buffer) 구조로 유지하며, 안전모 미착용 등의 이벤트 트리거 발생 시 버퍼의 과거 데이터(Pre-event)와 이후 데이터(Post-event)를 병합하여 비디오 클립으로 저장한다.

SUNG KYUN KWAN UNIVERSITY

3. 온프레미스 멀티모달 심층 분석 파이프라인[야간]

주간 운영 단계에서 큐레이션된 비디오 클립 데이터는 야간 또는 유휴 시간에 심층 분석 단계로 정해진 시각에 자동으로 실행하도록 되어 있으며 잠재적 위험을 규명하는 과정으로 구성되며, 분산 작업 스케줄링 시스템과 다단계 심층 추론 엔진이 온프레미스 환경에서 운용된다.

3.1 분산 작업 스케줄링 및 자원 할당 최적화

온프레미스 GPU 자원을 활용한 대량 비디오 처리를 위해 비동기형 분산 처리 아키텍처가 적용되었다. 인메모리 데이터 브로커(Redis)는 주간에 적재된 분석 요청 대기열(Queue)을 관리하고, Celery 워커(Worker) 프로세스가 이를 비동기적으로 처리하는 구조이다.

3.2 로컬 LLM 서빙 프레임워크 및 양자화

시스템은 Ollama 프레임워크 기반의 로컬 LLM 서빙 환경을 갖추고 있어서 단일 워크스테이션에서의 모델 운용을 위해 모델 가중치(Weights)를 4-bit 정수형으로 압축하는 양자화(Quantization) 기술이 적용되었고 이는 FP16(Half-precision) 대비 메모리 점유율을 낮추면서도 건설 안전 도메인에 필요한 추론 성능을 유지하여 엣지 환경에서의 운용성을 지원하는 방식이다.

3.3 다단계 심층 분석 파이프라인 (Multi-stage Deep Analysis Pipeline)

오프라인 분석 공정은 정량적 데이터 추출과 정성적 맥락 해석이 결합된 2단계 파이프라인으로 구성되며, 객체 탐지 기술과 VLM/sLLM이 상호 보완하는 구조를 취한다.

1단계(데이터 로드 및 정황 묘사): 분석은 주간에 축적된 정량적 데이터(DB 로그)

SUNG KYUN KWAN UNIVERSITY

와 미디어 파일(영상/스냅샷)을 로드하며 시작되며 VLMInferFactory 모듈은 이벤트 영상에서 키프레임을 추출하여 Gemma-3 27B (VLM)에 입력한 후 VLM은 이를 바탕으로 기존 객체 탐지 모델의 수치적 로그로는 표현하기 어려운 행동적·상황적 맥락을 텍스트 형태로 기술(Captioning)하는 역할을 수행한다.

2단계(법규 기반 종합 리포팅): 이어지는 단계에서는 시각적 묘사와 정량적 로그를 융합하여 결론을 도출하게 되는데 GPT-OSS (sLLM)는 통합 데이터를 기반으로 사건의 인과관계를 추론하며, RAG 모듈이 추출한 산업안전보건법 등 관련 법규를 적용하여 안전 리포트를 생성한다.

SUNG KYUN KWAN UNIVERSITY

4. 시스템 구현 상세

본 절에서는 앞서 설계한 아키텍처를 실제 소프트웨어로 구현한 핵심 기술 요소를 기술한다. 시스템은 Python 기반의 마이크로서비스로 구성되며, 각 모듈은 Docker 컨테이너로 패키징되어 배포 유연성을 확보하였다.

4.1 GStreamer 기반 고속 파이프라인 구현

실시간 영상 데이터의 고속 처리를 보장하기 위해 FFMpeg와 GStreamer를 융합한 하이브리드 파이프라인이 구축되었다. 구체적으로 'run_appsrc.py' 모듈은 'appsrc' 플러그인을 활용하여 메모리에 상주하는 영상 스트림 데이터를 파이프라인으로 직접 주입하는 역할을 수행하며, 'queue' 요소를 전략적으로 배치한 것은 데이터 처리 과정에서 발생할 수 있는 병목 현상을 구조적으로 완화하기 위한 조치로 해석할 수 있다.

Pipeline Structure: appsrc → queue → videoconvert → jpegenc → multifilesink

4.2 Redis-Celery 기반 비동기 작업 오케스트레이션

시스템의 확장성과 안정성을 위해 Redis를 메시지 브로커로, Celery를 분산 작업 큐로 활용하는 비동기 아키텍처를 구현하였다.

- Task Queue: tasks.py에 정의된 live_celery 태스크는 실시간 스트리밍 및 1차 감지를 담당하며, test_celery 태스크는 오프라인 심층 분석을 수행한다.
- State Management: 각 CCTV의 스트리밍 상태와 이벤트 쿨다운(Cooldown) 정보는 Redis의 Key-Value 저장소에 실시간으로 동기화되어, 다중 프로세스 간의 상태 불일치(Race Condition)를 방지한다.

SUNG KYUN KWAN UNIVERSITY

4.3 오프라인 분석 엔진 구현 (VLMInferFactory)

심층 분석 작업의 효율적인 수행을 위해 'VLMInferFactory' 클래스가 핵심 컴포넌트로 설계되었다. 본 모듈은 주간에 축적된 이벤트 로그(이벤트DB)와 영상 데이터(/data/playlist/accident/)를 입력으로 받아들이며, 저장된 스냅샷과 메타데이터를 별도의 재탐지 과정 없이 즉시 VLM/sLLM 파이프라인으로 주입하도록 설계되었는데 이는 대규모 야간 분석 작업의 처리 속도를 제고하기 위한 최적화 기제로 작용한다고 볼 수 있다.

4.4 RAG 기반 문서 탐지 및 법규 매칭

본 시스템에 구현된 RAG 모듈은 단순한 질의응답을 넘어, 상황에 적합한 안전 문서를 능동적으로 탐지(Document Detection)하는 기능적 특성을 갖는다. 벡터 저장소(Vector Store)의 경우 산업안전보건법 등 방대한 안전 관련 문서를 청크 단위로 분할하여 벡터화하는 과정을 거쳤다. 또한 문맥 인식 검색(Context-Aware Retrieval) 단계에서 VLM이 생성한 상황 묘사 텍스트를 쿼리로 변환하고 코사인 유사도에 기반하여 상위 법규 조항을 추출하는 방식은, 답변의 연관성을 확보하고 환각 현상을 억제하는 기술적 기제로 작용한다고 평가할 수 있다.

SUNG KYUN KWAN UNIVERSITY

5. 현장 적용 및 운영 결과

본 연구에서 제안한 인공지능 프레임워크 기반의 지능형 안전 관제 시스템의 실효성을 검증하기 위해, 실제 대규모 토목 공사 현장에 시스템을 배포하고 이에 대한 운영 결과를 분석하였다. 이 과정은 크게 현장 맞춤형 시스템 구축 및 인터페이스 통합 단계와, 수집된 데이터를 기반으로 한 심층 분석 및 그 효용성 검증 단계로 구분된다. 특히, 기존 컴퓨터 비전 기술의 한계로 지적되어 온 '맥락 부재로 인한 오알람(False Alarm)' 문제를 본 시스템이 얼마나 효과적으로 해결할 수 있는지에 중점을 두고 실증을 진행하였다.

5.1 실증 현장 특성 및 관제 시스템 구축

본 연구의 현장 적용성 검증은 2025년 10월, 국내 D사의 공사 현장에서 수행되었다. 굴착기, 덤프트럭, 콘크리트 펌프카 등 다종의 중장비와 작업 인력이 혼재된 해당 현장은 사고 발생 위험도가 높은 복합 공정 구간으로 분류된다. 공정 진행에 따라 작업 구간이 빈번하게 변경되는 특성을 고려하여 고정형 CCTV 대신 총 5채널의 이동식 CCTV가 운용되었는데, 이는 설치 위치와 촬영 각도(FoV)의 지속적인 변화로 인해 배경 모델링이 어렵다는 기술적 난이도를 수반한다. 이에 본 연구에서는 동일한 공정 조건 하에서 비교적 안정적인 영상 확보가 가능했던 2025년 10월 25일부터 26일까지 양일간의 데이터를 핵심 분석 대상으로 선정하였다.

시스템 운용 환경 측면에서는 현장의 열악한 네트워크 사정을 극복하고 데이터 보안을 강화하기 위해, 현장 관리 사무소 내에 독립적인 AI 추론서버 환경을 구축하였다. 5채널의 CCTV 영상은 외부망과 분리된 폐쇄형 로컬 네트워크를 통해 추론서버로 전송되어, 영상 데이터의 병목 현상 없이 실시간 추론이 가능하도록 설계되었다.

SUNG KYUN KWAN UNIVERSITY

그림 13 플랫폼내 AI 탐지 태스크

AI추론서버에 탑재된 1차 객체 탐지 모델(주간)은 건설 현장의 3대 핵심 위험 요소인 화재(Fire), 안전 장비 착용 상태(PPE), 중장비 접근(Heavy Equipment)을 실시간으로 감지하는 모델을 탑재하였다.

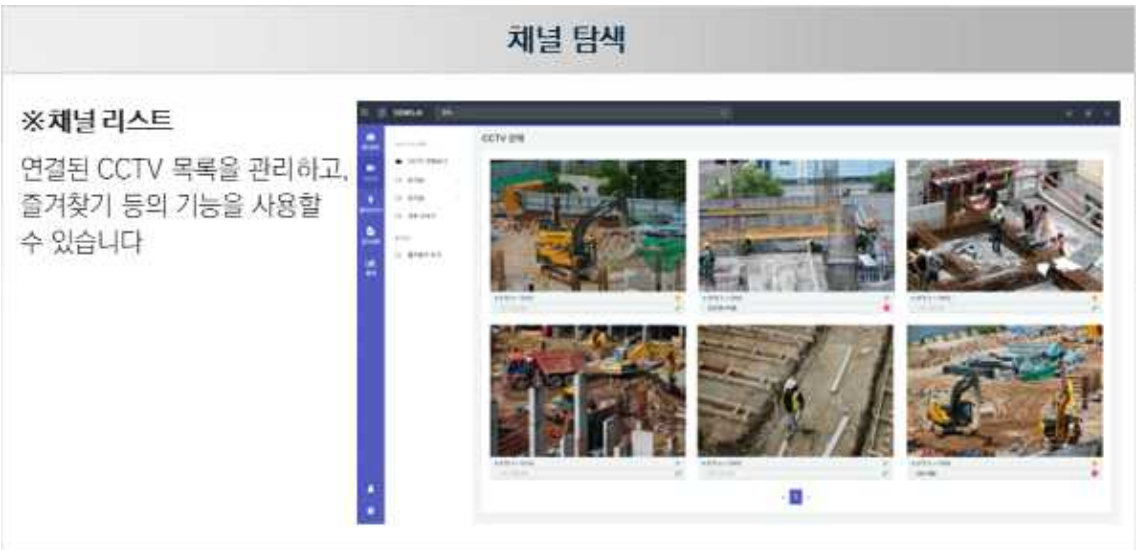


그림 14 실시간 감시 기능

또한 현장 안전 관리자가 복잡한 AI 모델을 직관적으로 통제할 수 있도록 웹 기반

SUNG KYUN KWAN UNIVERSITY

통합 관제 인터페이스가 개발되어 적용되었다. FastAPI 백엔드를 기반으로 구축된 이 인터페이스는 GStreamer 파이프라인과 연동되어 실시간 스트리밍 서비스를 제공하도록 설계되었다. 이를 통해 관리자는 5개 채널의 영상을 단일 화면에서 모니터링할 수 있으며, 위험 이벤트 발생 시 통계 기능을 통해 해당 시점의 영상과 분석 결과를 즉시 확인할 수 있는데, 이는 기존의 단순 모니터링 방식을 데이터 기반의 능동적 관제 체계로 전환하는 기능을 수행한다고 할 수 있다.

그림 15 저장된 감지 내역 목록

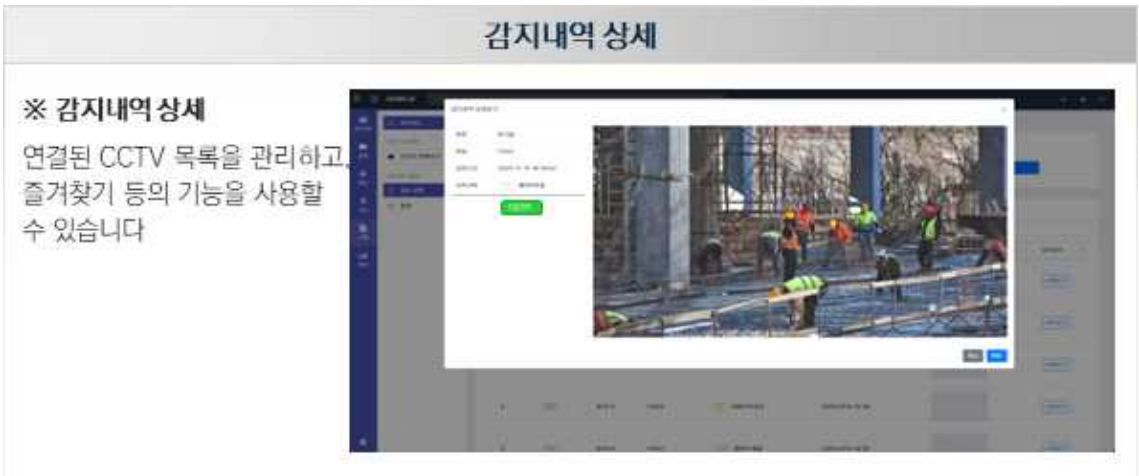


그림 16 감지내역 상세 확인

SUNG KYUN KWAN UNIVERSITY

5.2 온프레미스 프레임워크 심층 분석 적용결과[야간]

제안된 프레임워크의 핵심인 '맥락 인식(Context Awareness)' 성능을 검증하기 위해, 25일과 26일 양일간 현장에서 수집된 이벤트 전체를 대상으로 야간 배치 처리를 통한 오프라인 심층 분석을 수행하였다. 분석은 1차 객체 탐지 모델이 '위험'으로 판단하여 로깅한 이벤트를 입력으로 하여, VLM 파이프라인을 거쳐 최종 검증하는 방식으로 진행되었다.

해당 기간 동안 총 631건의 잠재적 위험 이벤트가 1차 모델에 의해 포착되었고, 본 시스템의 판단 결과와 인간 전문가의 정답(GT)을 대조 분석하였다. 그 결과는 단순한 정오답(True/False)을 넘어, 시스템의 기여도를 설명하는 세 가지 유의미한 그룹으로 분류하였다.

첫째, Group A: 의미론적 필터링을 통한 오알람 배제 (Correct Filter, 497건)

1차 객체 탐지 모델은 알람을 발생시켰으나, 본 시스템이 영상의 전후 맥락을 파악하여 '비위험'으로 필터링한 케이스가 497건(전체의 78.7%)에 달하였는데 영상 분석 결과, 이는 주로 '단순 이동 중인 작업자', '안전 수칙을 준수하며 휴식 중인 인원', '작업 구역 밖의 행인' 등 객체가 멀리 있는 중장비와 겹쳐서 발생하였다. 시스템은 497건의 무의미한 알람을 사전에 차단함으로써, 관리자가 겪는 '알람 피로(Notification Fatigue)'를 획기적으로 경감시키고, 실제 위험 상황에 집중할 수 있는 업무 환경을 조성하였다.

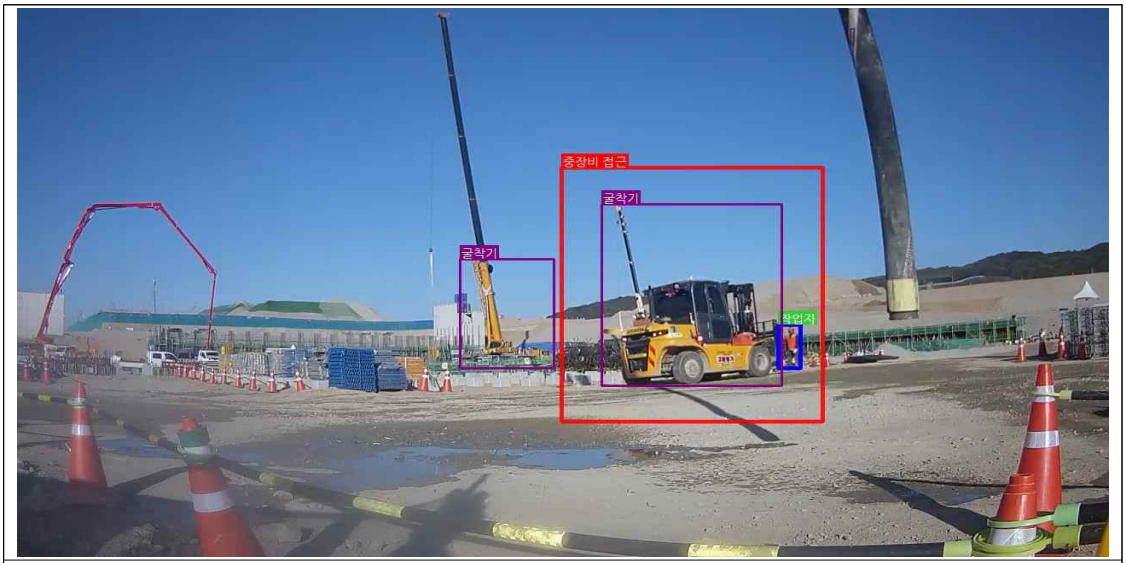
둘째, Group B: 실제 위험 상황의 정확한 포착 (Correct Alert, 99건)

1차 모델과 본 시스템 모두 '위험'으로 판단하고, 인간 전문가 역시 이에 동의한 99건의 사례이다. 이는 '중장비 반경 내 접근' 등 즉각적인 시정 조치가 필요한 명백한 위험 상황이었다.

SUNG KYUN KWAN UNIVERSITY

""“알람 정의는 ‘작업자가 중장비와 근접해 있어 충돌 위험이 존재한다’는 것이지만, 사진에서는 작업자가 현장 중앙에 서 있고 중장비(크레인)가 이미 사진의 오른쪽에 정지해 있는 상태입니다. 작업자와 중장비 사이의 거리가 명확히 떨어져 있으며, 중장비가 작동 중이거나 움직이는 모습이 전혀 보이지 않습니다. 또한 작업자는 안전모를 착용하고 있고, 다른 보호 장비가 없더라도 현장 안전 조건을 충족합니다.”“”

그림 17 VLM 심층 분석을 통해 발견한 오알람(Fasle Alarm) 사례



""“지게차는 팔레트에 실린 자재가 불안정하게 쌓여 있는 영역을 통과하며, 운전자는 차량 방향을 바라보는 모습으로 시야 확보가 제한된 상태. 인근에 작업자가 지게차의 접근 경로를 따라 움직이고 있었으며, 충돌 가능성이 확인됨. 안전모는 착용했으나, 안전대 등 추가 보호구 착용 여부는 확인할 수 없었으며, 주변 인원은 보호구 미착용 상태가 의심됨”“”

그림 18 VLM 심층 분석을 통해 발견한 유효 위험 사례

SUNG KYUN KWAN UNIVERSITY

셋째, Group C: 과잉 필터링 및 기술적 한계 (Over-Filtering, 35건)

시스템이 안전하다고 판단했으나, 전문가 재검토 결과 잠재적 위험이 존재했던 35건이다. 이는 주로 VLM이 건설현장 도메인에 적응하지 못한 오류에서 기인한 문제로 향후 데이터셋 보강 및 RLHF(인간 피드백 강화학습)를 통해 개선해야 할 기술적 과제(Trade-off)로 식별되었다.

"""근로자 행동: 일부 근로자는 서서 대화를 나누거나 휴식 중이었으며, 주변을 제대로 살피지 않은 채 중장비 이동 경로에 들어서거나 직진하고 있었습니다. 보호구(안전모)는 일부 착용했으나, 안전대·고시등·반사체크·추락방지 장치는 미착용 혹은 부적절하게 착용된 상태였습니다.
위험 상황: 중장비가 이동 중인 경로와 근로자의 동선이 겹치고, 시야 확보가 어려운 상황에서 충돌 위험이 명백히 존재했습니다."""

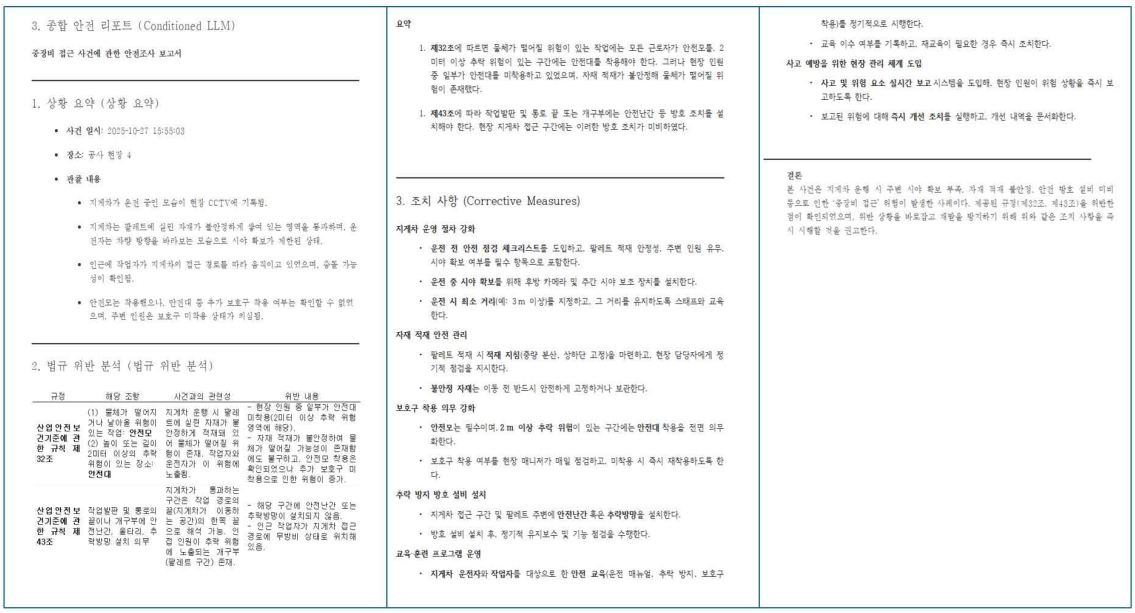
그림 19 VLM 심층 분석을 통해 발견한 위험탐지 실패 사례

이러한 결과를 종합하면, 본 시스템의 전체 상황 판단 정확도(Overall Accuracy)는 94.5%를 기록하였다. 특히, 시스템이 "이 알람은 무시해도 된다"고 판단했을 때 그 판단이 실제 정답일 확률인 필터링 정밀도(Filtering Precision)는 93.4%에 달하였다. 이는 본 연구가 제안하는 다단계 하이브리드 파이프라인이 현장의 고질적인 문제인 오알람을 구조적으로 해결하고, 실무에 즉시 투입 가능한 수준의 성능을 확보했음을 강력하게 시사한다.

또한 본 시스템은 심층 분석 결과를 단순 저장하는 데 그치지 않고, 관리자가 신

SUNG KYUN KWAN UNIVERSITY

속하게 활용할 수 있도록 자동화된 이메일 레포팅 기능을 구현하였다. 매일 새벽, 야간 배치 분석이 완료되면 시스템은 전날 발생한 모든 이벤트를 요약하여 마크다운(Markdown) 형식의 일일 안전 보고서를 생성하고, 안전 담당자의 이메일로 자동 전송한다. 보고서에는 각 이벤트의 발생 시각, 검출된 위험 요소, VLM이 분석한 현장 상황 요약, 그리고 매칭된 관련 법규 위반 사항이 포함된다. 이를 통해 관리자는 출근 직후 별도의 시스템 접속 없이도 전일의 안전 현황을 브리핑 받듯 확인할 수 있을 것으로 기대된다.



3. 종합 안전 리포트 (Conditioned LLM)

중장비 접근 사건에 관한 안전조사 보고서

1. 상황 요약 (상황 요약)

- **사건 일시**: 2025-10-27 15:55:03
- **장소**: 공사 현장 4
- **관찰 내용**
  - 지게차가 운전 중인 모습이 현장 CCTV에 기록됨.
  - 지게차는 팔레트에 실린 자재가 불안정하게 쌓여 있는 영역을 통과하며, 운전자는 차량 방향을 바라보는 모습으로 시야 확보가 제한된 상태.
  - 인근에 작업자가 지게차의 접근 경로를 따라 움직이고 있었으며, 충돌 가능성이 확인됨.
  - 안전모는 착용했으나, 안전대 등 추가 보호구 착용 여부는 확인할 수 없었으며, 주변 인원은 보호구 미착용 상태가 의심됨.

2. 법규 위반 분석 (법규 위반 분석)

규정	해당 조항	사건과의 관련성	위반 내용
산업안전보건기준에 관한 규칙 제32조	(1) 물체가 떨어지거나 날아올 위험이 있는 작업: **안전모** (2) 높이 또는 깊이 2미터 이상의 추락 위험이 있는 장소: **안전대**	지게차 운행 시 팔레트에 실린 자재가 불안정하게 적재돼 있어 물체가 떨어질 위험이 존재. 작업자와 운전자가 이 위험에 노출됨.	- 현장 인원 중 일부가 안전대 미착용(2미터 이상 추락 위험 영역에 해당).
산업안전보건기준에 관한 규칙 제43조	작업발판 및 통로의 끝이나 개구부에 안전난간, 울타리, 추락방망 설치 의무	지게차가 통과하는 구간은 작업 경로의 끝(지게차가 이동하는 공간)의 한쪽 끝으로 해석 가능. 인접 인원이 추락 위험에 노출되는 개구부(팔레트 구간) 존재.	- 해당 구간에 안전난간 또는 추락방망이 설치되지 않음. - 인근 작업자가 지게차 접근 경로에 무방비 상태로 위치해 있음.

요약

1. **제32조**에 따르면 물체가 떨어질 위험이 있는 작업에는 모든 근로자가 안전모를, 2미터 이상 추락 위험이 있는 구간에는 안전대를 착용해야 한다. 그러나 현장 인원 중 일부가 안전대를 미착용하고 있었으며, 자재 적재가 불안정해 물체가 떨어질 위험이 존재했다.
1. **제43조**에 따라 작업발판 및 통로 끝 또는 개구부에는 안전난간 등 방호 조치를 설치해야 한다. 현장 지게차 접근 구간에는 이러한 방호 조치가 미비하였다.

3. 조치 사항 (Corrective Measures)

지게차 운영 절차 강화

- **운전 전 안전 점검 체크리스트**를 도입하고, 팔레트 적재 안정성, 주변 인원 유무, 시야 확보 여부를 필수 항목으로 포함한다.
- **운전 중 시야 확보**를 위해 후방 카메라 및 주간 시야 보조 장치를 설치한다.
- **운전 시 최소 거리**(예: 3 m 이상)를 지정하고, 그 거리를 유지하도록 스태프와 교육한다.

자재 적재 안전 관리

- 팔레트 적재 시 **적재 지침**(중량 분산, 상하단 고정)을 마련하고, 현장 담당자에게 정기적 점검을 지시한다.
- **불안정 자재**는 이동 전 반드시 안전하게 고정하거나 보관한다.

보호구 착용 의무 강화

- **안전모**는 필수이며, **2 m 이상 추락 위험**이 있는 구간에는 **안전대** 착용을 전면 의무화한다.
- 보호구 착용 여부를 현장 매니저가 매일 점검하고, 미착용 시 즉시 재착용하도록 한다.

추락 방지 방호 설비 설치

- 지게차 접근 구간 및 팔레트 주변에 **안전난간** 혹은 **추락방망**을 설치한다.
- 방호 설비 설치 후, 정기적 유지보수 및 기능 점검을 수행한다.

교육·훈련 프로그램 운영

- **지게차 운전자**와 **작업자**를 대상으로 한 **안전 교육**(운전 매뉴얼, 추락 방지, 보호구 착용)를 정기적으로 시행한다.
- 교육 이수 여부를 기록하고, 재교육이 필요한 경우 즉시 조치한다.

사고 예방을 위한 현장 관리 체계 도입

- **사고 및 위험 요소 실시간 보고** 시스템을 도입해, 현장 인원이 위험 상황을 즉시 보고하도록 한다.
- 보고된 위험에 대해 **즉시 개선 조치를** 실행하고, 개선 내역을 문서화한다.

결론

본 사건은 지게차 운행 시 주변 시야 확보 부족, 자재 적재 불안정, 안전 방호 설비 미비 등으로 인한 '중장비 접근' 위험이 발생한 사례이다. 제공된 규정(제32조, 제43조)을 위반한 점이 확인되었으며, 위반 상황을 바로잡고 재발을 방지하기 위해 위와 같은 조치 사항을 즉시 시행할 것을 권고한다.

그림 20 자동 생성된 종합안전 리포트

5.3 정성적 평가 및 논의

정성적 관점에서 본 시스템은 단순한 '객체 검출기'의 한계를 넘어 '지능형 안전 보좌관'으로서의 역할을 성공적으로 수행함을 입증하였다. 1차 객체 탐지 모델이 "무엇이 어디에 있는가"를 식별하는 데 그쳤다면, 제안된 프레임워크는 "그 상황이 왜 위험한가"를 설명하고, "어떤 규정"에 위배되는지를 판단하는 고차원적 추론 능

SUNG KYUN KWAN UNIVERSITY

력을 실증하였다.

결론적으로, 본 실증 연구는 제안된 하이브리드 안전 시스템이 복잡하고 역동적인 건설 현장 환경에서도 높은 신뢰성을 유지하며 작동함을 확인시켜 주었다. 특히, VLM의 시각적 추론 능력과 LLM의 논리적 분석 능력이 결합됨으로써, 기존 시스템의 난제였던 오알람 문제를 구조적으로 해결함과 동시에 안전 관리의 질적 수준을 한 차원 격상시킬 수 있음을 명확히 증명하였다.

SUNG KYUN KWAN UNIVERSITY

제5장 결론

본 연구는 건설 현장이 내재적으로 안고 있는 통신 인프라의 불안정성, 데이터 보안의 폐쇄성, 그리고 엣지 컴퓨팅 자원의 물리적 제약이라는 삼중고(Trilemma)를 해결하기 위한 공학적 대안으로서, '온프레미스 기반의 지능형 안전 인지 프레임워크'를 제안하고 그 실효성을 검증하는 데 주력하였다.

기존의 센서 및 비전 AI 기술은 "작업자가 있다" 또는 "안전모를 쓰지 않았다"와 같은 단편적인 사실 감지에 그쳐, "작업자가 왜 위험한가"에 대한 인과적 맥락을 설명하지 못하는 한계가 있었다. 또한, 최근 등장한 고성능 인공지능 모델(VLM 및 LLM)은 뛰어난 추론 능력을 보유하고 있으나, 막대한 연산 비용과 클라우드 의존성으로 인해 통신이 불안정하고 데이터 보안이 필수적인 건설 현장에 직접 적용하기에는 현실적인 장벽이 존재하였다.

이에 본 연구는 다음과 같은 핵심 접근법을 통해 이러한 기술적 난제를 해결하고자 하였다.

첫째, 시공간 분리 처리 아키텍처를 설계하였다. 즉각적인 대응이 필요한 주간에는 경량화된 객체 탐지 모델을 통해 실시간 위험 이벤트를 포착하고, 시스템 부하가 적은 야간에는 고성능 VLM과 sLLM을 활용하여 축적된 데이터에 대한 심층 분석을 수행하는 이원화된 파이프라인을 구축하였다.

둘째, 온프레미스 최적화 기술을 적용하였다. 수십억 파라미터 규모의 모델(Gemma-3 27B, GPT-OSS 20B)을 단일 워크스테이션에서 운용하기 위해 양자화(Quantization) 모델을 선택/도입하였으며, 이를 통해 추론 성능의 저하를 최소화하면서도 메모리 점유율을 절감하였다.

SUNG KYUN KWAN UNIVERSITY

셋째, RAG 기반의 신뢰성 있는 리포팅 체계를 구현하였다. 생성형 AI의 환각(Hallucination) 현상을 억제하기 위해, VLM이 생성한 시각적 묘사를 쿼리로 변환하여 산업안전보건법 등 관련 법규를 검색(Retrieval)하고, 이를 근거로 안전 리포트를 생성이 가능케 하였다.

실제 토목 공사 현장에서의 실증 실험 결과, 제안된 시스템은 94.5%의 종합 판단 정확도를 기록하였으며, 특히 기존 시스템의 고질적 문제였던 오알람(False Alarm)에 대해 93.4%의 필터링 정밀도를 보임으로써 관리자의 업무 피로도를 유의미하게 경감시킬 수 있음을 확인하였다. 또한, RAG 기술을 접목하여 단순한 위험 감지를 넘어 산업안전보건법에 근거한 구체적인 위반 사유와 시정 조치를 텍스트로 자동 생성함으로써, AI 시스템이 단순 관제 도구를 넘어 안전 관리자의 의사결정을 보조하는 '지능형 에이전트'로 진화할 수 있는 가능성을 제시한 것으로 평가된다.

본 연구의 성과에도 불구하고, 다음과 같은 한계점이 존재하며 이는 향후 연구를 통해 발전되어야 한다.

- 본 시스템은 단일 GPU 워크스테이션을 기준으로 설계되어, 수십 대 이상의 CCTV가 설치된 대규모 현장의 데이터를 야간 시간 내에 전수 분석하기에는 처리 용량(Throughput)의 물리적 한계가 존재한다. 향후 연구에서는 다수의 엣지 노드를 연결하는 분산 처리(Distributed Computing) 아키텍처나 쿠버네티스(Kubernetes) 기반의 오케스트레이션 기술 도입이 필요하다. 현장의 엣지 디바이스들은 네트워크 단절이나 전력 불안정으로 인해 가용성이 수시로 변동될 수 있으므로, 이러한 장애 상황을 능동적으로 감지하고 작업 부하(Workload)를 가용 노드로 자동 재할당하는 자율 복구(Self-healing) 메커니즘이 필수적이다. 또한, 최신 AI 모델의 지속적인 업데이트와 배포(CI/CD)를 자동화하여 현장 관리자의 기술적 개입을 최소화하는 MLOps 체계의 구축 또한 중요한 과제로 남아 있다.

SUNG KYUN KWAN UNIVERSITY

- 본 연구에서는 범용적으로 학습된 파운데이션 모델을 프롬프트 엔지니어링과 RAG로 보정하여 사용하였다. 건설 현장의 특수한 용어(예: 비계, 동바리, 슬래브 등)나 뉘앙스를 완벽하게 이해하기 위해서는, 건설 도메인 데이터셋을 구축하여 모델 자체를 미세 조정(Fine-tuning)하거나 LoRA(Low-Rank Adaptation)과 같은 효율적인 튜닝 기법을 적용하는 후속 연구가 요구된다.

결론적으로, 본 연구는 최신 AI 기술을 건설 안전이라는 보수적인 도메인에 접목하여, 기술적 진보와 현장의 실용성을 동시에 추구한 시도라는 점에서 의의가 있다. 제안된 프레임워크가 향후 스마트 건설 안전 기술의 표준 모델로 발전하여, 건설 현장의 무재해 달성에 기여하기를 기대한다.

SUNG KYUN KWAN UNIVERSITY

참 고 문 헌

김명호, 신성우, 서용윤, 2019. 건설현장 근로자의 안전모 착용 여부 검출을 위한 컴퓨터 비전 기반 딥러닝 알고리즘의 적용, 한국안전학회지 제34권 제6호, pp.29-37

김현우, 손영두, 이찬, 박지혁, 정상국, 김대근, 2025. 멀티 모달 신경망을 활용한 영상·소리 분석 기반 경계 감시용 지능형(AI) 보안시스템, 국방로봇학회 논문지, Vol.4(1), pp.1-10

전소연, 2020. 딥러닝 기반 영상 분석 알고리즘을 이용한 실시간 작업자 안전관리 시스템 개발, 스마트미디어저널, Vol.9(3), pp.1-9

Adil, M., Lee, G., Gonzalez, V.A. and Mei, Q., 2025. Using Vision Language Models for Safety Hazard Identification in Construction, arXiv preprint arXiv:2504.06649

Cheng, T. and Teizer, J., 2013. Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications, Automation in Construction, Vol.34, pp.3-15

Ding, L., Fang, W., Luo, H., Love, P.E., Zhong, B. and Ouyang, X., 2018. A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory, Automation in Construction, Vol.86, pp.118-124

SUNG KYUN KWAN UNIVERSITY

Fang, W., Ding, L., Luo, H. and Love, P.E., 2018a. Falls from heights: A computer vision-based approach for safety harness detection, Automation in Construction, Vol.91, pp.53-61

Gemini Team, Google, 2023. Gemini: A Family of Highly Capable Multimodal Models, arXiv preprint arXiv:2312.11805

Girshick, R., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.580-587

Jiang, A.Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D.S., de Las Casas, D., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., Lavaud, L.R., Lachaux, M.-A., Stock, P., Le Scao, T., Lavril, T., Wang, T., Lacroix, T. and El Sayed, W., 2023. Mistral 7B, arXiv preprint arXiv:2310.06825

Kim, H., 2023. Image Generation of Hazardous Situations in Construction Sites for DNN Model Training using Text-to-Image Synthesis, Master's Thesis, Ewha Womans University

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S. and Kiela, D., 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, Advances in Neural Information Processing Systems (NeurIPS), Vol.33, pp.9459-9474

Liu, H., Li, C., Wu, Q. and Lee, Y.J., 2023. Visual Instruction Tuning, Advances in Neural Information Processing Systems (NeurIPS),

SUNG KYUN KWAN UNIVERSITY

Vol.36, pp.34892-34916

Mneymneh, B.E., Abbas, M. and Khoury, H., 2017. Automated hardhat detection for construction safety applications, Procedia Engineering, Vol.196, pp.895-902

Nath, N.D., Behzadan, A.H. and Paal, S.G., 2020. Deep learning for site safety: Real-time detection of personal protective equipment, Automation in Construction, Vol.112, p.103085

OpenAI, 2023. GPT-4 Technical Report. arXiv preprint arXiv:2303.08774.

Park, M., Kulinan, A.S., Tran, D.Q., Bak, J. and Park, S., 2024. Preventing falls from floor openings using quadrilateral detection and construction worker pose-estimation, Automation in Construction, Vol.165, 105226

Park, M., Tran, D.Q., Bak, J. and Park, S.-H., 2023. Small and overlapping worker detection at construction sites, Journal of Building Engineering, Vol.76, 107286

Park, M., Tran, D.Q., Bak, J., Kulinan, A.S. and Park, S., 2023a. Real-time monitoring unsafe behaviors of portable multi-position ladder worker using deep learning based on vision data, Journal of Building Engineering, Vol.68, 106096

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.779-788

Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal

SUNG KYUN KWAN UNIVERSITY

Networks, Advances in Neural Information Processing Systems (NeurIPS), Vol.28, pp.91-99

S. Al-Azani, H. Luqman, M. Alfarraj, A. A. I. Sidig, A. H. Khan and D. Al-Hamed,, 2024. "Real-Time Monitoring of Personal Protective Equipment Compliance in Surveillance Cameras," in IEEE Access, vol. 12, pp. 121882-121895,

Seo, J., Han, S., Lee, S. and Kim, H., 2015. Computer vision techniques for construction safety and health monitoring, Advanced Engineering Informatics, Vol.29(2), pp.239-251

Teizer, J., Cheng, T. and Fang, Y., 2013. Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity, Automation in Construction, Vol.35, pp.53-68

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Canton Ferrer, C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B., Gao, C., Goswami, V., Goyal, N., Hartshorn, A., Hosseini, S., Hou, R., Inan, H., Kardas, M., Kerkez, V., Khabsa, M., Kloumann, I., Korenev, A., Koura, P.S., Lachaux, M.-A., Lavril, T., Lee, J., Liskovich, D., Lu, Y., Mao, Y., Martinet, X., Mihaylov, T., Mishra, P., Molybog, I., Nie, Y., Poulton, A., Reizenstein, J., Rungta, R., Saladi, K., Schelten, A., Silva, R., Smith, E.M., Subramanian, R., Tan, X.E., Tang, B., Taylor, R., Williams, A., Kuan, J.X., Xu, P., Yan, Z.,

SUNG KYUN KWAN UNIVERSITY

Zarov, I., Zhang, Y., Fan, A., Kambadur, M., Narang, S., Rodriguez, A., Stojnic, R., Edunov, S. and Scialom, T., 2023. Llama 2: Open Foundation and Fine-Tuned Chat Models, arXiv preprint arXiv:2307.09288

Uhm, M., Kim, J., Ahn, S., Jeong, H. and Kim, H., 2025. Effectiveness of retrieval augmented generation-based large language models for generating construction safety information, Automation in Construction, Vol.170, 105926

Wang, Y., Xiao, B., Bouferguene, A. and Al-Hussein, M., 2024. Proactive safety hazard identification using visual - text semantic similarity for construction safety management, Automation in Construction, Vol.166, 105634

Wu, J., Cai, N., Chen, W., Wang, H. and Wang, G., 2019. Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset, Automation in Construction, Vol.97, pp.22-37

Zhang, J., Pan, Y. and Zhang, L., 2024. Construction-GPT: A domain-specific multimodal generative model for automated hazard identification and description. Automation in Construction, 161, p.105342.

Zhang, Y., Wang, X., Hu, M. and Zhang, Z., 2024. BifrostRAG: Bridging Dual Knowledge Graphs for Multi-Hop Question Answering in Construction Safety, Automation in Construction, Vol.168, 105789

SUNG KYUN KWAN UNIVERSITY

ABSTRACT

A Study on a Artificial Intelligence Framework for Safety Perception in Understanding and Responding to Risk Situations in Construction Sites

Taeheon Kim
Department of Global Smart City
Sungkyunkwan University

Construction site accidents and disasters are intricate causal events stemming from the complex interplay of worker behavior and environmental factors. Despite this inherent complexity, conventional sensor and computer vision technologies have largely been confined to phenomenological data acquisition, such as basic object localization or hard hat detection, thus failing to provide a comprehensive contextual understanding of accident risks and leaving a significant 'semantic gap'. While large-scale artificial intelligence models (e.g., VLM, LLM) offer the potential to overcome these limitations, their inherent high computational costs and cloud dependency present considerable practical impediments for direct application in construction environments, which are typically characterized by inadequate communication infrastructure and stringent data security requirements. Consequently, this study proposes an 'On-Premise

SUNGKYUNKWAN UNIVERSITY

Hybrid Safety Awareness Framework′ that strategically converts the inherent constraints of edge environments into opportunities for optimized computational efficiency.

The core of this research lies in the spatio-temporal partitioning of computational resources. A dual-architecture has been meticulously implemented: lightweight object detection models are employed for real-time risk event capture during daytime, when immediate response is paramount, while high-performance VLM and sLLM are engaged for in-depth analysis of accumulated data during nighttime, utilizing dormant computational resources. Furthermore, model quantization enables deployment on a single workstation, obviating the need for high-specification servers. The integration of Retrieval-Augmented Generation (RAG) effectively mitigates generative AI hallucinations and ensures the generation of safety reports compliant with the Industrial Safety and Health Act. Empirical validation in a real-world construction setting demonstrated that this approach accurately filters 93.4% of false alarms—a pervasive challenge in conventional object detection systems—thereby confirming its capacity to significantly alleviate the alarm fatigue experienced by safety managers.

Ultimately, this research advances an optimized engineering solution by adapting cutting-edge generative AI to the specific physical and environmental constraints of the construction domain, thereby aiming to establish a novel paradigm for smart construction safety technology.

Keywords: Construction Safety, Artificial Intelligence, Small Large Language Model (sLLM), False Alarm Reduction

SUNG KYUN KWAN UNIVERSITY