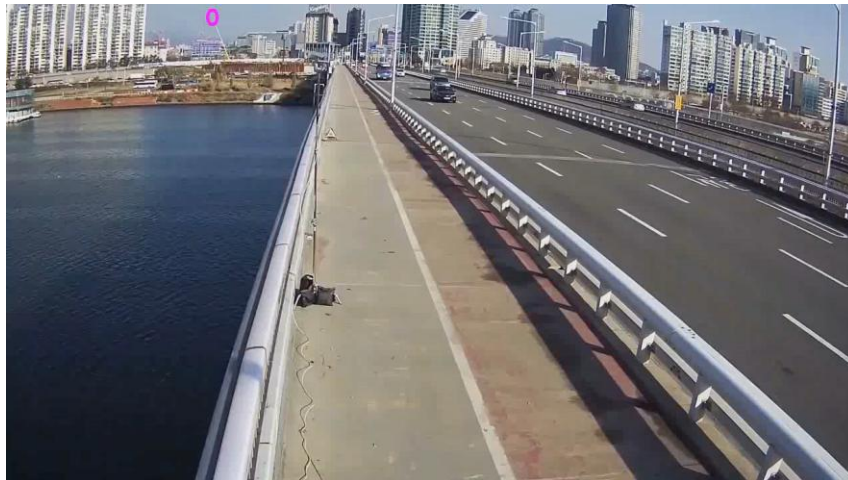
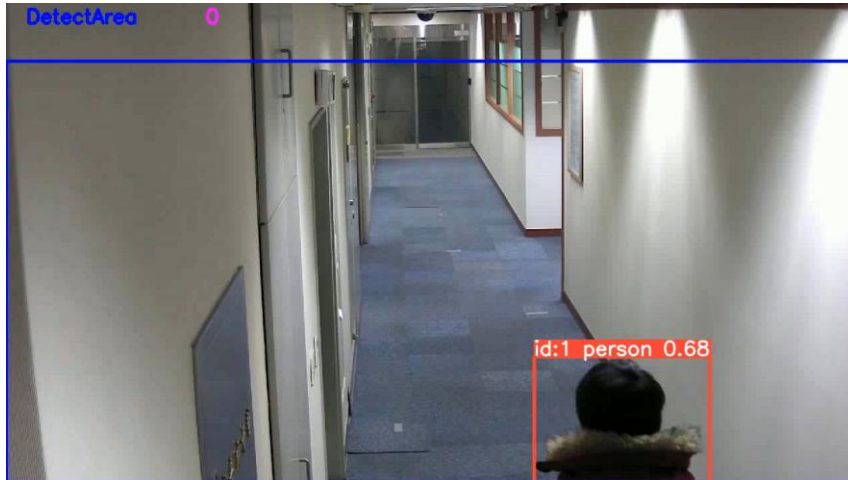


K-국방 융합데이터 기반 인공지능 학습데이터셋 구축

- 데이터 관리와 품질관리를 중심으로

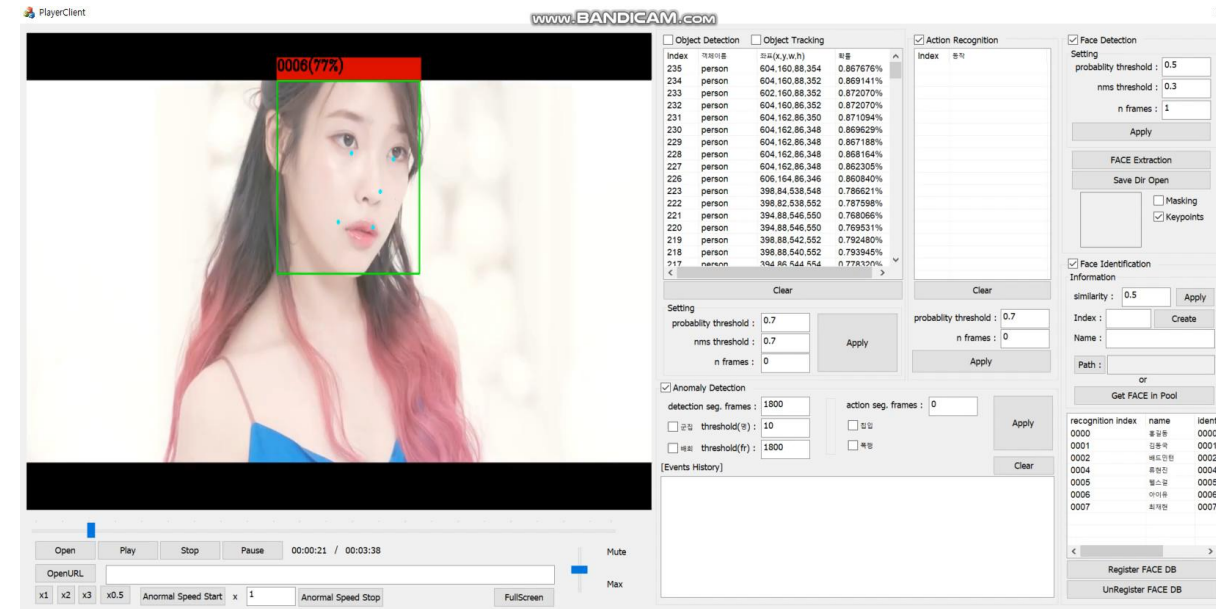
I. 회사 소개

○ CCTV Event Detection(falldown)



I. 회사 소개

● CCTV 실시간 비식별화



I. 회사 소개

Context Learning 기반 상황인식 AI comparision 2024.10.30



지금 뭐하고 있는 상황인지 한국어로 설명해줘

안전한 작업 환경을 확보하기 위해 안전대를 설치했습니다. 작업자는 안전대에서 작업물을 설치하거나 조작하고 있습니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다. 작업자는 안전대에서 작업물을 설치하거나 조작하고 있습니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다. 작업자는 안전대에서 작업물을 설치하거나 조작하고 있습니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다.

안전벨트를 착용하고 있나요?

네, 안전벨트를 착용하고 있습니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다. 작업자는 안전복을 착용하고 있습니다. 작업장에서 안전대를 설치하는 것은 안전 작업을 위해 필수입니다.

ChatGPT 4o



지금 뭐하고 있는 상황인지 한국어로 설명해줘

이 사진은 건물 외벽에서 작업을 하는 근로자가 안전 벨트를 착용한 상태에서 유리창을 설치하거나 보수 작업을 하고 있는 모습으로 보입니다. 그는 높은 곳에서 작업 중이므로 안전 장비를 착용하고, 비계(foot scaffold) 위에서 작업을 수행하고 있습니다. 안전 벨트는 그의 작업 안전을 확보하기 위해 벽에 고정되어 있는 것으로 보이며, 이와 같이 높은 곳에서 작업할 때는 추락 방

안전벨트를 착용하고 있나요?

이 사진에서는 근로자가 안전벨트를 착용하고 있지 않은 것으로 보입니다. 허리에 벨트가 있긴 하지만, 벽이나 비계에 고정된 모습이 보이지 않습니다. 안전벨트가 고정되지 않은 상태로 비계 위에서 작업하는 것은 매우 위험할 수 있습니다. 고소 작업 시에는 반드시 안전벨트를 구조물에 제대로 고정하여 추락 위험을 방지하는 것이 중요합니다.

메시지 ChatGPT

II. 현재의 AI

What is AI? ANI vs. AGI vs. ASI



Artificial narrow intelligence (ANI)

Designed to perform specific tasks



Artificial general intelligence (AGI)

Can behave in a human-like way across all tasks



Artificial super intelligence (ASI)

Smarter than humans—the stuff of sci-fi



9 : 1
8 : 2



Chat GPT
로봇

특수한 서비스
- 드론
- HMD
- MUMT

II. 현재의 AI



Detection Prompt

soldier|

Run



Detection Prompt

soldier.

Run

Advanced options



III. 학습 데이터

- 원천데이터로 활용 가능한 데이터(수집 후 정제한 데이터)
- 지도학습을 위한 가공 데이터(원천데이터를 전처리, 가공한 학습데이터)

User					
UserID	User	Address	Phone	Email	Alternate
1	Alice	123 Foo St.	12345678	alice@example.org	alice@neo4j.org
2	Bob	456 Bar Ave.		bob@example.org	
...
99	Zach	99 South St.		zach@example.org	

Order	
OrderID	UserID
1234	1
5678	1
...	...
5588	99

LineItem		
OrderID	ProductID	Quantity
1234	765	2
1234	987	1
...
5588	765	1

Product		
ProductID	Description	Handling
321	strawberry ice cream	freezer
765	potatoes	
...	...	
987	dried spaghetti	

정형 데이터

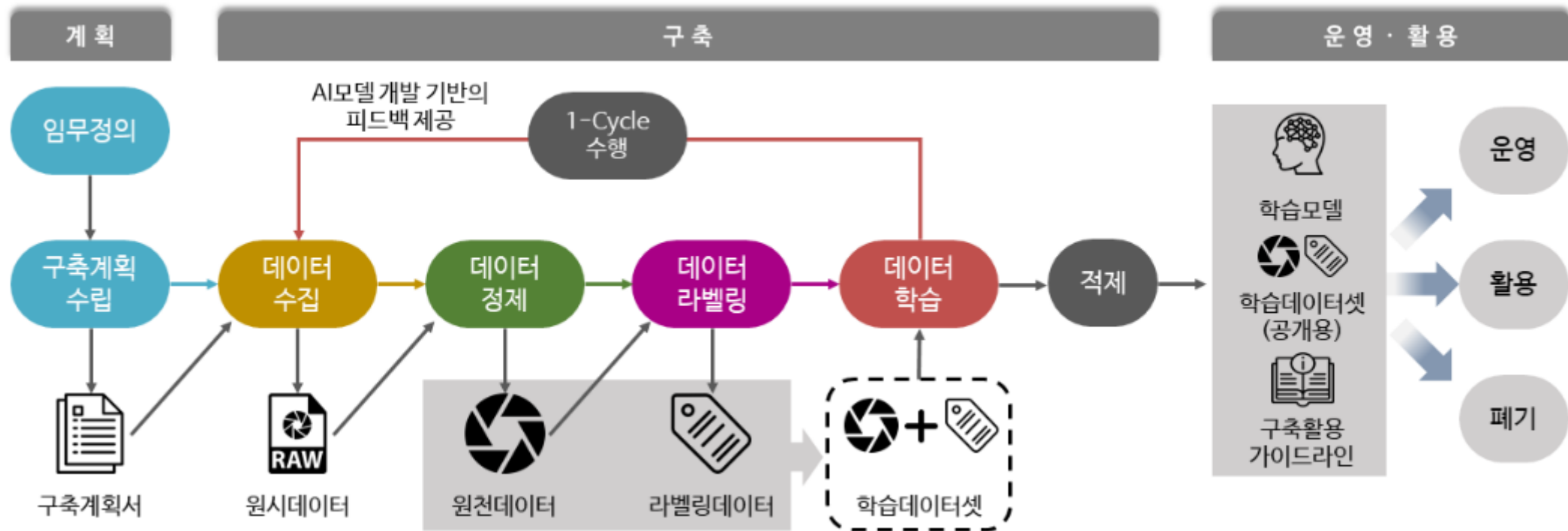
```
"header": {  
  "title": "The JSON example",  
  "descriptionText": "This is some title text",  
},  
"content": {  
  "title": "The content example text",  
  "elements": [  
    {  
      "title": "The first element",  
      "mainText": "First element main text",  
      "additionalText": "First element additional text",  
    },  
    {  
      "title": "The second element",  
      "mainText": "Second element main text",  
      "additionalText": "Second element additional text",  
    }  
  ]  
}
```

반정형 데이터



비정형 데이터

Ⅲ. 학습 데이터



AI Training DATA
->AI의 성능을 좌지우지하는 것은 DATA

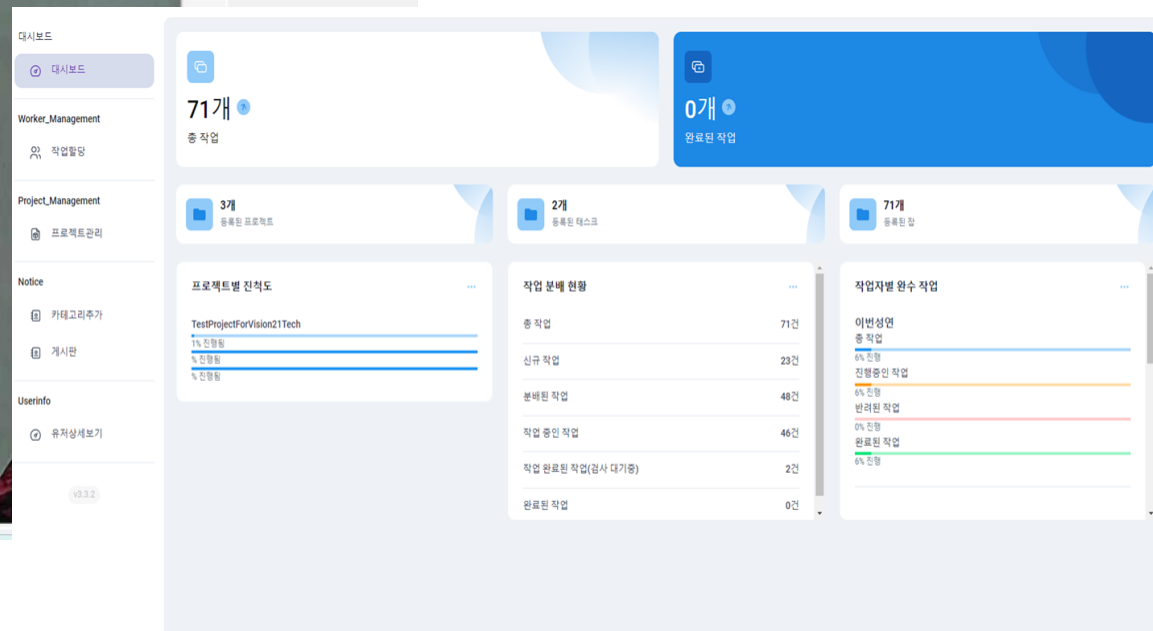
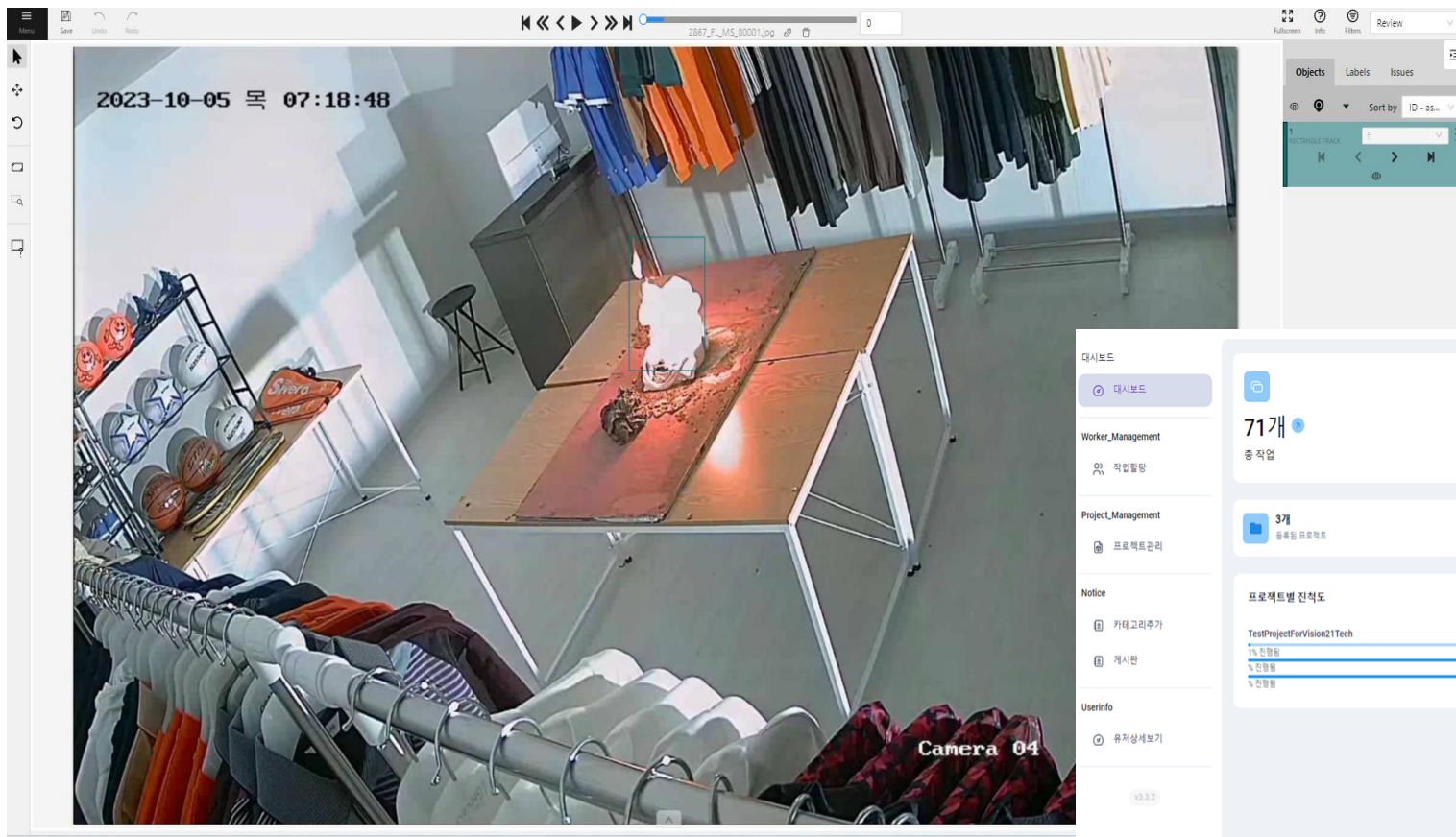
Ⅲ. 가공(Annotation)



Ⅲ. 가공(Annotation)

가공 도구 필요성

- 고비용, 저효율-> 저비용 고효율



IV. 군의 데이터를 활용한 AI 서비스 수요(사례)

수요 기술(제품)명	폐수처리시설 실시간 모니터링 시스템	
수요자	구분	■ 육군, □ 해군, □ 공군
개발 희망(예상) 기간	■ 12개월, □ 18개월, □ 24개월	
기술(제품) 키워드	누출감지, 폐수처리, 환경, 화학, 자동화, 실시간	
제안 배경 및 필요성	<ul style="list-style-type: none"> ○ 폐수처리시설 유해화학약품관리 수작업 수행 <ul style="list-style-type: none"> - 시설 내 화학약품(황산 등 4종) 탱크 및 위탁·자가 폐수처리 수조 덮개를 작업자가 제거하여 수작업으로 수위확인(실시간 확인 불가) - 유해화학물질 누출 등 인명사고 발생 가능성 상존 ○ 화학물질 누출 시 즉시 대처 불가 <ul style="list-style-type: none"> - 누출 모니터링 및 경고시스템 부재, 사고 발생시 초동조치 지연으로 인한 인명사고 및 토양·수질오염 발생 우려 	
	○ 폐수처리시설 실시간모니터링 시스템 및 누출감지 문자발송시스템 구축	
요구기능	항 목	요구내용 및 수준
	실시간 모니터링시스템	<ul style="list-style-type: none"> • 위탁 및 자가처리 수조 내 수위 실시간 모니터링 • 화학약품 탱크 내 수위 실시간 모니터링 • 시설 내 펌프 등의 작동 유·무 상태 실시간 모니터링 • 실시간 모니터링을 위한 상황실 구축 • 장비 전반 CCTV 설치
	화학물질 누출감지 문자발송시스템	<ul style="list-style-type: none"> • 약품탱크(황산 등 4종) 누출감지 즉시 담당자에게 경고문자가 개인 휴대폰으로 자동발송되도록 구축 • 예시) 0000시 00분에 누출감지 되었습니다. 즉시 초동조치할 것.

수요 기술(제품)명	이미지 데이터셋 자동구축 솔루션	
수요자	구분	□ 육군, □ 해군, ■ 공군
개발 희망 (예상) 기간	□ 12개월, □ 18개월, ■ 24개월	
기술(제품) 키워드	오토라벨링, AI 학습용 데이터셋 구축(이미지영상), 전처리, 품질관리	
제안 배경 및 필요성	<ul style="list-style-type: none"> · 기 구축 AI 서비스의 고도화를 위해 지속적으로 발생하는 신규 데이터의 전처리 (정제, 증강, 레이블링 등) 작업 소요 지속 발생 · 사람의 수작업으로 진행하는 레이블링 작업으로 인한 개발비용* 증가 * AI 개발비용의 약 80% 이상을 데이터셋 구축비용 차지(개발기간, 인건비 등) · 학습데이터 생산과정에서 동일인의 반복작업에 따른 집중력 저하와 다수의 인력 참여 시, 숙련도에 따른 품질이슈(높은 검수비용 등) 발생 	
	ex) 지능형 공대지사격 자동채점체계 <div data-bbox="1549 861 2300 1089"> </div> <ul style="list-style-type: none"> ① 원본 데이터 영상 추출 / 이미지 분할 후 AI 모델을 활용한 이미지 자동분류 및 이미지 內 영역분리(바운딩박스, 폴리곤, 세그멘테이션 등) ② 데이터를 이용해 AI 모델이 자동으로 생성한 데이터셋 이상 여부 점검 툴 개발 * 점검과정에서 모델이 잘못 제작한 데이터셋만 수작업으로 수정(AI 모델 성능 개선) ③ 데이터셋 자동 합성 및 증강(합성 및 증강기법 사용자설정 기능 제공)을 통해 데이터셋 수 증대 	

IV. 군의 데이터를 활용한 AI 서비스 수요(사례)

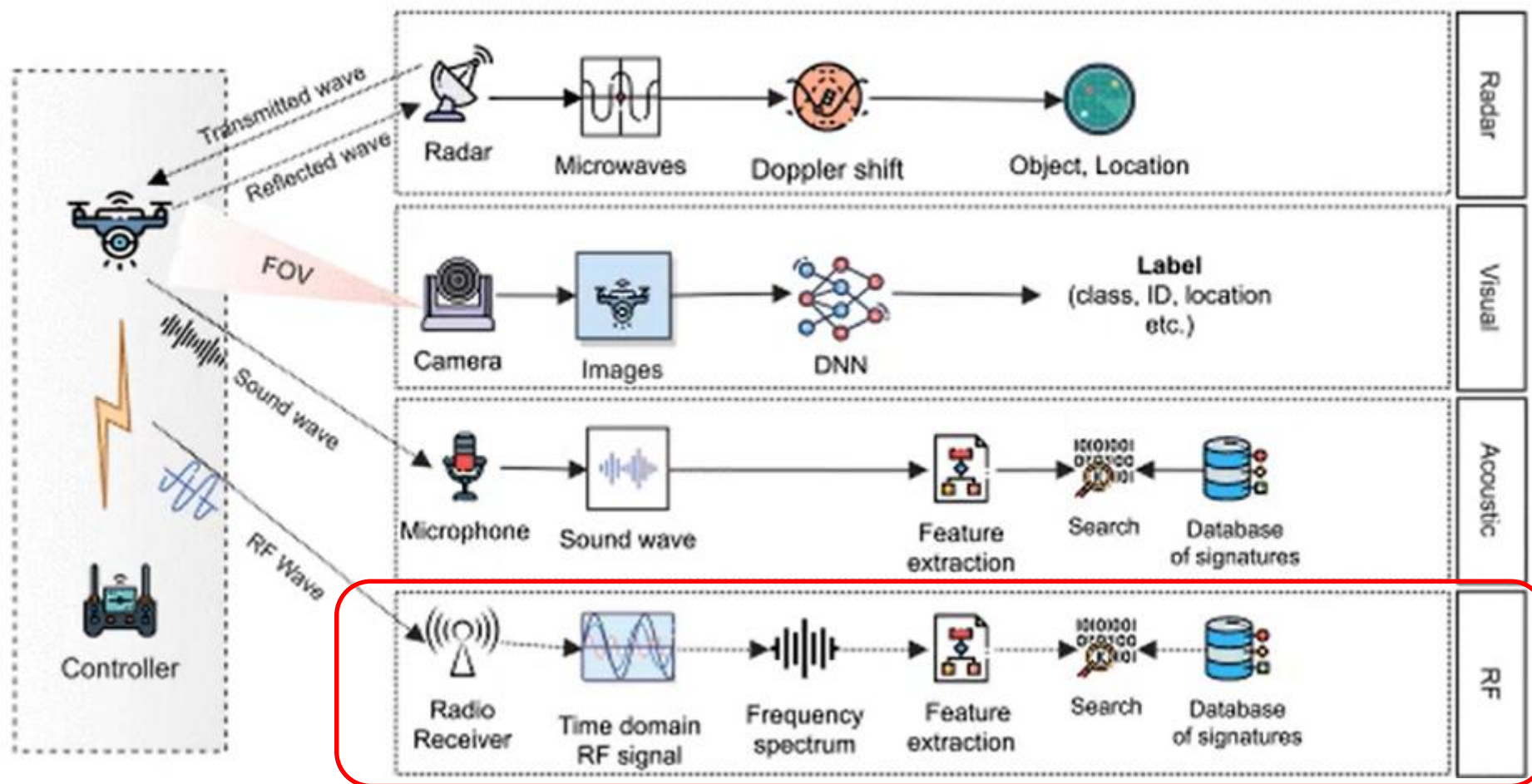
● AI 기반 항행안전시설 위험예측 시스템
- 김해공항



V. 군에서 융합 데이터 수집

● 드론 탐지 AI 학습 데이터

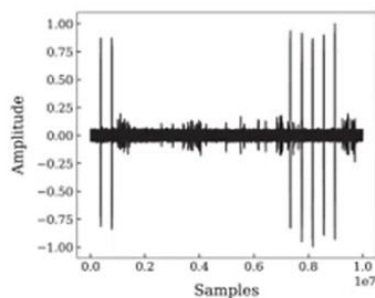
- Microwaves, Images, Sound wave, Time domain RF signal



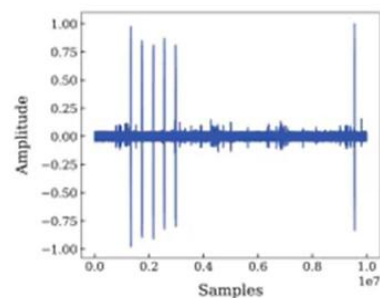
V. 군에서 융합 데이터 수집

● 드론 탐지 AI 학습 데이터

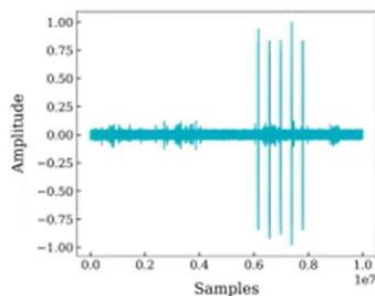
- Class Definition
- Preprocessing [Time domain RF signal + Frequency Spectrum(Spectrogram)]



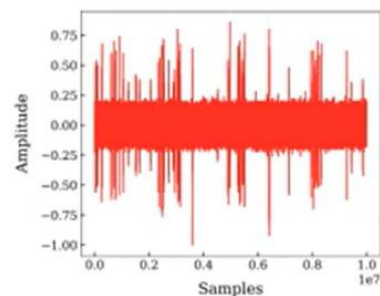
(a) Mode: On and connected.



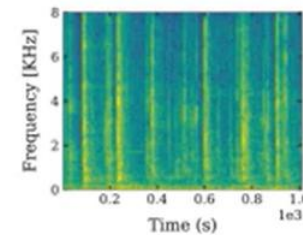
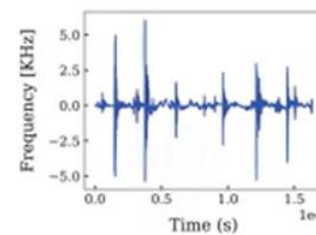
(b) Mode: Hovering.



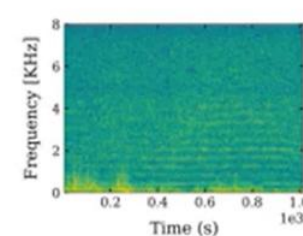
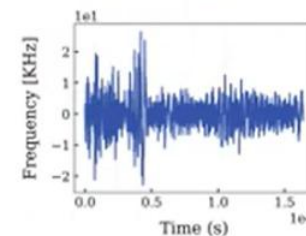
(c) Mode: Flying.



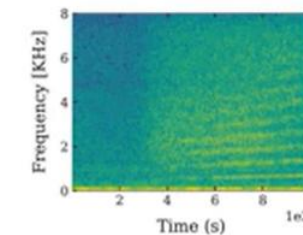
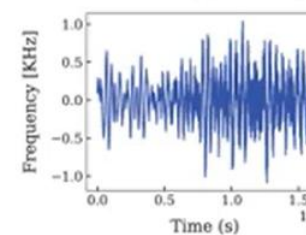
(d) Mode: Flying and video recording.



(a) Background noise (amplitude and spectrogram).



(b) Bebop drone's signature (amplitude and spectrogram).



(c) Membo drone's signature (amplitude and spectrogram).

VI. 인공지능 학습용 데이터 품질관리

- 인공지능 학습용 데이터의 구축과정과 이를 통해 생성되는 원시데이터, 원천데이터, 라벨링데이터 등 인공지능 학습용 데이터 특성 파악
- 데이터 구축과정을 기반으로 데이터의 생애주기를 운영 및 활용 영역까지 확대하여 정의하고, 생애주기 관점에서 데이터의 품질관리 영역 제시
- 구축한 데이터를 학습 모델에 적용하여 인공지능 학습용 데이터로서 유효성을 확보했는지 자체 검사

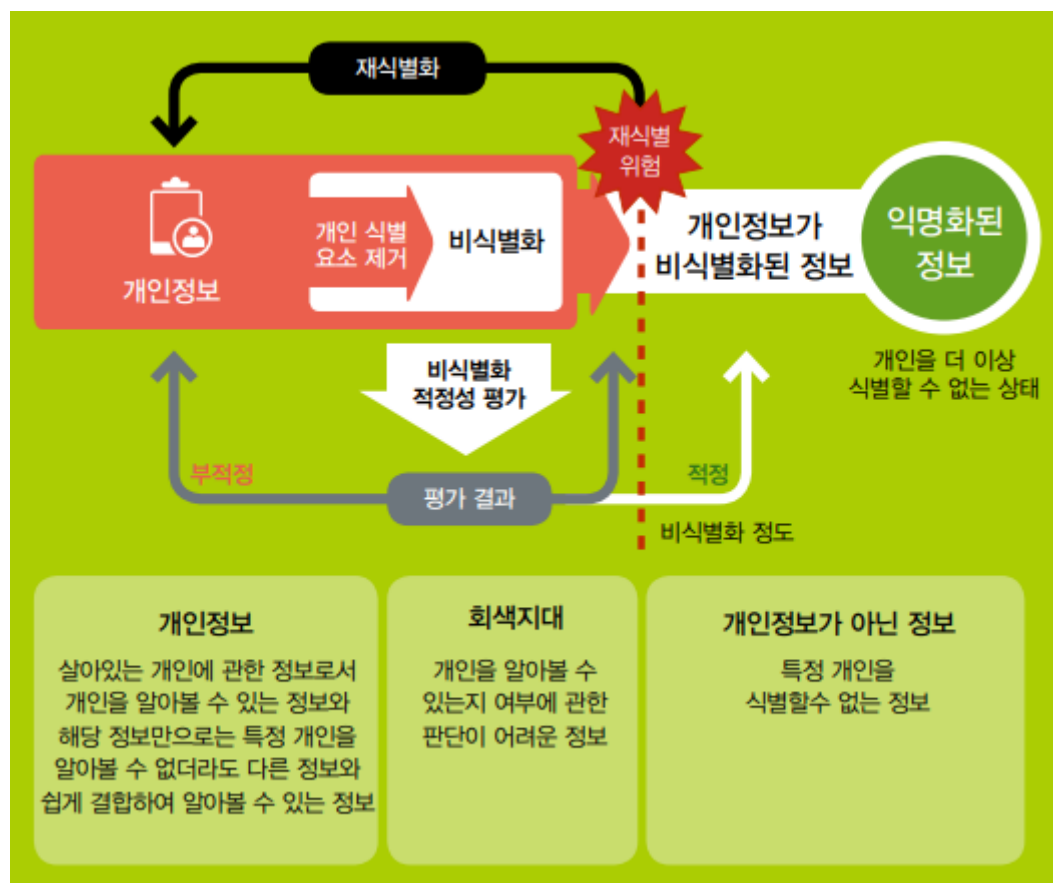
구분	품질 검사	품질 검증	품질 관리
정의	인공지능 학습용 데이터 구축 및 모델 관점에서 주요 품질 문제를 식별하고 문제의 근원적 원인을 파악하여 개선 기회를 도출하는 기법을 의미	인공지능 학습용 데이터의 품질보장을 위해 사업수행기관이 구축한 데이터의 품질을 제3자가 품질 검증	인공지능 학습용 데이터의 구축시 생산되는 원시데이터, 원천데이터, 라벨링 데이터와 구축공정의 품질을 확보하기 위한 제반 활동
중점	데이터 특성을 고려한 기준을 반영하여 지속가능한 품질확보에 중점	데이터의 결함을 예방하고 오류를 방지하는데 중점	데이터의 부적합, 결함을 감지하고 시정조치 활동에 중점
대상	특정 주요 지표 및 절차	특정 주요 지표 및 절차	원칙, 대상, 기준, 조직, 절차 등

VI. 데이터 담당 부서 필요

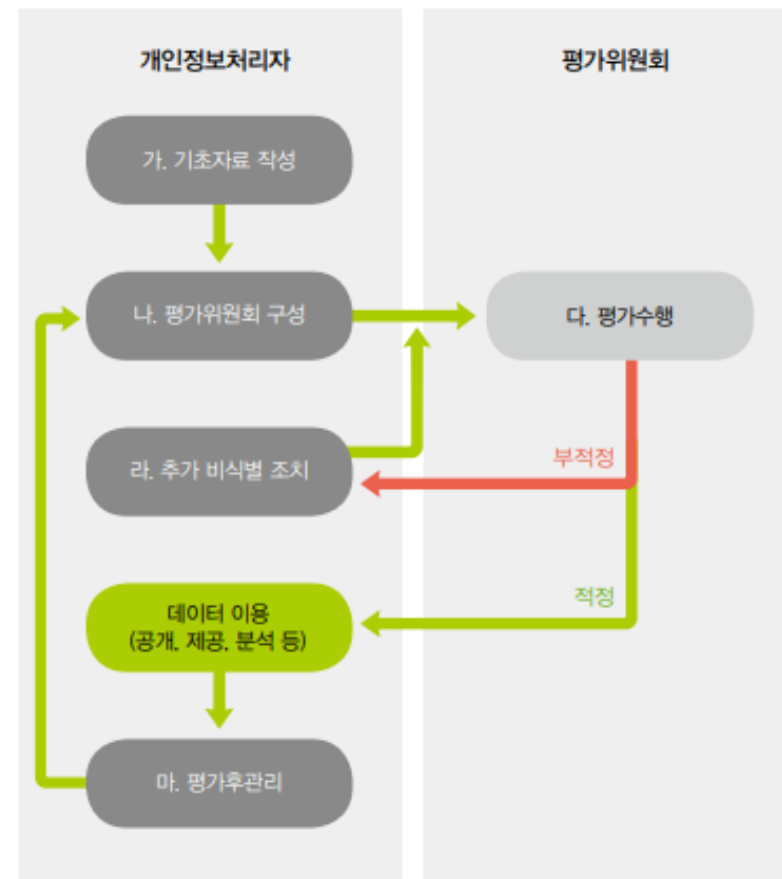
- 육군 정책실 AI 정책과
- 육군 인공지능 센터
- 인사사 AI 인재관리시스템 구축 TF
- 미래혁신 연구센터 AI/사이버기술연구과
- 교육사 AI 연구개발센터
- AI 기술 소요과
- AI 유무인복합체계전투개념발전과
- 군수사 AI/빅데이터분석과
- 육군 정작부 AI과학화 경계발전 TF

VII. 데이터 보안

- 군 기밀 정보 비식별화 : 보안성 강화



<개인정보 비식별화에 대한 적정성 자율평가 안내서>



<평가 절차>

Q&A