인공지능 데이터셋 구축 가이드북

AI DATASET CONSTRUCTION
GUIDEBOOK

과학기술정보통신부
Ministry of Science and ICT
NIA 한국정보화진흥원
NATIONAL INFORMATION SOCIETY AGENCY
TTA 한국정보통신기술협회
Telecommunications Technology Association

인공지능 데이터셋 구축 가이드북

과학기술정보통신부
Ministry of Science and ICT
NIA 한국정보화진흥원
NATIONAL INFORMATION SOCIETY AGENCY
TTA 한국정보통신기술협회
Telecommunications Technology Association

AI DATASET CONSTRUCTION GUIDEBOOK

가이드북 발간 소개

정부는 「데이터·AI경제 활성화 계획(2019.1.16.)」을 통해 4차 산업혁명의 성공이 고도의 인공지능 기술 확보 및 데이터와 인공지능 간 유기적인 융합에 달려있다고 선언한바 있습니다.

본격적인 AI와 데이터 시대를 맞아 TTA는 과기정통부와 한국정보화진흥원의 지원을 받아 인공지능 데이터 구축지침 표준화 및 품질검증에 앞장서왔으며 「지능정보산업인프라조성사업」의 2019년도 인공지능 학습 데이터를 대상으로 데이터셋 구축 가이드라인을 엮어서 발간하게 되었습니다.

올해 구축된 인공지능 학습 데이터는 이미지 5종, 영상 4종, 텍스트 1종 등 총 10종입니다. 이미지의 경우 랜드마크 인식, 손글씨 인식, 질환 진단, 엑스레이 위험물 탐지, 다양한 각도의 얼굴인식용 데이터를 수집 및 가공하였습니다. 그리고 영상 데이터는 이상행동탐지, 동작분석, 보행보조, 감정인식을 위한 영상을 제작 및 가공하였습니다. 마지막으로 대규모 한영 병렬 말뭉치를 제작하여 기계번역 성능향상에 도움을 주고자 하였습니다.

본 가이드북은 인공지능 기술개발에 활용되는 다양한 데이터셋의 실제 학습용 데이터 구축 경험 및 활용을 통해 축적한 노하우를 집약한 결과물입니다.

또한 여기서 만들어진 모든 인공지능 학습 데이터는 「AI 허브(aihub.or.kr)」를 통해 공개하여 산학연에서 자유롭게 활용할 수 있도록 하였습니다.

2019년 11월

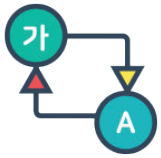
Contents

인공지능 데이터셋 구축

한국어-영어 번역말뭉치 데이터

고품질 AI 번역 엔진 개발과 다양한 언어 지능 기술 개발,
자동번역 챗봇 평가/개발, 커뮤니티 서비스에 활용되는 데이터

01



한국형 사물 이미지 데이터

국내 장소, 객체에 대한 인공지능 인식기술 개발 및
성능강화를 위해 한국형 사물이미지인 국가지정문화재, 국내 상품,
도시별 랜드마크 이미지 AI 데이터

02

한국어 글자체 이미지 데이터

자율주행차 이정표 인식, 증강현실 및 IoT 산업의 문자 판독,
전자상거래 산업의 상품 라벨 및 도서표지 검색 등에 활용되는 데이터

03

인도(人道) 보행 영상 데이터

인도보행 장애물 객체 인식을 위한 29종 객체 어노테이션, 보행 노면의
마스킹과 Depth Prediction ZED Stereo Camera 촬영 이미지 등이
포함되어 보행 보조 및 자율주행 기술 개발에 활용되는 데이터

04

멀티모달 영상 데이터

감정 분석, 개체 인식, 인물 인식, 관계 분석, 영상 기반 질의응답,
대화처리, 음성 인식, 영상 상황·의도 분석에 활용되는 데이터

05

가이드북 데이터 소개 –AI Hub(aihub.or.kr)

06

사람동작 영상 데이터

사람 동작 영상은 20종의 기초동작과 30종의 의미론적 동작에 대하여 16관절의 학습 데이터를 만들고 있으며 샘플로 앉기, 팔짱끼기, 식사동작, 방향지시, 웨이트트레이닝의 동작을 담고 있음.
이 중 웨이트트레이닝에 대해서는 3D 데이터를 포함하여 오픈함

07

한국인 안면 이미지 데이터

한국인 얼굴에 대한 국내 특화 시각지능기술 및 서비스 개발을 위해 필요한 이미지 데이터의 제작, 수집, 분류 체계를 마련하고 관련 정보를 레이블링

08

위해물품 엑스레이 이미지 데이터

관공서 한국 공항 공사, 인천 공항 공사, 공세청, 기타 공항 항만의 AI 기반 자동 탐지 시스템 개발에 활용되는 데이터

09

질병 진단 이미지 데이터

의료 지식베이스와 인공지능 판독 시범 모델 개발을 위해 유방암 진단 및 안저 질환 진단 영상에 의료정보를 레이블링

10

이상행동 CCTV 영상 데이터

동작, 사건, 이상행동과 사건 또는 이상행동의 비디오 장면을 묘사하는 기술에 활용되는 데이터

출처 - http://www.aihub.or.kr

인공지능 데이터셋 구축 가이드북

AI DATASET CONSTRUCTION GUIDEBOOK

AI 데이터 구축 가이드라인

- 한국어·영어 번역 말뭉치
- 한국형 사물 이미지
- 한국어 글자체 이미지
- 인도(人道) 보행 영상
- 멀티모달 영상
- 사람동작 영상
- 한국인 안면 이미지
- 위해물품 엑스레이 이미지
- 질병 진단 이미지
- 이상행동 CCTV 영상

01

한국어 · 영어 번역 말뭉치 AI 데이터 구축 가이드라인

Saltlux Partners

	성명	소속	직위
가이드라인 제안	신석환	솔트룩스 파트너스	대표
가이드라인 작성	남보라	솔트룩스 파트너스	주임
전문가 검토 위원	최미란	ETRI	박사
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 솔트룩스파트너스/ 참여: 에버트란, 플리토

서 문

1. 가이드라인의 목적

이 가이드라인의 목적은 한국어-영어 번역 말뭉치 구축을 위한 프로세스와, 이 데이터를 활용하는 방법을 보고 독자가 각 단계를 잘 수행하기 위함이다. 크게는 5단계로 나누어 서술하였으며, 각 단계는 2~3개의 세부 단계로 구성하였고, 더 세부적으로는 문어체, 구어체, 대화체에 관한 내용을 담았다.

문어체는 가장 일반적인 코퍼스로 활용하는 범위가 넓기 때문에 주로 문어체 중에서도 뉴스를 예로 들어 설명하였으며, 이를 기반으로 다른 분야의 코퍼스도 쉽게 활용할 수 있게끔 하였다. 본 가이드라인에서의 구어체는 범용적으로 쓰이는 문장을 중심으로 하여 별도의 상황 데이터를 태깅하지 않고 구축하는 것으로 하며, 대화체는 상황 별 대화셋을 구축하여 향후 챗봇 인공지능의 학습 데이터로써 활용될 수 있게 하는 것을 목적으로 구축한다.

2. 주요 내용 요약

이 가이드라인은 고품질의 번역 말뭉치 데이터 셋을 구축하기 위한 데이터 수집, 정제, 가공, 검사, 활용 단계까지의 데이터 공급자가 고려해야 할 점과 이에 대한 예시를 설명한다.

구축 단계	세부 절차	설명
1. 수집	1-1 원천 데이터 선정	원천 데이터를 수집하는 대상 선정(저작권 확인, 문장 형식 검토)
	1-2 원천 데이터 수집	텍스트(문장) 형태의 원천 데이터 수집
2. 정제	2-1 부적합 데이터 삭제	수집 과정에서 형식 오류가 생긴 문장 삭제
	2-2 데이터 정제	목적에 맞게 데이터 수정
3. 가공	3-1 MTPE 또는 번역	1차 번역물 생성, 기계번역을 활용하거나 사람이 초벌 번역
	3-2 번역물 리뷰	초벌 번역 오류 수정 및 문장 개선. 필요에 따라 리뷰를 2단계로 수행
4. 검사	4-1 외부 검사	전체 번역 문장 평가/검증을 위해 별도 기관이 검사, 검사 지침 필요
	4-2 검사 결과 보고 및 조치	검사 결과를 가공 담당자에게 전달하여 조치 요청. 조치 후 결과 확인
5. 활용	5-1 번역 업무 활용	결과를 기존 번역 업무에 활용. NMT 학습 후 활용하거나 번역메모리(TM)로 활용
	5-2 코퍼스 연구 활용	단일 언어 또는 이중 언어 연구에 활용

인공지능 학습용 문어체/구어체/대화체 한영번역 말뭉치 구축 가이드라인

Guideline for Building Korean-English Parallel Corpus for AI Training in Literary, Colloquial, and Conversational Style

01

1. 적용 범위

본 가이드라인은 인공지능 학습을 위한 번역 말뭉치 구축에 있어서 작업 공정과, 상세 지침을 기술한다. 말뭉치의 형태는 문어체, 구어체, 대화체 총 3가지로 구축하며, 이를 영문으로 번역하는 내용이다. 본 가이드라인에서 정의하는 문어체의 예시로는 뉴스, 정부 지자체 홈페이지/간행물, 조례, 한국문화 분야 등이 있으며, 뉴스 분야를 전체 프로세스에서의 예시로 활용하였다. 구어체는 범용적으로 쓰이는 문체로 그 범위를 명확히 한정하지 않았으나 데이터 공급자는 수요자의 필요에 따라 그 범위를 한정할 수 있다. 본 가이드라인의 대화체는 챗봇 학습을 그 목적으로 하므로, 상황 정보가 있는 문장 세트로 그 내용을 한정한다.

2. 인용 표준

3. 용어 정의

● 3.1 기계번역 포스트에디팅 (MTPE, Machine Translation Post-Editing)

기계가 자동으로 번역한 번역문을 사람이 후편집(검수)하는 과정

● 3.2 문어체

문어(文語)로 쓰인 문체를 의미하며, 말 그대로 문서 작성에 적합한 어투와 용어를 사용하는 문체

● 3.3 구어체

문장에서만 쓰이는 특별한 말이 아닌 일상적인 대화에서 쓰는 말로 된 문체

● 3.4 대화체

위 구어체 문장 4개로 이루어진 세트를 본 가이드라인의 대화체로 정의

4. 약어

- NMT Neural Machine Translation
- AI Artificial Intelligence
- PE Post-Editing

5. 데이터 수집

수집은 크게 원천 데이터 선정 및 수집 크게 2단계로 나뉜다.

세부 절차	작업
1. 원천 데이터 선정	• 저작권 확인 • 문장 적절성: 분야, 길이, 분량 • 기술 문제 검토: 수집 작업 적절성
2. 원천 데이터 수집	• 수집 방법 결정 • 수집 기준 설정: 문장 분절, 길이, 기호 처리 등 • 메타데이터 결정 • DB화

● 5.1 원천 데이터 선정

원천 데이터를 선정할 때 가장 먼저 데이터 수요자의 필요에 맞춰 구축 분야를 정한다. 적절한 문장이 있는 데이터원을 온라인 혹은 오프라인에서 찾아 저작권을 확인한다. 저작권 계약이 필요한 경우 원천 데이터 소유자와 구축 목적, 범위에 맞게 계약하여 저작권을 확보한다.

저작권과 더불어 데이터 수요자와 협의한 문장의 길이, 분량을 해당 데이터원에서 획득할 수 있는지 확인하고, 부족할 경우 추가 데이터원을 선정한다. 이때 분량은 구축 목표량의 1.1배 이상 확보되어야 이후 가공 단계에서 어려움이 적다.

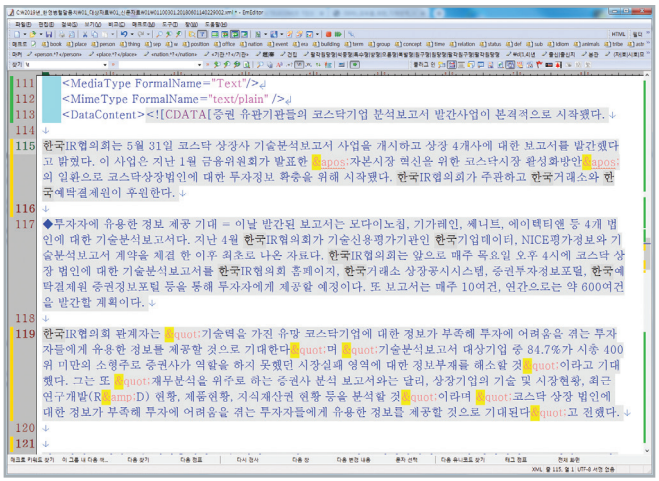
선정 시 수집 기술도 같이 검토가 되어야 한다. 예를 들면, 온라인에서 크롤링해야 하는 경우에는 크롤링 기술이 확보되고, 해당 기술이 효과적으로 적용될 수 있는 형태의 웹사이트인지 확인해야 한다. 원천 데이터가 파일의 형태인 경우에는 해당 파일을 가공하거나, 문장을 추출할 기술이 있는지도 같이 검토가 필요하다.

● 5.2 원천 데이터 수집

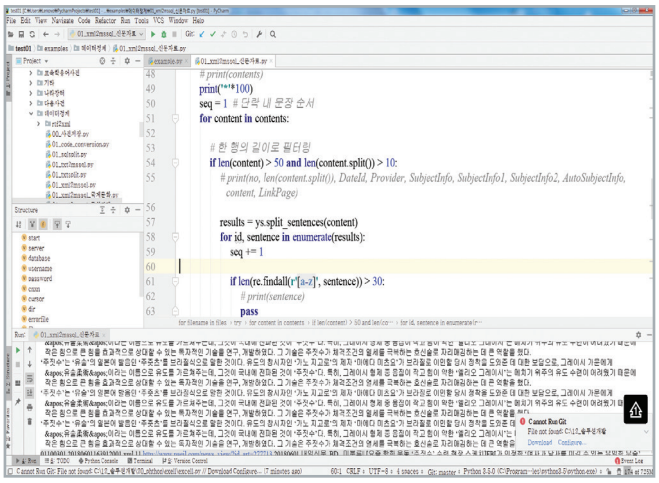
원천 데이터를 선정한 후 데이터 수요자의 확인이 완료되면 이를 수집하는데, 먼저 원천 데이터 형태에 따라 수집 방법을 결정한다. 예를 들어 파일인 경우 문장을 추출하는 방법을 결정한다. 데이터 수요자와 합의한 수집의 기준이 있어야 하는데 수집에 필요한 기준은 문장의 길이와, 여러 가지 기호의 통일 방법 등이 있다. 데이터 구축 목적에 따라 각 분야 별 메타데이터를 정의한다. 이러한 기준에 맞춰 메타데이터와 함께 수집된 문장은 정제 전 데이터베이스(DB)에 저장한다.

(예) 뉴스 분야 데이터 수집(XML 파일 형태)

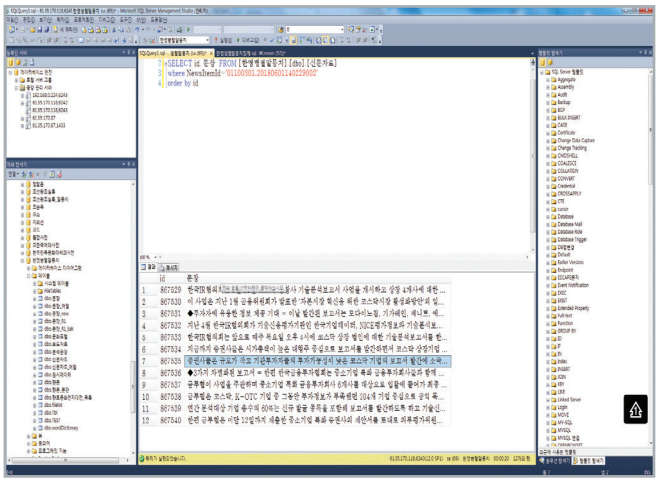
1) 파이썬 lxml 모듈을 이용하여 XML 문서로 된 파일을 파싱



2) 기사를 문장 단위로 분할하여 DB에 저장



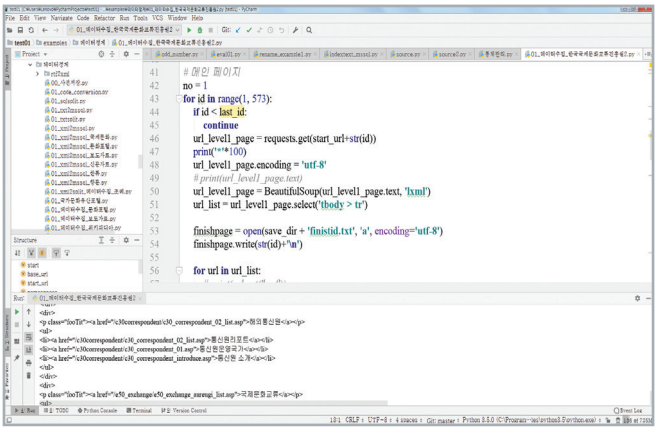
— 문장 단위 변환 결과



(예) 크롤링을 통한 수집 방법 예시

- 파이썬 라이브러리 중 Beautifulsoup와 requests를 이용
- 파이썬을 이용하여 프로그램을 개발할 때 웹 페이지마다 그 구조가 다르기 때문에 웹 사이트별로 최적화된 프로그램이 필요하므로 이를 맞게 개발

(예) 파이썬을 이용한 크롤링 프로그램의 일부



- 게시판 등에서 수집할 때는 페이지 별로 이동하면서 해당 링크 정보를 추출 → 추출한 링크 정보를 이용하여 최종 페이지로 이동 → 해당 페이지를 수집
- 웹 페이지에서 크롤링 할 엘리먼트의 선택은 css selector를 사용
- HTML 문서에서 핵심이 되는 부분을 XML 문서로 저장
- XML 문서를 파싱하여 XML 문서의 구조를 충족시킬 수 있도록 수정
- 문서의 스타일과 체제를 XML 문서에서 직접 확인
- 문장을 추출할 때 대상이 되는 단락과 메타데이터가 정확히 구분이 되도록 체제 정리
- 본문의 문단을 수집하고 수집된 문장을 대상으로 문장 단위로 분할
- 웹 사이트에서 직접 복사한 텍스트와 파이썬을 이용하여 크롤링한 텍스트를 비교하여 텍스트의 누락이나 오류가 있는지 확인

[표 5-1] 실제 구축한 분야 별 메타데이터 예시

구분	문장번호	구분	출처	특성
뉴스	O	분야(3단계)	기사 URL	기사 작성일
지자체 웹사이트	O	지자체	-	-
조례	O	지자체	-	-
한국문화	O	분야	URL, 출판물	-
구어체	O	-	-	-
대화체	O	분류/상황	-	대화세트/화자

메타 데이터는 데이터 수요자와 협의하여 결정하며, 문장 번호의 경우 구축하는 분야 별로 이를 별도로 관리하고, 관리에 용이하게 정수로 이루어진 숫자로 작성한다(예: 1~100,000,000).

● 5.3 특이사항

본 가이드라인의 대화체는 다른 인공지능 기술에도 사용될 것을 고려하기 때문에, 분야 선정 및 협의 시 인공지능 스피커, 챗봇, 음성 인식 등 수요 기관으로부터 최근 수요 요청이 많았던 분야 및 향후 수요가 많을 것으로 예상하는 분야 중 데이터 수요자와 협의한 분야를 대상으로 수집한다. 대화체 세트는 대화를 구성하는 화자의 수를 정하고 이 화자들이 서로 주고 받는 문장으로 구성하기 때문에, 각 세트의 문장 수를 한정하고 각 문장 별 화자 및 세트번호가 구분되어야 한다.

6. 데이터 정제

본 단계에서는 수집된 데이터에 대하여 가공 전, 작업에 용이한 형태로 데이터를 변형하거나, 변형이 어려운 문장을 제외하는 등의 과정을 수행한다.

세부 절차	작업
1. 부적합 데이터 삭제	• 분절 오류 문장 삭제 또는 수정 • 고유명사 처리 기준 미달 문장 삭제(개인정보, 미등록어 과다 등) • 심각한 비문 삭제
2. 데이터 정제	• 번역 부적합 기호 처리(이중 꺽쇠, 중점 등) • 문장 종료 처리(분리, 병합, 마침표 등) • 오타, 인코딩 오류 등 교정

● 6.1 부적합 데이터 삭제

수집한 문장을 정제하기 전에 효율적인 정제를 위해 구축 대상에 적합하지 않은 문장을 먼저 삭제한다. 삭제해야 하는 문장은 대표적으로 도구를 이용하여 문장이 강제로 분절되었을 때 오류가 생긴 경우, 공개/사용하기에 적절하지 않은 개인정보가 있는 경우, 가공이 어려운 미등록어가 한 문장 내에 많은 경우를 예로 들 수 있다. 국문이 비문일 때에도 데이터 가공 단계에서의 오류를 발생하므로 해당 문장을 삭제한다.

(예) 문장 분절이 적절하지 않은 문장

– 광경화성수지투사방식 DLP(Digital Light Processing) 엔비전텍사의 퍼팩토리 장비들은 1997년 미국 텍사스 인스트루먼트의 래리 흔백(Dr.

● 6.2 데이터 정제

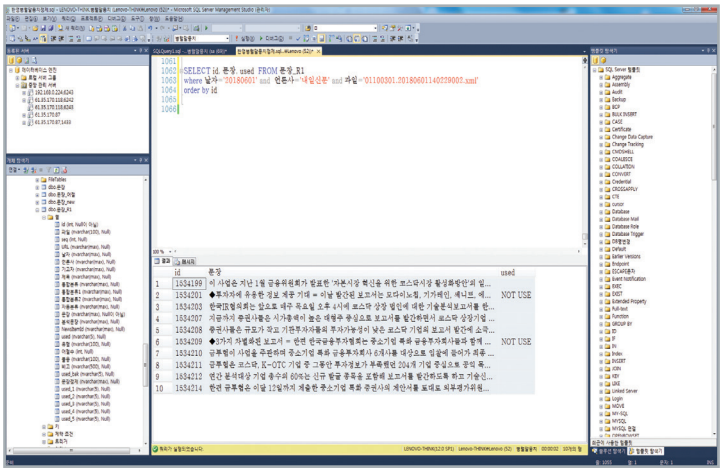
부적합한 문장을 제외하고 나면 일반적으로 중복된 문장을 삭제하고, 번역에 부적합한 기호를 교체하거나 삭제하고, 문장 부호 오류, 오타, 인코딩 오류 등을 수정하는 정제 과정을 거친다. 인코딩의 경우 특히 초기 웹에서 사용하는 인코딩 방식(euc-kr)이 나중에 utf-8로 변경되면서 이전 방식의 웹페이지가 깨지거나, 아예 유니코드 표준을 제대로 표현할 수 없는 인코딩 ks5601 등인 경우 모두 utf-8이나 utf-16으로 변환해주는 과정을 거친다.

(예) 기호 정제 예시

문자	문자 명칭	통일
• 아래아(318D)	HANGUL LETTER ARAEA	·
• (22C5)	DOT OPERATOR	·
• (30FB)	KATAKANA MIDDLE DOT	·
• (2219)	bullet operator	·
• (2022)	bullet operator	·
• (0387)	GREEK ANO TELEIA	·
• (1427)	CANADIAN SYLLABICS FINAL MIDDLE DOTGREEK ANO TELEIA	·
• (2024)	ONE DOT LEADER	·
• (2027)	HYPHENATION POINT	·
~ (FF5E)	FULLWIDTH TILDE	~
(FEFF)	ZERO WIDTH NO-BREAK SPACE	삭제

(예) 문어체 뉴스분야 정제

— 정제 이전 DB 조회 결과

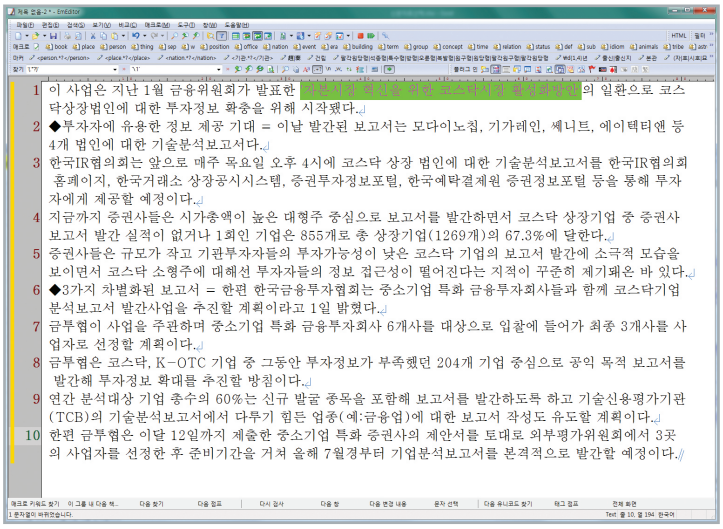


1) 번역으로 적절치 않은 문장을 제외하고, 어절 수 등 기본적인 요건을 충족시키지 못하는 문장은 used 컬럼에 'NOT USE'를 표시

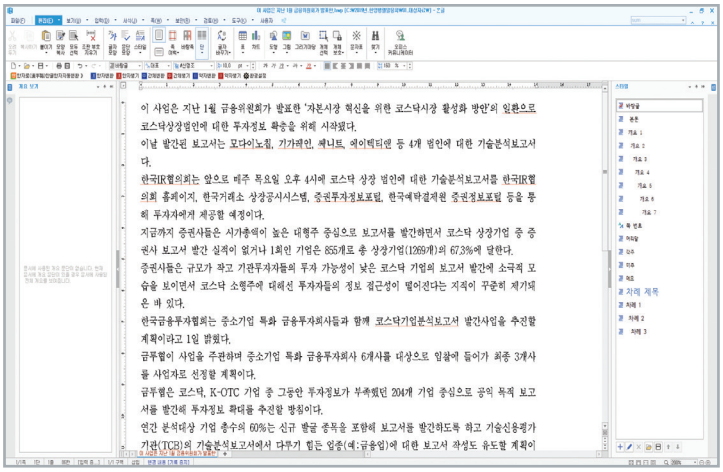
(문장 어절 수 예시)

— 10-30 어절 범위 내 문장으로 제한하되, 구어체, 대화체는 이를 예외로 한다.

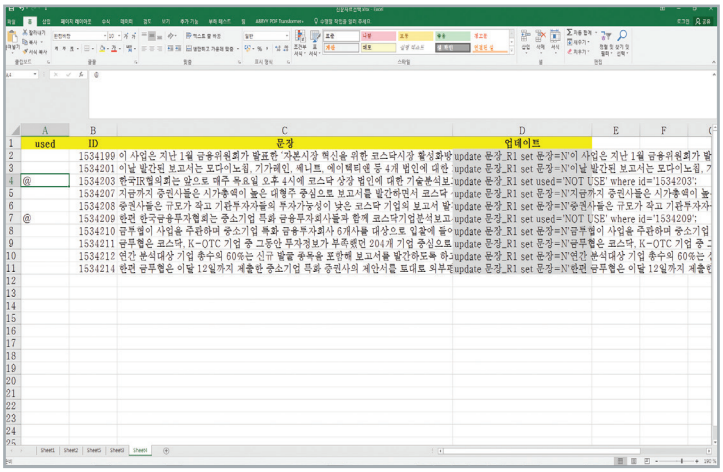
2) 기본적인 요건을 충족시킨 문장을 추출하여 엑셀과 EmEditor, 아래한글에서 정제하고, EmEditor에서 '자본시장 혁신을 위한 코스닥시장 활성화방안' 등의 부호가 제대로 처리되지 않은 것, '◆투자자에 유용한 정보 제공 기대 = ', '◆3가지 차별화된 보고서 = 한편' 등 요약 제목이나 접속사 등을 정제하거나 정제가 어려운 경우 used 컬럼에 사용하지 않음을 표시한 후 엑셀에서 DB에 업데이트



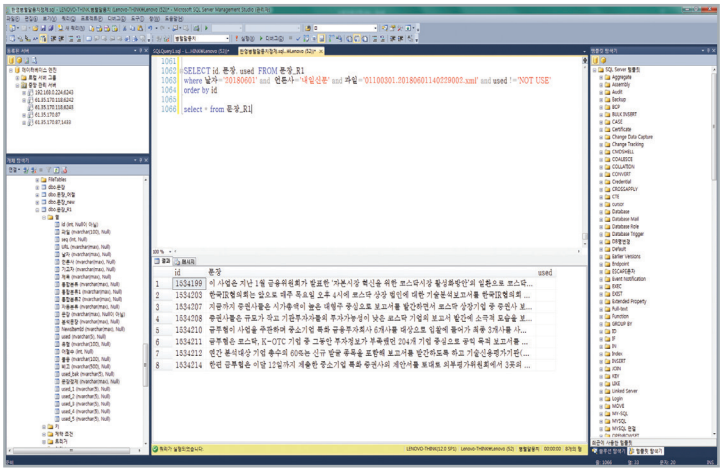
3) 아래한글에서 맞춤법 체크 기능을 활용하여 문장을 이해하는데 지장이 있는 오류를 수정(부산대학교 인공지능연구실과 나라인포테크가 개발한 맞춤법 검사기를 이용)



4) 정제된 결과 문장을 엑셀에서 쿼리로 만들어 DB에 업데이트



— 정제 이후 DB 조회 결과



7. 데이터 가공(번역)

본 가이드라인에서의 가공 방법은 기계번역 결과를 활용하는 기계번역후편집(MTPE)과, 처음부터 사람이 번역하는 2가지로 정한다. 데이터 수요자와의 협의 및 분야 특성에 따라 효율적인 방법을 선택하여 가공한다. 어떤 방법을 선택하더라도 이를 리뷰하는 과정은 반드시 후행되어야 함을 유의한다.

세부 절차	작업
1. MTPE 또는 번역	• 적합한 기계번역(MT) 선택 • MT 후 오류 수정 • MT 품질이 낮은 문장은 분류하여 사람이 초벌 번역
2. 번역물 리뷰	• 리뷰 단계 설정 (1단계 또는 2단계) • 지침에 따른 번역 문장 수정 • (2단계 수정 시) 번역 문장 다듬기 및 문장 대조 확인

● 7.1 MTPE 또는 번역

MTPE 방법을 활용하여 가공하는 경우 분야 별 특성에 따라 알맞은 기계번역을 선택해야 한다. 기계번역의 결과를 놓고 사람이 원문과 대조하여 번역문의 오류를 수정한다. 오류를 수정하는 것뿐 아니라, 이후 구축된 데이터로 NMT 학습 등의 활용을 고려하여 형식적 일관성 등에 대한 작업 지침을 마련하여 이에 맞춰 수정한다. 기계번역에 따라 몇몇 문장은 심각한 오류를 나타내는 경우가 있는데 이때는 사람이 초벌 번역하거나, 정제된 문장 수가 충분하다면 이를 보류하는 등 효과적인 프로세스를 만들어 가공한다.
사람이 직접 번역할 때에는 MTPE와 마찬가지로 일관성이 일정 수준 확보된 형태로 작업을 하는 것이 중요한데, 이를 위한 작업 지침 예시는 아래와 같다.

(예) 문어체 뉴스 분야 작업 지침

1) 단위는 앞 핵심어에 붙여 씀
   (예) 120nm, 120kg
2) 국립국어원 외래어/로마자 표기법 준수
3) (회사명, 모델명) 공식 이름에 따라 작성
4) 원문에 강조나 인용 등을 위해 쓰인 따옴표도 번역문에서 따옴표를 사용하여 작성
5) 원문에 약어와 풀이가 병기되어 있을 때 그 순서와 형식을 유지하여 번역
   (예) 한국정보화진흥원(NIA) → National Information Society Agency (NIA)

(예) 문어체 뉴스 분야 MTPE 문장

- 원문: 금투협은 코스닥, K-OTC 기업 중 그동안 투자정보가 부족했던 204개 기업 중심으로 공익 목적 보고서를 발간해 투자정보 확대를 추진할 방침이다.
- NMT 결과: The association plans to publish public interest reports on 202 companies among KOSDAQ and K-OTC companies, which had lacked investment information so far, to push for the expansion of investment information.
- MTPE 결과: KOFIA plans to expand investment information by publishing reports for public interest on 202 companies, which lacked investment information all these time, among KOSDAQ and K-OTC companies.

7.2 리뷰

MTPE, 번역 방법에 관계없이 앞단계의 결과물을 확인하여 오류를 수정하고 지침에 맞지 않는 경우 이를 적절하게 수정한다. 데이터 공급자는 분야의 특성, 작업 품질에 따라 리뷰 단계를 1단계 또는 2단계로 구성하여 데이터 수요자와 협의된 품질을 확보한다.

(예) 문어체 뉴스 분야 MTPE 문장

- 원문: 금투협은 코스닥, K-OTC 기업 중 그동안 투자정보가 부족했던 204개 기업 중심으로 공익 목적 보고서를 발간해 투자정보 확대를 추진할 방침이다.
- NMT 결과: The association plans to publish public interest reports on 202 companies among KOSDAQ and K-OTC companies, which had lacked investment information so far, to push for the expansion of investment information.
- MTPE 결과: KOFIA plans to expand investment information by publishing reports for public interest on 202 companies, which lacked investment information all these time, among KOSDAQ and K-OTC companies.
- 리뷰 결과: KOFIA plans to expand investment information by publishing reports for public interest focusing on 204 companies, which lacked investment information all these time, among KOSDAQ and K-OTC companies.

8. 데이터 검사

본 단계는 데이터 수요자가 자체 검사까지 완료한 내용에 대하여 제 3 기관이 이를 검사하여 번역 인공지능이 학습을 함에 있어서 데이터가 그 수준을 충족하는지를 확인한다.

세부 절차	작업
1. 외부 검사	• 검사 기관 선정 • 검사 방법/기준 설정 및 테스트 • 번역물 검사
2. 검사 결과 보고 및 조치	• 검사 결과 보고서 작성 및 공유 • 보고서 리뷰 및 조치 방법 합의 • 조치(구축 기관) 및 확인(검사 기관)

● 8.1 외부 검사

자체적으로 리뷰가 완료되었다 하더라도 이를 객관적인 관점에서 검사를 수행한다. 이때, 검사 기관을 선정하고, 검사 기관과 협의된 범위 안에서 방법과 기준을 설정한다. 구축 기준에 대해서는 테스트를 통해 검토하여 현실적이고 효율적인 기준을 정해야 한다. 이후에는 협의된 검사 항목, 방법에 맞춰 검사 기관을 통해 번역물을 검사한다. 가공과 동시에 검사를 병행하여 가공 단계에서의 품질을 동시에 확인한다. 검사 기준 예시는 부속서 A를 참고한다.

● 8.2 검사 결과 보고 및 조치

검사 기관은 검사 결과 보고서를 데이터 공급자와 공유한다. 데이터 공급자는 보고서를 검토하고, 작업물에 대한 추가 조치가 필요한 경우 이를 검사 기관과 합의하여 조치를 이행한다. 조치가 완료된 작업물은 검사기관의 확인을 거쳐 이상이 없음을 확인한다.

9. 데이터 활용

본 가이드라인으로 구축한 한영 번역 데이터는 번역 업무에 활용하거나, 코퍼스 연구에 활용할 수 있다.

세부 절차	작업
1. 번역 업무 활용	• 기존 번역 업무에 번역메모리(TM)로 활용 • 데이터를 활용하여 NMT 학습 • NMT를 초벌번역으로 사용하여 번역 업무 적용
2. 코퍼스 연구 활용	• 단일 언어 연구 활용(자연어 처리, 언어 지능 등) • 이중 언어 연구 활용

● 9.1 번역 업무 활용

데이터 수요자가 번역 서비스 제공회사라면 본 가이드라인에서 구축하는 분야와 동일한 분야의 원문을 한국어에서 영어로 번역할 때에 번역 메모리(TM)로 활용할 수 있다. 데이터 수요자가 자체적으로 개발한 NMT 기술을 보유하고 있거나, 공개되어 있는 NMT 기술을 활용할 수 있는 경우 NMT 엔진을 학습할 때 코퍼스로도 활용할 수 있고, 학습한 NMT 엔진을 초벌 번역으로 사용하여 번역 업무에 적용할 수도 있다.

● 9.2 코퍼스 연구 활용

연구에 활용할 때에는 단일 또는 이중 언어 연구에 활용할 수가 있는데, 단일 언어 연구경우에는 자연어 처리, 언어 지능 등에 활용할 수 있다.
이중 언어 연구 예시로는 본 가이드라인을 통해 구축한 데이터로 학습한 NMT 엔진과, 기존 구글, 파파고 엔진을 비교하거나, 자동번역 챗봇 평가/개발에 활용하는 등이 있다.

부속서 A

품질 기준 명세서 예시

A.1 목적

본 부속서의 목적은 한국어-영어 번역 말뭉치 데이터 구축 품질 평가 기준을 정의하여 구축 시 발생할 수 있는 오류를 최소화하고, 99.9% 이상의 품질 보장을 확보하기 위한 품질 기준을 제시하는 것이다.

A.2 오류의 종류

A.2.1 명백한 문장 오류

- 원문과 명확히 다른 문장

  (예) 2001년 맥도널드, 버거킹, P&G등 감자를 대량으로 이용하는 기업들은 소비자들의 반발을 고려해 GM 감자를 이용하지 않겠다고 발표했고, 몬산토는 결국 그해 3월 해당 GM감자의 상업화를 철회한 바 있다.
  → In 2001, he married a businessman, but he disappeared in 19 months and in 2002, he married businessman who is 11 years older and divorced in 2013.

- 문장이 잘리거나 미완성인 경우

  (예) Type 1 diabetes is diabetes that does not make enough insulin and is distinguishable from type 2 diabetes, where cells.

- 명확한 오역인 경우

  (예) IEA는 태양광발전의 LCOE가 큰 폭으로 줄어든 이유로 초기투자비 감소를 꼽았다. → The IEA cited initial investment costs as a reason for the large amount of LCOE.

A.2.2 수치 오류

(예) APR1400 → APR1440

A.2.3 지나친 의역

- 원문과 번역문의 문장 길이 차이가 심할 정도의 누락이나 추가는 지양한다.

A.2.4 맞춤법 및 오탈자

(예) with a problemaic performance -> with a problematic performance

A.2.5 기호 및 부호 오류

- 문장 부호 작성 오류 및 기호 오류

  문장의 맨 마지막은 마침표로 작성한다.

A.3 품질 순도율 산정 기준

- 총 순도율 = Σ문장별 순도율 / 문장 수 X 100(%)
- 문장별순도율 = 1 - S감점지수 (단, 문장별 순도율 ≥ 0)

[표 A-1] 품질 기준 항목

오류 유형	감점 지수	비고
명백한 문장 오류(오역, 누락)	1.0	
숫자 오류	0.3	
지나친 의역	0.2	
맞춤법, 문법 오류	0.1	복수 감점 적용
기호, 부호 오류	0.1	복수 감점 적용

02

한국형 사물 이미지 AI 데이터 구축 가이드라인

	성명	소속	직위
가이드라인 제안	한윤기	㈜미디어그룹사람과숲	대표
가이드라인 작성	류동철	㈜미디어그룹사람과숲	부장
	강진규	㈜인피닉	부장
	이종민	㈜솔트룩스	과장
전문가 검토 위원	장하영	써로마인드로보틱스	이사
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 사람과숲/ 참여: 인피닉, 솔트룩스

서 문

1. 가이드라인의 목적

본 구축 지침의 목적은 한국형 사물 이미지 학습용 데이터 구축을 위한 객체 및 속성정보를 취득하여 정보이용자(산업계, 학계 및 연구소)가 연구개발에 쉽고 효율적으로 활용할 수 있는 데이터를 제작하고자 함이다. 이를 위해 데이터의 설계를 기반으로 필요한 이미지 데이터를 직접 촬영 등을 통해 수집하고 온톨로지 데이터 사전을 통한 이미지 의미 확장, 그리고 어노테이션 도구를 활용한 메타데이터 입력 및 레이블링을 통해 고품질의 데이터를 확보할 수 있는 가이드를 제시하는 것이 목적이다.

2. 주요 내용 요약

본 구축 지침의 주요 내용은 데이터 설계 및 수집 방법, 데이터 정제, 의미 확장 방법, 데이터 가공 방법, 데이터의 품질 확보 방법으로 구성되었다. 이 중 데이터 수집을 위한 촬영 장비 및 방법을 자세히 제시하였고 데이터의 정제를 위한 유사도 이미지 분류 및 중복제거를 위한 솔루션과 데이터 의미 확장을 위한 온톨로지 검색 활용 방법을 제시하고 있다. 또한 이미지에 대한 메타데이터 정보의 효율적인 입력과 편집을 위한 어노테이션 도구의 활용 방법에 대한 지침을 제시하고 있다. 끝으로 각 지침을 기반으로 구축된 데이터에 대한 검수 절차와 방법, 목표기준율을 정의함으로서 고품질 데이터 확보 방안 또한 제시하였다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

한국형 사물 이미지 AI데이터 구축 지침

Development of a database
for korean type object detection image

02

1. 적용 범위

본 지침은 한국형 사물이미지 인공지능 데이터를 제작하는 절차와 구축 방법을 제시하며, 공정별 구축 표준을 제시한다. 이미지의 수집 중 촬영 방법은 인공지능 알고리즘이 대상체를 인식할만한 수준의 유의미한 이미지 수집 방법을 제시한다.
또한 온톨로지 데이터 사전을 연계한 의미 확장 방법은 본 사업에서 제안된 ADAM 온톨로지 사전[1)]을 기준으로서 지침을 제시한다. 인공지능 학습을 위한 어노테이션 제작 역시 어노테이션 도구를 활용한 입력 기준을 제시한다.

2. 인용 표준

- 해당 사항 없음

3. 용어 정의

● 3.1 스키마

— 논리적인 데이터 베이스 레코드를 매핑하는 데 사용되는 정의 정보
— 클래스, 프로퍼티, 제약조건으로 구성됨

● 3.2 클래스

— 공통되는 것들을 묶어서 대표적인 이름을 붙인 것
— 인스턴스가 해당되는 분류를 의미
예) 버락 오바마(인스턴스) → 사람(클래스)

1) Tony Lee et al, Brochette: Toward A Continuous Learning Platform for Knowledge Acquisition and Integration, Association for the Advancement of Artificial Intelligence, 2015.

● 3.3 프로퍼티

– 인스턴스를 설명하기 위한 속성 값
예) 이름, 생일, 직업 등

● 3.4 인스턴스

– 클래스와 프로퍼티로 설명되는 의미 객체
예) 버락 오바마(사람), 삼성전자(기업)

● 3.5 SPARQL

– 온톨로지 질의 언어

4. 약어

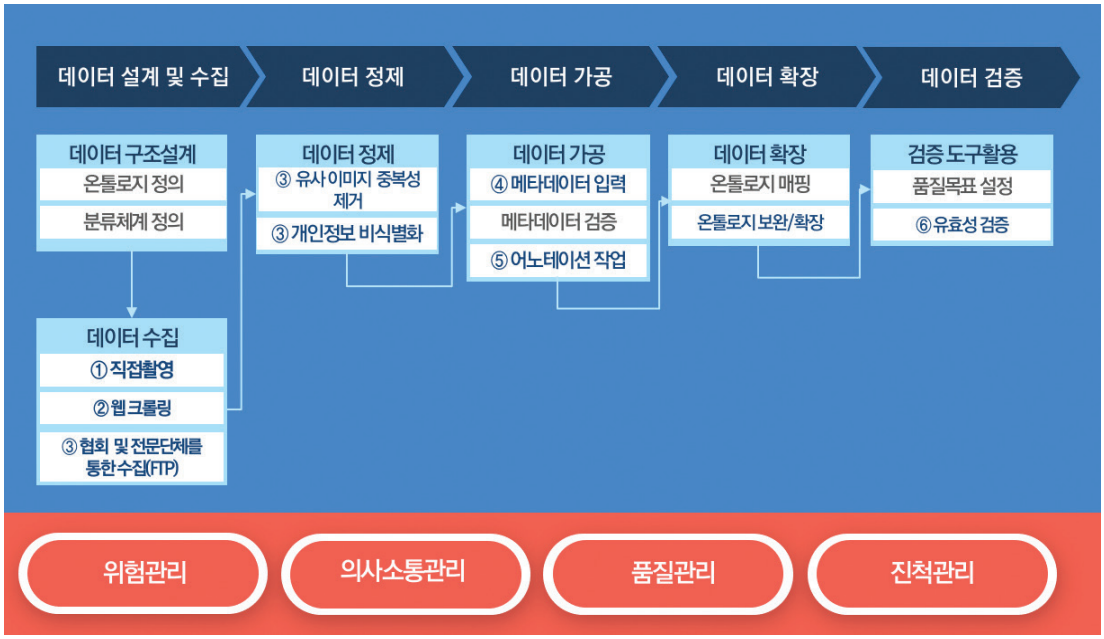
• 해당 사항 없음

5. 인공지능 학습용 데이터 설계

● 5.1 전체공정도

1) 전체 공정도

[그림 5-1]



5.2 분류체계

1) 분류체계 기준 및 목적

[그림 5-2] 분류체계 수립 기준

360만장 이미지 수집방안

- 인공지능 학습을 위한 한국형 사물이미지 종별 수량, 다양성, 품질수준을 만족시킬 수 있는 분류체계 기준 수립

분류기준 수립

종별로 학습가능한 데이터 구축/활용

- 한국형 사물이미지 종별 타케팅을 통한 일정추진이 가능하도록 분류체 계 수립
- 시범서비스 모델 발굴을 통한 활용 가능한 분류체계 수립

분류체계 수립기준

- 종별 수량, 다양성, 품질수준을 만족시킬 수 있는 분류체계 기준 수립
- 종별로 학습가능한 데이터 구축을 위한 분류체계 기준수립
- 활용 가능한 모델 구축을 고려한 분류체계 기준 수립

가) 한국형 사물 이미지 정의 및 범위

- 정의 : 대한민국내에 존재하고 유통되는 한국형 사물에 국한되며 학습용데이터로 적합한 대상체
- 범위 : 국가지정 유적건조물(탑, 성곽, 가옥, 궁궐, 사찰 등), 한국에서 제조되었거나 유통되는 상품(신발, 가방, 지갑, 잡화 등), 한국의 35개 도시 랜드마크

※ 이번과제에서는 국가지정 유적건조물은 서울·경기·인천·경주에 한정함

2) 분류체계 구성

가) 분류체계는 3개 분야로 구분

[표 5-1]

분야	설명	분류기준
유적 건조물	서울/경기/인천/경주에 위치한 국가지정문화재	종교신앙, 정치국방을 포함한 15개 분야
상품	외국관광객 쇼핑구매 우선순위 고려	신발, 가방, 화장품, 악세서리 포함 8개 분야 28개 아이템
랜드마크	각 도시의 주요 상징물	총 35개 도시 랜드마크 대상체

나) 대상목록 선정

총 대상체수는 약 1,157개로 결정됨

[표 5-2]

구분	대상체 수량	비고
유적건조물	약 840 개	촬영 대상체 공사 중 등의 이유로 수량 변동 가능성 있음
랜드마크	약 66개	
상품	약 251개	

다) 분류체계 목록

[표 5-3] 유적건조물

대분류	중분류	소분류	대상체명
유적건조물	과학기술	천문지리시설	서울 관상감 관천대
	교육문화	교육시설	배재학당동관, 연세대학교 스팀슨관
	교통통신	교통 · 통신시설	서울 우정총국, 창경궁 옥천교, 봉천대
	무덤	무덤	화의군이영묘역, 여계묘역, 방이동고분군
		봉토묘	류순정 · 류홍부자묘역, 이규보묘, 정제두묘
		왕실무덤	영산군이전묘역, 양녕대군이제묘역, 서울정릉
	불교공예	의식법구	선덕대왕신종
	불교조각	금속조	백률사금동약사여래입상, 불국사금동비로자나불좌상
		석조	영지석불좌상, 백운대마애불입상, 보리사마애석불
	산업생산	산업시설	구용산수위관측소, 광통관, 구세군중앙회관
		요업시설	조선백자요지, 남현동요지, 동방동와요지
	서각류	금석각류	중성리신라비, 이씨삼강묘비, 손종로정충비각
	인물사건	역사사건시설	아담스기념관, 사육신묘, 삼전도비
		인물기념시설	안창호묘소, 최규하가옥, 홍파동홍난파가옥
	일반조각	암벽조각	석장동암각화, 안심리암각화
	정치국방	궁궐·관아	창경궁근정전, 경복궁수정전, 창덕궁
		근대정치국방	구서대문형무소, 구러시아공사관, 구벨기에영사관
		성	흥인지문, 숭례문, 몽촌토성
		전쟁관련시설	포천방어벙커, 영국군설마리전투비
	종교신앙	제사유적	선농단, 서빙고동부군당, 강화참성단
		종교시설	흥천사극락보전, 봉은사선불당, 약현성당
		탑	거동사지원공국사탑, 봉림사지진경대사탑
	주거생활	조경건축	탑골공원팔각정, 봉황각, 용양봉저정
		주거건축	도정궁경원당, 김좌근고택, 월계동각심재
	기타	공용시설	청량리역검수차고, 한강철도교
		기타 시설물	방화수류정, 수광리오름가마
		문화집회시설	건국대학교구서북학회회관, 창경궁대온실
		상업시설	구부국원,선린동공화춘, 구대화조사무소
		업무시설	수원문화원, 기상관측소 · 구미국문화원
		의료시설	구용산철도병원본관, 대한의원

[표 5-4] 상품

대분류	중분류	소분류	대상체명
상품	귀금속	귀걸이	골드귀걸이, 실버귀걸이, 피어싱
		반지	패션반지, 골드반지, 실버반지
		목걸이/팬던트	패션목걸이, 골드목걸이, 실버목걸이
	화장품	스킨케어	수분크림, 아이크림, 앰플
		선케어/메이크업	파운데이션, 컨실러, 립글로즈
		클렌징/마스크팩	클랜징폼, 마스크팩, 코팩
		향수	여성향수, 남성향수, 미니어쳐
		남성화장품	남성선케어, 남성스킨, 에프터쉐이브
	시계	패션시계	가죽시계, 메탈시계, 팔찌시계
		기타 시계	목걸이시계, 전자시계, 탁상시계
	신발	운동화/캐주얼화	패션운동화, 하이탑, 키높이운동화
		여성화	플랫슈즈, 슬리퍼샌들, 앵클부츠
		남성화	남성구두, 로퍼, 캐주얼화
		슬리퍼/기능화	일반슬리퍼, 털슬리퍼, 고무신
		아동신발	아쿠아슈즈, 발광운동화, 롤러운동화
	가방	캐쥬얼가방	백팩, 숄더백, 크로스백
		여성가방	클러치백, 토트백, 퀼팅백
		남성가방	서류가방, 맨스백, 노트북가방
		여행용가방	보스톤백, 여행용백팩
		운동용가방	수영쇼퍼백, 테니스라켓가방
		아동가방	캐릭터책가방, 손가방, 미아방지가방
	지갑	여성지갑	장지갑, 중지갑, 반지갑
		남성지갑	장지갑, 반지갑
		기타지갑	카드지갑, 명함지갑, 동전지갑
	모자	패션모자	스냅백, 야구모자, 군모
		아동모자	아동비니, 썬캡, 귀달이모자
	아이웨어	선글라스	여성선글라스, 남성선글라스
		안경테	뿔테, 금속테, 무테

[표 5-4] 랜드마크

대분류	중분류	소분류	대상체명
랜드마크	강릉시	동상/빌딩	오죽헌율곡이이동상, 경포대스카이베이호텔
	경주시	타워	경주타워
	공주시	정자/탑	공산선공산정, 곰나루국민관광지웅비탑
	과천시	센터	국립과천과학관
	광양시	빌딩	이편한세상
	광주광역시	탑/빌딩	518민주묘지상징탑, KDB생명빌딩
	구리시	타워/동상	구리타워, 광개토대왕동상
	군산시	센터	근대역사박물관
	나주시	빌딩	한국전력공사 본사
	대구시	타워/빌딩	83타워
	대전시	탑/타워	수성SK리더스뷰
	목포시	센터	김대중노벨평화상기념관
	부산시	타워/다리/빌딩	용두산부산타워
	서울시	빌딩/동상/타워	63빌딩, 이순신장군상, 서울남산타워
	성남시	전망대	율동공원번지점프
	세종시	센터	정부세종컨벤션센터
	속초시	타워	엑스포타워
	수원시	전망대	광교프라이부르크전망대
	안동시	동상/다리	월영교
	안성시	동상	바우덕이사당조각상
	양산시	타워	양산타워
	여수시	타워/등대	하멜등대, 여수엑스포역스카이타워전망대
	용인시	타워/센터	아프리아타워, 백남준아트센터
	울산시	전망대/센터	울산대교전망대, 이안태화강엑소디움
	인천시	다리/타워	인천대교, 송대G타워, 경인운하아라타워
	전주시	성문/타워	풍남문, 코오롱스카이타워
	제주시	봉우리/빌딩	성산일출봉, 롯데시티호텔제주
	진주시	전망대	촉석루
	창원시	천/다리/빌딩	저도연륙교, 여좌천
	천안시	탑/빌딩	독립기념관겨레의탑, 펜타포트
	청주시	동상/빌딩	용두사지철당간, 초정약수동상
	춘천시	동상/다리/빌딩	소양강처녀상, 공지천구름다리
	파주시	동상	임진각평화누리공원 통일바라기조각상
	포항시	등대/빌딩	호미곶등대, 장성두산위브더제니스
	화성시	빌딩	메타폴리스

5.3 데이터셋

1) 데이터셋

가) 데이터셋 구성

데이터셋 구성은 아래와 같은 형태를 지님

구분	필드	속성
image	identifier	파일명
	imsize	이미지 파일 크기
	copyright	이미지 저작권 소유자
	date	데이터 취득 일자
	resolution	해상도
	bit	비트값
	F-Stop	조리개 투과량
	exposure time	노출 시간
	ISO	ISO 감도
	focal length	초점 거리
	full aperture	조리개 최대 개방 수치
	view angle	화각
	white balance	화이트 밸런스
	depth	이미지 mode(RGB, Grayscale, bitmap)
regions	type	어노테이션 종류
	boxcorners	어노테이션 좌표 값
	ansize	이미지 영역 사이즈
	class	클래스명
	tags	분류항목
	truncated	대상체 잘림 여부
	종ID	종 아이디 쩡보
	대분류	대분류 정보
	중분류	중분류 정보
	소분류	소분류 정보
	Instance	대상체
	Instance Upper	대상체 상위 레벨
	GPS 위도	GPS 위도 값
	GPS 경도	GPS 경도 값
	instance_uri	온톨로지 uri 정보
	sem_ext	의미확장 정보
	property : locatedIn	의미확장 지역정보
	property : relatedTerm	의미확장 연관정보
	property : description	의미확장 상세정보

6. 인공지능 학습용 데이터 수집/정제

● 6.1 원천 데이터 확보 방안

- 원천 데이터는 직접촬영을 기본으로 하며, 환경조건 및 수집 불가능한 항목에 대한 보완책으로 크롤링, 기존 크라우드 플랫폼 수집을 통한 데이터는 사진가 협회를 통해 제공 받음
- 수집 데이터는 장축 800 픽셀 이상, 촬영 데이터는 Full HD 장축 1920픽셀 이상 확보

가) 직접 촬영을 통한 확보

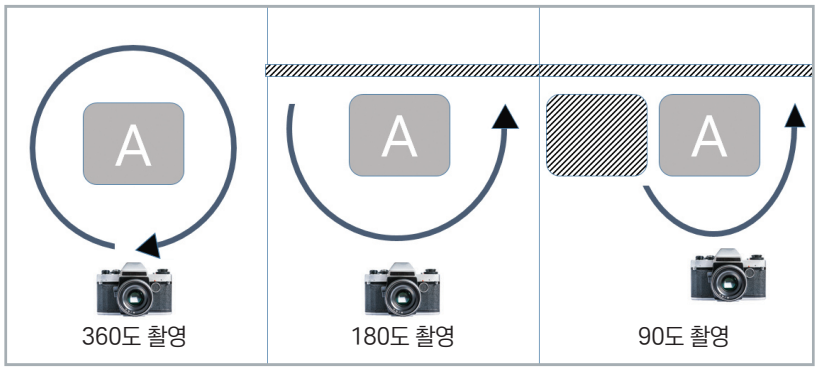
- 직접 촬영은 유적/문화재 전문 촬영 경험 인력 또는 사진전공 인력을 투입하여 사진촬영의 전문성 확보.
- 촬영 각도별 수량을 다르게 촬영진행. 예를들어 360도, 180도, 90도로 촬영각도를 구분 하여 각 각도별 수량범위를 정하여, 수량 일관성 유지

[그림 6-1]

직접 촬영을 통한 수집		
유적건조물	상품	랜드마크
		
• GPS 장착 카메라 촬영 • 상품 촬영을 위한 오브젝트 스튜디오 셋팅 • 촬영자 교육 및 분야별 촬영 진행		

(1) 유적/건조물 촬영 방안

]그림 6-2]



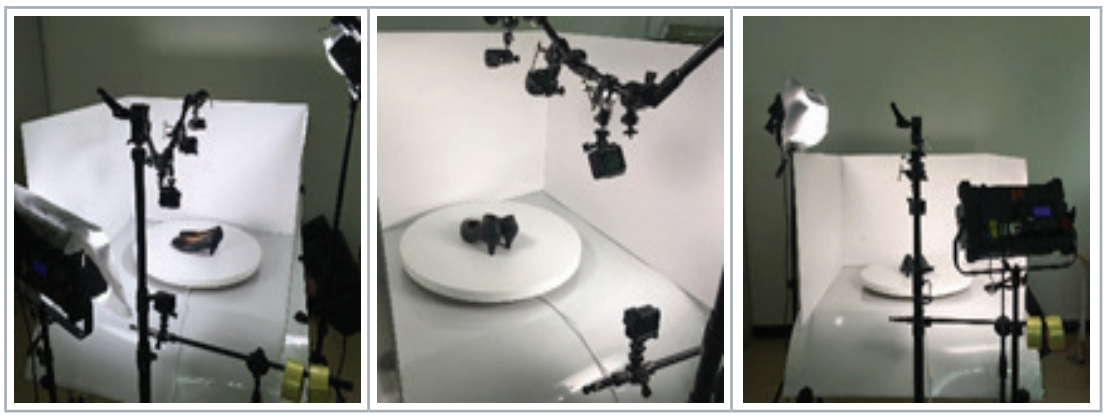
- 학습용 이미지 데이터의 확보를 위하여 다양한 촬영높이와 촬영거리를 기준으로 다량의 이미지를 획득하기 위하여 대상체를 중심으로 촬영 가능한 각도 내에서 이동하며 촬영
- 이미지 인식률을 향상을 위해 건축물의 경우, 촬영시간대 등 다양한 주변환경에 맞는 촬영 일정계획 수립
- 고프로 4대를 각기 다른 수직 각도로 위치시킨후, 사람 크기 기준으로 1번카메라 약 1.8m 위치(남성), 2번카메라 약 1.6m 위치(여성), 모니터카메라 약 1.7m 위치, 3번카메라 약 1.4m 위치(아동) 모노포드에 고정 후 사람이 들고 대상을 360도 도보로 촬영

[그림 6-3] 유적건조물 촬영

(2) 사물(상품) 촬영 방안

- 고프로 4대를 각기 다른 수직 각도로 위치시킨 후 턴테이블 위에서 돌아가는 상품을 타임랩스 방식으로 촬영
- 턴테이블이 한 바퀴 도는 시간이 약 2분(120초) 소요 되므로 2초 간격으로 촬영하여, 한 개의 세팅 값에서 촬영

[그림 6-4] 오브젝트 촬영 전용 장치



(3) 랜드마크 촬영 예시

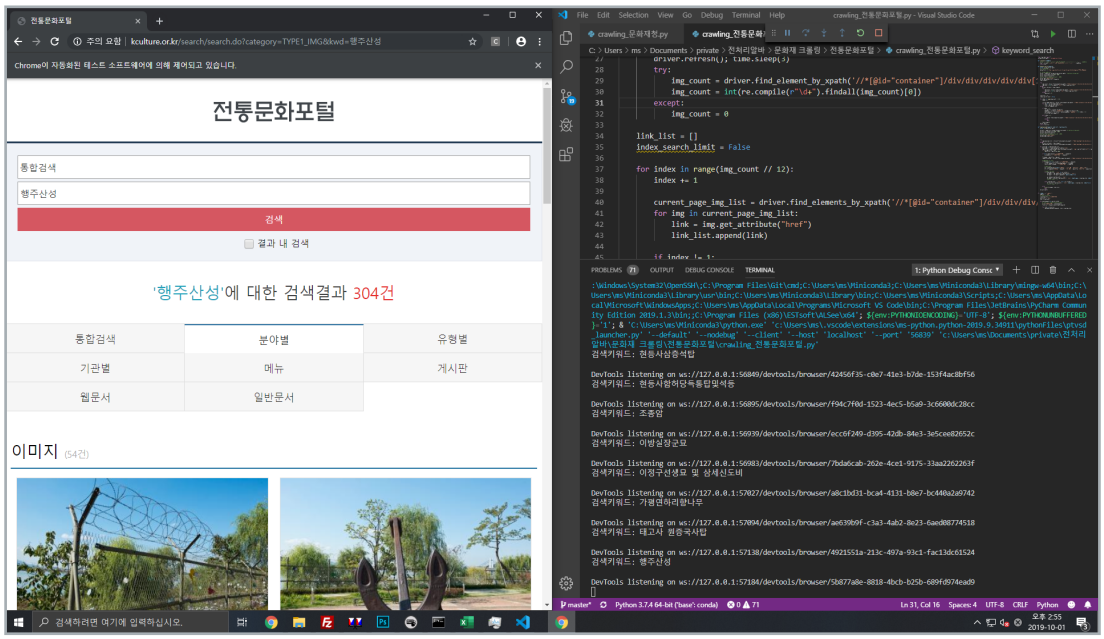
- 건물이나 상징조형물의 360도 외관에서 다양한 각도와 앵글로 촬영(단 예외적인 것은 360도가 불가능한 것은 다양한각도와 앵글로 촬영, 드론촬영도 고려)

[그림 6-5] 랜드마크 촬영

나) 웹 크롤링 방법을 통한 수집

- 자체보유 크롤링 Tool 활용하여 이미지 수집

[그림 6-6] 크롤링 도구 예시



(1) 기능 및 내용

[표 6-1]

구분	내용
수집 기술	Selenium-Chrome 크롤링(자동수집) 기술기반
수집 조건	이미지 활용동의가 필요한 경우 동의 후 수집 저작권 이용가능 범위 대상 이미지 수집 (저작권 이용가능범위 벗어나는 이미지 제외)
수집 내용	이미지 파일, 이미지 제목 개별 이미지 저작권자
다운로드	링크 포털사이트와 비밀번호로 보호된 사이트의 다운로드 기능
	동시에 30개의 파일 다운로드 가능
미리보기	썸네일과 슬라이드쇼 방식으로 미리보기 기능 내장.
동시설정	다수의 URL 동시 설정 가능
기타	사용자가 설정한 URL에서 월페이퍼, 디지털 영상, 사진, 음악, 비디오 등 다양한 형태의 데이터를 크롤링 할 수 있음
	사용자 설정을 통한 목적화 된 데이터 수집 가능 ※ 이미지 사이즈, 파일명변경, URL 필터 등.

(2) 운영 프로세스

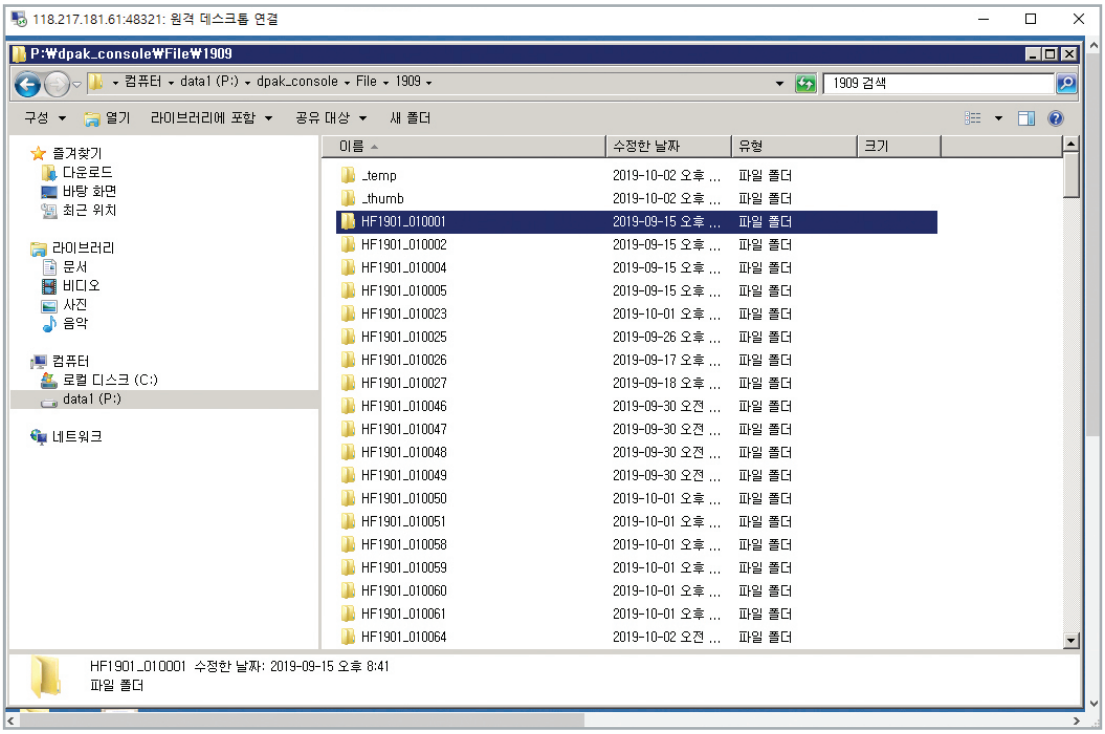
- 수집 프로세스
  - 검색 키워드 정제 (띄어쓰기 수정 및 지역명 제거)
  - 각 개별사이트 수집자동화 구축 (속도보단 안전성 우선순위)
  - 키워드 검색결과항목이 중복인 경우 수기로 지정 후 데이터수집
  - 조건에 맞는 데이터 모두수집 (이미지, 이미지제목, 해당 문화재코드, 저작권자)
  - 개별 검색키워드마다 이미지파일 분류
  - 각 키워드마다 발생되는 중복 이미지 및 해당 키워드조건에 맞지 않는 데이터 제거
  - 개별 사이트 수집완료 및 검수후 데이터 종합확인

※ 웹크로링 타게팅 대상은 사용권이 해결된 한국문화정보원, 한국저작권위원회, 특허청, 국가기록원 등 공공포털내의 이미지로 설정하고 관련 기관과 협의하여 진행

다) 협회 및 전문단체를 통한 수집(FTP)

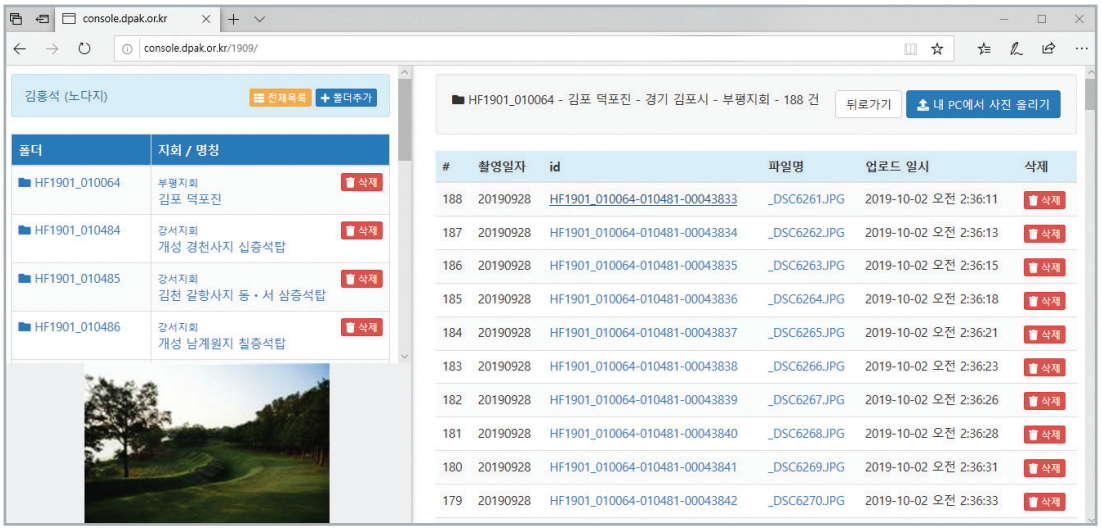
(1) 협회 및 전문단체 FTP서버

[그림 6-7]

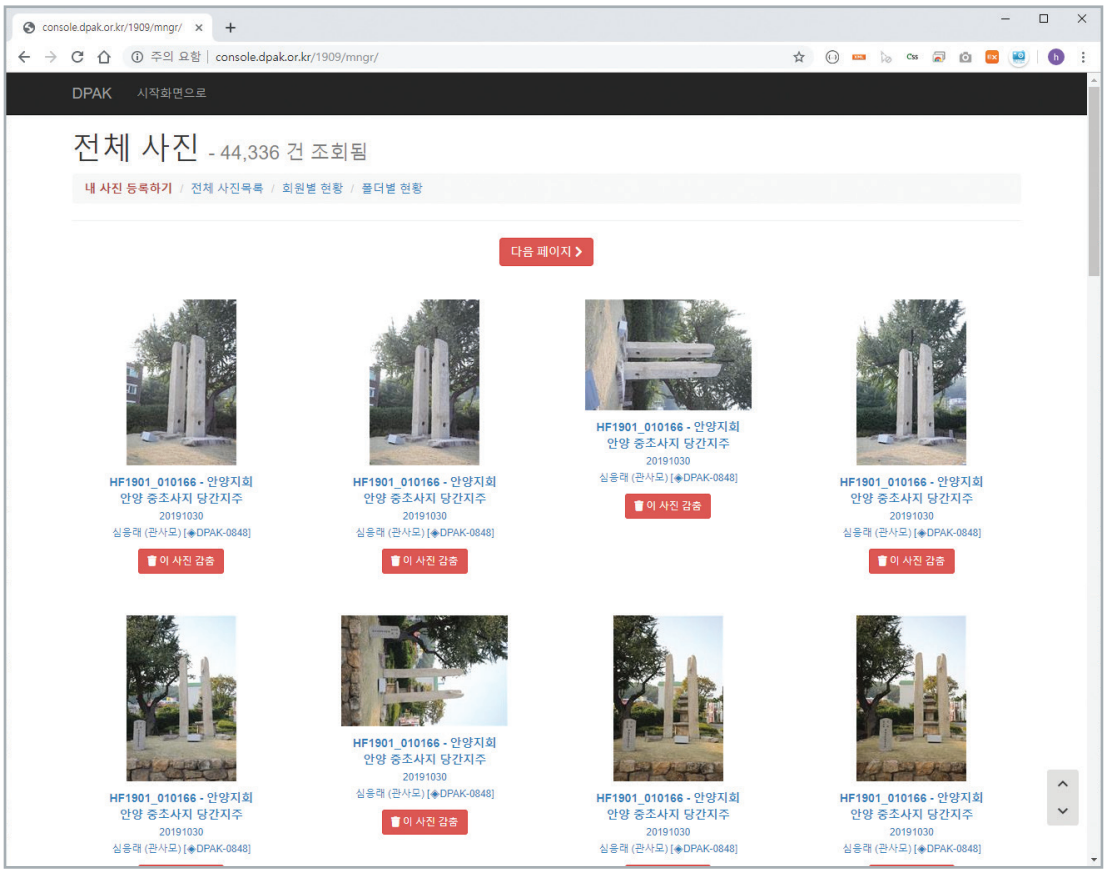


(2) 협회 및 전문단체 회원망을 통한 수집

[그림 6-8]



[그림 6-9]

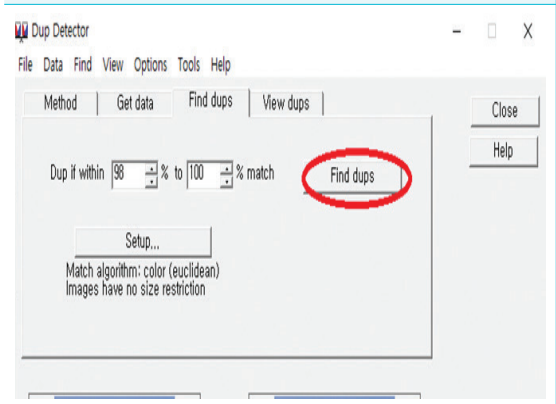
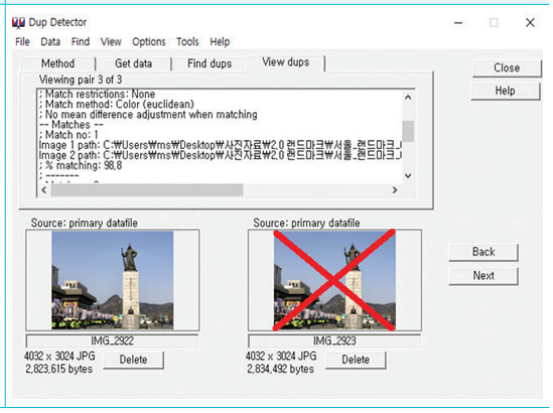


6.2 원천 데이터 정제 방안

가) 유사한 수집 이미지의 자동 태깅을 통한 후보 이미지 확보

- 이미지 유사도 분류 Tool을 활용하여 카테고리 분류 기준 후보 이미지 확보 진행

[그림 6-10]

이미지 유사도 분류 실시	유사 이미지 중복성 제거
	

(1) 기능

[표 6-2]

구분	내용
분류/구별	• 검색 필터 레벨 설정 가능 : 완전 동일한 사진부터 유사한 사진까지 분류 가능 ※ 유사도 99% 이상 범위를 설정하여 정제 진행
	• 다수의 저장소 대상 동시에 유사도 분류 가능
선택/삭제	• 중복사진 자동 선택 및 삭제 가능
이동	• 특정 카테고리 내 유사도 분류 시 대량의 후보사진을 원하는 위치로 이동 가능

[그림 6-11]




유사 이미지 ○　　중복 이미지 ×



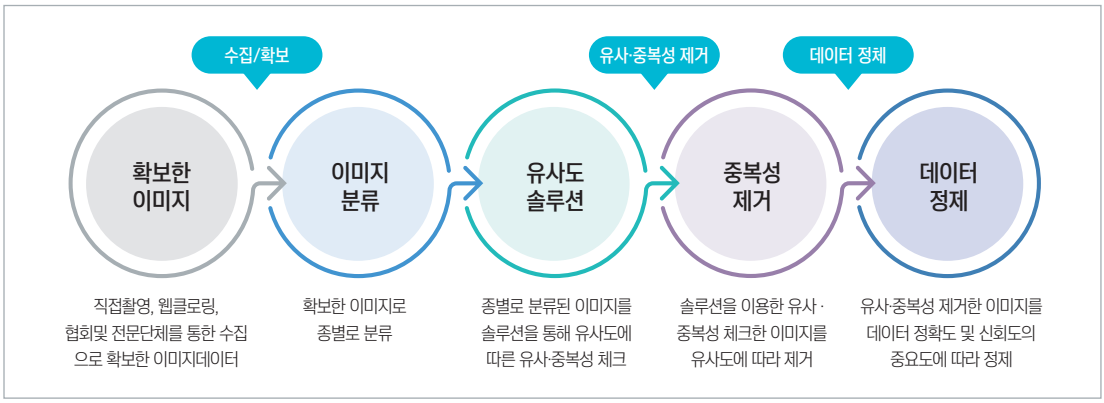


카테고리 내 오류이미지

(2) 유사이미지 간 중복성 제거 프로세스

- 프로그램 실행 : Dup Detector

[그림 6-12]



- 결과 확인 : 데이터 정제 인원이 Dup Detector 사용시, 작업완료 메시지 반드시 확인후 별도 파일목록 및 폴더 정리

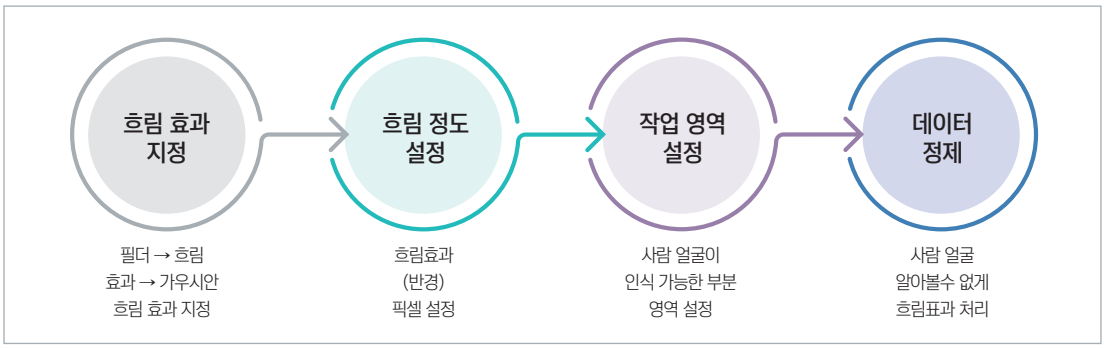
※ 작업완료 메시지 ※

- Consider a duplicate if match is between 99.0% and 100,0%.
- Matches (No duplicates)

(3) 개인정보 비식별화 작업 프로세스

- 프로그램 실행 : 포토샵

[그림 6-13]



- 결과 확인 : 데이터 정제 인원이 포토샵 사용시, 작업완료후 흐림정도 픽셀 값 및 사람얼굴 인식여부 육안 확인후 별도 파일목록 및 폴더 정리

나) 개인정보 비식별화

- 비식별화 범위 : 인물 얼굴, 자동차 번호판
- 개인정보에 해당되는 얼굴 부문, 자동차번호판 흐림 효과(blur)를 통해 비식별화

[그림 6-14]

인물얼굴 비식별화	인물얼굴 비식별화 확대
	

[그림 6-15]

자동차번호판 비식별화	자동차번호판 비식별화 확대
	

다) Annotation

- 한국형 사물이미지 구축 대상 객체에 대한 Bounding box 처리
- 객체 별 속성정보 입력, 산출물 형태는 Json 파일로 처리
- 메타데이터 오류 발생 예방을 위하여 일부 속성값 자동입력으로 도구 개발

[그림 6-16]

메타데이터 입력

<메타데이터 입력 - 유적건조물 예시>

Property	Value
Truncated	0
copyright	미디어그룹사람과...
date	2019-09-25
resolution	72 DPI
bit	24
F-Stop	F/2.8
exposure time	1/1159초
ISO	100
focal length	3mm
full aperture	2.97
view angle	15
white balance	Auto
depth	3
종ID	HF019141
대분류	유적건조물
중분류	무덤
소분류	무덤
Instance	경주구정동방형분
Instance Upper	
GPS 위도	35; 46; 42.31415989...
GPS 경도	129; 17; 48.9807600...
instance_uri	

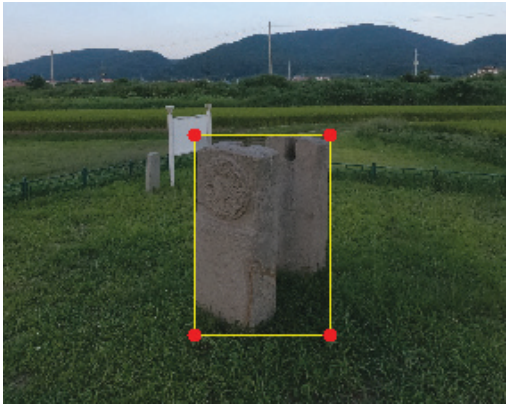
<메타데이터 입력- 상품 예시>

Property	Value
Truncated	0
copyright	미디어그룹사람과...
date	2019-09-25
resolution	72 DPI
bit	24
F-Stop	F/2.8
exposure time	1/55초
ISO	100
focal length	3mm
full aperture	2.97
view angle	15
white balance	Auto
depth	3
종ID	HF022055
대분류	상품
중분류	신발
소분류	남성화
Instance	컴포트화
Instance Upper	
instance_uri	

[표 6-3] Json 파일 속성 정보

속성 구분	속성값	속성값 상세	입력 방안
환경속성	유적건조물 & 상품 분류구분	대분류, 중분류, 소분류, 간략분류	폴더Text 추출입력
객체속성	종 표제/ID	1046 종	폴더Text 추출입력
	Class	유적건조물: 현판 등 상품: 브랜드명 글자 등 거리: 간판상호 등	Text 입력
	Instance	Instance 분류명	폴더Text 추출입력
	Instance Upper	Instance 상위레벨	폴더Text 추출입력
	Instance uri	객체에 대한 의미확장 정보	온톨로지 uri 입력
	인식좌표	(x,y) 좌표	Annotation Tool 좌표설정
이미지 데이터 속성정보	이미지 ID	KIMG_Location/종ID_YYMMDD	이미지데이터 파일명 추출
	환경생성이력	YYMMDD	이미지데이터 속성 추출
	기록 유형	.jpg/.gif	이미지데이터 속성 추출
	이미지 크기	800×600(SVGA), 1920×1080(FHD) 등	이미지데이터 속성 추출
	GPS	위도, 경도 정보	이미지 데이터 속성 추출
	저작권 정보	유(저작권 정보 입력), 무	Text 입력

[그림 6-17]




```
{
  "image": {
    "identifier": "HF019094_0201_0153.JPG",
    "imsize": [
      1,
      1
    ],
    "copyright": "미디어그룹사람과숲(컨)",
    "date": "2019-09-19",
    "resolution": "72 DPI",
    "bit": "24",
    "F-Stop": "F/2.8",
    "exposure time": "1/962초",
    "ISO": "386",
    "focal length": "3mm",
    "full aperture": "2.97",
    "view angle": "15",
    "white balance": "Auto",
    "depth": "3"
  },
  "regions": [
    {
      "type": "box",
      "boxcorners": [
        839,
        772,
        1167,
        1253
      ],
      "ansize": [
        1,
        1
      ],
      "class": "당간",
      "tags": [
        "truncated:0",
        "종ID:HF019094",
        "대분류:유적건조물",
        "중분류:종교신앙",
        "소분류:종교시설",
        "Instance:보문사지연화문당간지주",
        "Instance Upper:",
        "GPS 위도:35; 49; 43.9129199000018",
        "GPS 경도:129; 15; 11.8897199999774"
      ],
      "instance_uri": "",
      "sem_ext": []
    }
  ]
}
```



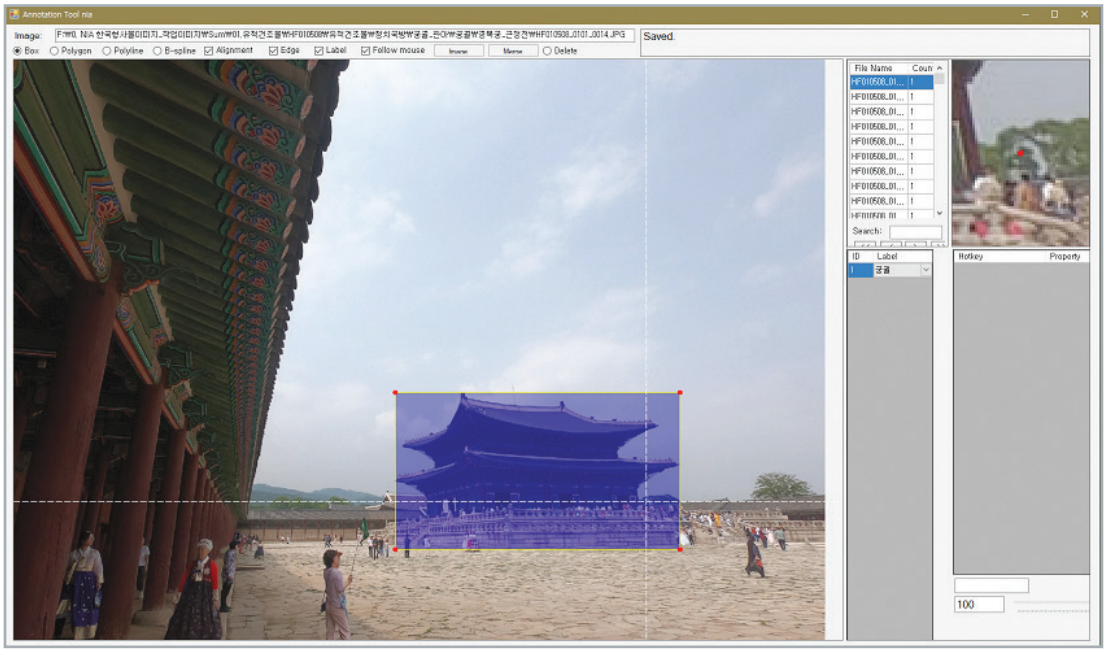
```
{
  "image": {
    "identifier": "HF021056_0101_0020.JPG",
    "imsize": [
      1,
      1
    ],
    "copyright": "미디어그룹사람과숲(컨)",
    "date": "2019-09-16",
    "resolution": "72 DPI",
    "bit": "24",
    "F-Stop": "F/2.8",
    "exposure time": "1/51초",
    "ISO": "100",
    "focal length": "3mm",
    "full aperture": "2.97",
    "view angle": "15",
    "white balance": "Auto",
    "depth": "3"
  },
  "regions": [
    {
      "type": "box",
      "boxcorners": [
        424,
        221,
        1407,
        1086
      ],
      "ansize": [
        1,
        1
      ],
      "class": "남성캐주얼화",
      "tags": [
        "truncated:0",
        "종ID:HF021056",
        "대분류:상품",
        "중분류:신발",
        "소분류:남성화",
        "Instance:샌들",
        "Instance Upper:"
      ],
      "instance_uri": "",
      "sem_ext": []
    }
  ]
}
```

7. 인공지능 학습용 데이터 가공/검증

● 7.1 Annotation 개요/내용

- 사람이 대상체를 인식하는 기준으로 Annotation
- 대상 객체의 명확한 인식을 위하여 객체별로 직접 Bounding box 처리
- 인식률 확보를 위하여 대상 객체의 Annotation 크기는 25 * 25 Pixel을 기준
  : 이미지 내 사람이 식별할 수 있는 기준 사이즈(정확한 Annotation 가능 여부 기준)
- Bounding box 처리 된 객체에 학습에 필요한 메타데이터 속성 부여
- 정확한 Annotation 작업을 위한 가이드라인(X,Y축) 및 View Focus 추가
- 작업의 효율성을 위한 기본 메타 속성 자동입력 기능 구현

[그림 7-1]



● 7.2 Annotation 프로세스

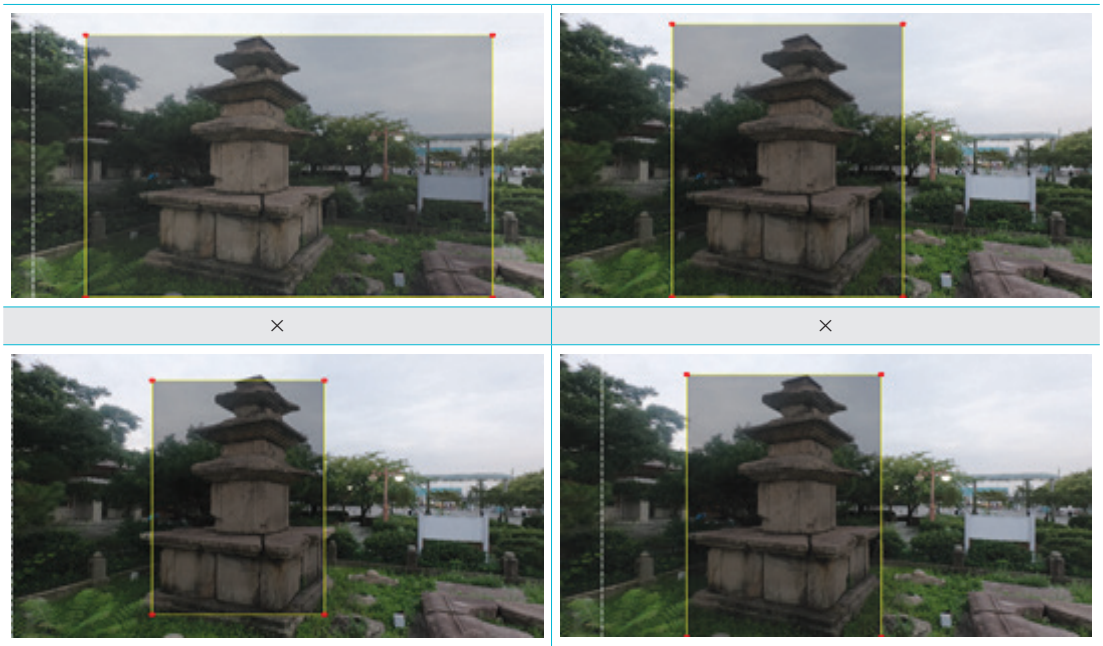
본 내용에서는 Annotation 작업을 진행함에 있어 사전 고려되어야 하는 Risk를 고려하여, 작업 간 해당 Risk들이 자연적으로 대응될 수 있는, 효율적인 프로세스를 안내합니다.

○ 7.2.1 Annotation 방법

- Annotation의 목적이 객체에 대한 Detection 이므로 Bounding box로 진행
- 주요 속성의 경우 객체의 분류체계 기준 폴더 구조를 통하여 자동입력
- 대상 객체의 외곽 기준 Box 처리(사람이 대상체를 인식하는 기준)

— 객체 외곽 기준 Box가 안쪽으로 들어가게 Box를 그리면 안됨
— 객체 외곽 기준은 최대한 타이트하게 box처리 (약 2 Pixel)
— 객체 외의 사물이 함께 box 처리 되면 안 됨

[그림 7-2]



×	×
×	○

no			항목명	설명	필수여부		
					유적 건조물	상품	랜드 마크
1			**image**	**이미지 정보**	Y	Y	Y
	1-1		identifier	파일명	Y	Y	Y
	1-2		imsize	이미지 파일 크기	Y	Y	Y
	1-3		copyright	이미지 저작권 소유자	Y	Y	Y
	1-4		date	데이터 취득 일자			
	1-5		resolution	해상도	Y	Y	Y
	1-6		bit	비트값	Y	Y	Y
	1-7		F-Stop	조리개 투과량			
	1-8		exposure time	노출 시간			
	1-9		ISO	ISO 감도			
	1-10		focal length	초점 거리			
	1-11		full aperture	조리개 최대 개방 수치			
	1-12		view angle	화각			
	1-13		white balance	화이트 밸런스	Y	Y	Y
	1-14		depth	RGB 여부	Y	Y	Y

no			항목명	설명	필수여부		
					유적 건조물	상품	랜드 마크
2			**regions**	**관련 정보**	Y	Y	Y
	2-1		type	어노테이션 종류	Y	Y	Y
	2-2		boxcorners	어노테이션 좌표 값	Y	Y	Y
	2-3		ansize	이미지 영역 사이즈	Y	Y	Y
	2-4		class	클래스명	Y	Y	Y
	2-5		tags	분류항목	Y	Y	Y
		2-5-1	truncated	대상체 잘림 여부	Y	Y	Y
		2-5-2	종ID	종 아이디 정보	Y	Y	Y
		2-5=3	대분류	대분류 정보	Y	Y	Y
		2-5-4	중분류	중분류 정보	Y	Y	Y
		2-5-5	소분류	소분류 정보	Y	Y	Y
		2-5-6	Instance	대상체	Y	Y	Y
		2-5-7	Instance Upper	대상체 상위 레벨			
		2-5-8	GPS 위도	GPS 위도 값			
		2-5-9	GPS 경도	GPS 경도 값			
	2-6		instance_uri	온톨로지 uri 정보	Y	Y	Y
	2-7	2-7-1	sem_ext	의미확장 정보	Y	Y	Y
		2-7-2	property : locatedIn	의미확장 지역정보			
		2-7-3	property : relatedTerm	의미확장 연관정보	Y	Y	Y
		2-7-4	property : description	의미확장 상세정보	Y	Y	Y

[그림 7-3]


```
{
  "image": {
    "identifier": "HF010508_0101_0369.JPG",
    "imsize": [
      1920,
      1440
    ],
    "copyright": "미디어그룹사람과숲(주)",
    "date": "2019-09-30",
    "resolution": "72 DPI",
    "bit": "24",
    "F-Stop": "F/2.8",
    "exposure time": "1/962초",
    "ISO": "182",
    "focal length": "3mm",
    "full aperture": "2.97",
    "view angle": "15",
    "white balance": "Auto",
    "depth": "3"
  },
  "regions": [
    {
      "type": "box",
      "boxcorners": [
        573,
        472,
        1401,
        820
      ],
      "ansize": [
        1920,
        1440
      ],
      "class": "궁궐",
      "tags": [
        "truncated:0",
        "종ID:HF010508",
        "대분류:유적건조물",
        "중분류:정치국방",
        "소분류:궁궐관아",
        "Instance:근정전",
        "Instance Upper:경복궁",
        "GPS 위도:37; 53; 15.9496799999906",
        "GPS 경도:126; 48; 19.4058000000078"
      ],
      "instance_uri": "",
      "sem_ext": []
    }
  ]
}
```


❶ 이미지 정보

❷ 저작권, 촬영 정보

❸ 카메라 정보

❹ 어노테이션 정보

❺ 클래스 및 하위 정보

❻ GPS 정보

❼ 의미 확장 정보

- 입력 필요 속성 처리, Json 파일 내 속성 정보 확인 진행

○ 7.2.2 Annotation 작업 간 Risk 발생가능 항목

가. 작업자의 숙련도 부족에 따른 Error 발생

- 인공지능 학습용 데이터를 가공하는 기준은 알고리즘 개발 방향 및 학습방법론에 따라 차이가 있으며, 학습용 데이터는 해당 기준에 부합되는 기준(Labeling 방법, 메타데이터 유형/종류/단계 등...)에 따라 가공이 되어야 함.
- 가공에 대한 기준이 명확히 숙지되지 않으면 기준에 맞지 않는 가공에 따른 Error가 발생함

나. 작업 간 발생할 수 있는 Human Error

- 인공지능 학습용 데이터 가공 방법은 Auto detection 도구를 활용한 자동화와 사람중심의 수동으로 크게 2가지로 나눠짐
- 알고리즘의 학습방법론에 따른 가공데이터를 자동으로 추출하기 위해선, 해당 알고리즘에 대한 맞춤형 학습이 필요. 해당 학습에 필요한 데이터는 사람 중심의 수동으로 가공되고 있음
- 실제로 오픈소스 형태의 Auto detection 알고리즘/도구의 경우 범용으로 사용되기에는 정확도가 많이 부족함(산출물 그대로 학습용 데이터로 활용하기 불가능)
- 따라서 정확도 99% 이상의 학습데이터를 생성하는데 사람 중심의 작업이 유효하며, 이 과정에서 작업자의 컨디션, 주변상황 들에 의한 Error가 발생할 수 있음

다. 다량의 데이터에 대한 Annotation 진행에 따른 Error 발생 가능성

- 알고리즘 학습에 있어, 특정 객체를 학습시키는데 필요한 인공지능 학습용 데이터는 최소 1천개부터 수십만에 이르기까지 많은 양의 데이터가 필요함
- 다량의 데이터를 가공함에 있어 학습방법론 변경 등에 따른 기준변경 등이 발생할 수 있으며, 이를 작업에 반영하는 과정에서 Error가 발생할 수 있음

○ 7.2.3 Risk 발생가능 항목에 대한 대응방안

가. 작업자의 숙련도 부족에 따른 Error 발생

1) 작업자에 대한 업무적합도 확인 절차 진행(작업자 추가 확보 필요 시)

- Annotation은 집중력, 판단력, 지속력이 중요한 작업으로 업무에 적합한 인원을 선발하고 작업에 투입하는 것이 중요한 사항임.
- 인성면접 후 Pilot 교육/평가를 통해 학습용 데이터 가공업무에 적합한 인원선발 진행

2) 교육 및 실습

- 학습용 데이터 가공 기준에 대한 사전 교육을 진행하여 기준에 대한 지식베이스의 숙련도를 확보.
- 교육진행 후 해당 내용에 대한 가공대상 객체별 실습 및 피드백을 통하여 지식베이스의 숙련도를 실제 작업 간 판단에 대한 숙련도로 확장.

3) 정기평가 실시

- 작업자의 숙련도 확인 및 숙련도 향상을 확인하기 위한 정기평가 실시
- 단위로 평가하며, 평가 결과에 따라 인원교체 및 업무 수량 조정에 활용

나. 작업 간 발생할 수 있는 Human Error

1) 교차검수

- 학습용 데이터 가공 작업 간 작업자 2인 이상을 1개의 단위(조, 파트 등...) 해당 단위 내에서 작업자 간 교차검수를 통하여 검수 진행

2) 작업도구 기능에 따른 기본 대응

- 메타 데이터 값을 입력하는 과정에서 유효한 속성 값 유형만 입력될 수 있도록 도구 내 Error 방지 기능 탑재
- 예를 들어 A, B, C 객체에 대한 Annotation을 진행한다는 가정 하에 1번부터 10번까지의 전체 속성에 있고, A라는 객체에 해당하는 속성이 1~3번까지라면, A라는 객체를 선택하였을 시 1~3번만 입력가능 속성으로 활성화.
- 또는, 이미지에 종속 되는 속성의 경우 폴더구조 등을 활용하여 자동으로 입력되도록 도구에 기능 탑재

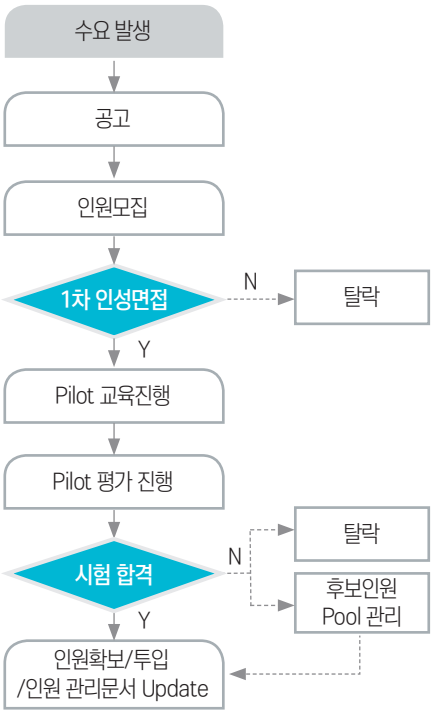
다. 다량의 데이터에 대한 Annotation 진행에 따른 Error 발생 가능성

1) 도구에 의한 검수/자동변경

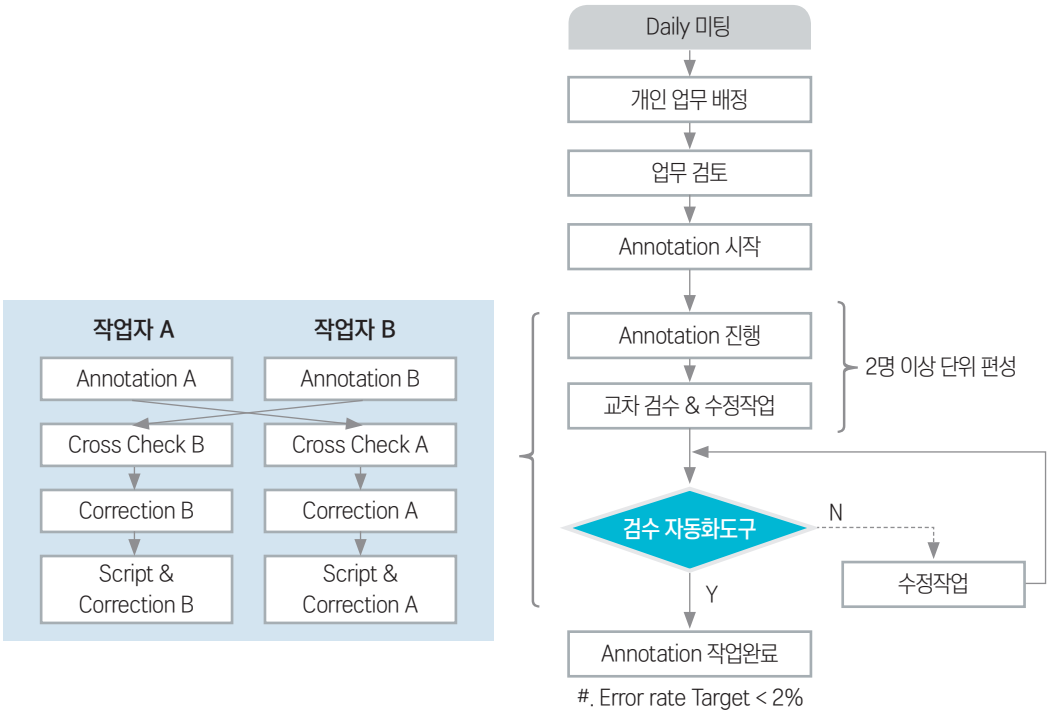
- 다량의 학습용 데이터의 경우 작업도중 기준변경 등이 발생하였을 경우 이미 작업된 학습용 데이터에 대하여 전수 검수/수정은 많은 리소스를 필요함.
- 가공 기준을 기준으로 한 검수도구를 활용하여 Error 발생 위치 및 변경된 기준에 따른 수정이 필요한 항목을 파악하고 수정/변경작업 진행
- 속성 값 변경 등의 사항은 도구 내에서 자동변경 될 수 있도록 지원(스크립트 활용도 가능함)

7.2.4 유형 별 프로세스

[그림 7-4] 인원확보



[그림 7-5] Annotation



8. 인공지능 학습용 데이터 확장

● 8.1 온톨로지 연계를 통한 이미지 의미 확장

이 전 단계에서 구축된 이미지와 어노테이션 정보에 대해 온톨로지와 연계하여 어노테이션 정보의 확장과 온톨로지를 매핑하는 과정에 대해 설명한다.

○ 8.1.1 온톨로지 매핑

기 구축된 온톨로지가 없는 경우 온톨로지 구축 일반 방법론에 따라 필요한 온톨로지를 구축하고 그 외의 경우 온톨로지 매핑을 수행한다. 1차 어노테이션된 레이블 정보를 기반으로 온톨로지를 조회할 수 있는 도구를 활용하거나 온톨로지 질의(SPARQL)를 생성하여 각 레이블에 해당하는 온톨로지 클래스, 프로퍼티, 인스턴스를 조회 한다. 조회된 결과에 대해 수동 검수를 수행하고 해당 레이블과 온톨로지 클래스 URI를 1차 어노테이션 정보에 추가한다. 이 때, 어노테이션 과정에서 부여된 레이블과 같은 의미이지만 다른 레이블로 생성된 클래스가 존재 할 수 있으므로 온톨로지 상의 상,하위 의미 관계를 수동으로 점검하여 동일한 의미의 개념이 없는지 수동으로 확인해야 한다.

[그림 8-1] 온톨로지 연계를 통한 의미 확장 과정

[그림 8-2] 온톨로지 조회를 위한 질의(SPARQL)

```
SELECT ?s ?o
WHERE {
        OPTIONAL{ ?s rdfs:label " 궁궐"@ko . }
        OPTIONAL{ ?s adp:name " 궁궐"^^xsd:string .}
        OPTIONAL{ ?s adp:description ?o . }
}
LIMIT 100
```

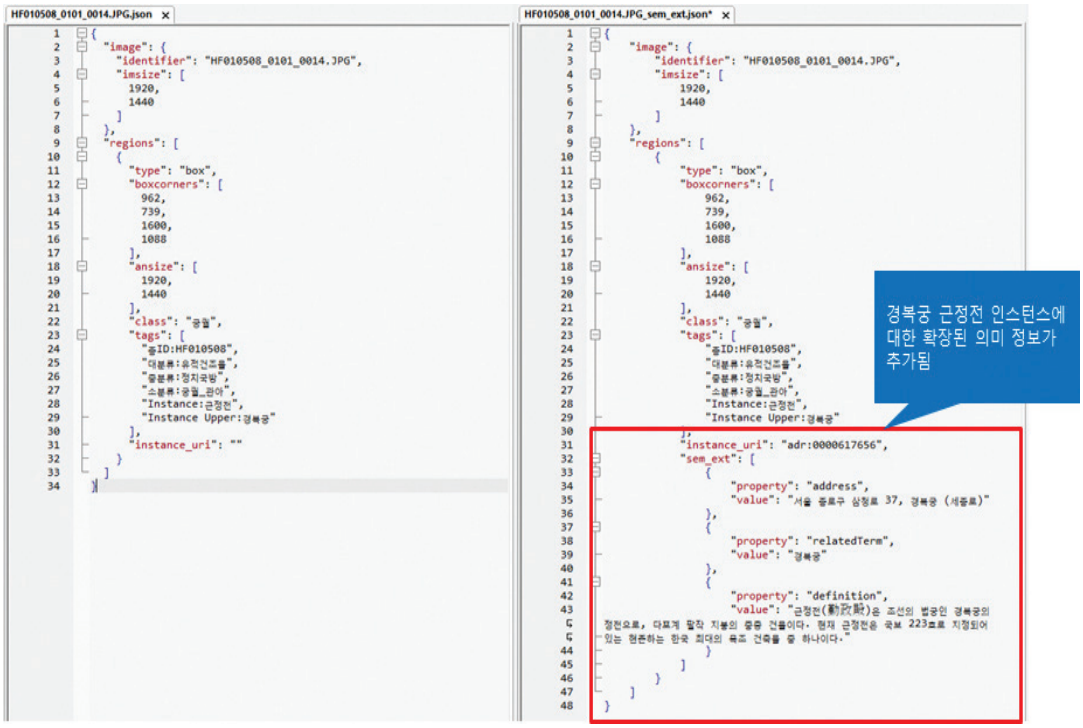
온톨로지 질의 언어를 작성할 때, 각 온톨로지에 따라 개념의 명칭을 조회하는 속성이 달라질 수 있다. 위 그림에서 보이는 바와 같이 rdfs:label을 사용하거나 name, nickname 등의 여러 가지 속성으로 표현되어 있을 수 있으므로 연계하고자 하는 온톨로지의 특성에 맞추어 SPARQL을 작성해야 한다.

○ 8.1.2 온톨로지 보완/확장

온톨로지를 보완하고 확장하는 과정정은 구축된 데이터의 활용 및 공개를 위한 측면에서 요구되는 과정이다. 전 단계에서 기 구축한 온톨로지의 표현으로 이미지 데이터의 표현이 부족한 경우에 온톨로지 구축 일반 방법론에 따라 필요한 개념을 추가하고 확장한다.

직접 구축한 온톨로지가 아닌 공개되어 있는 LOD(Linked Open Data)를 활용하는 경우에는 해당 개념을 포함하는 상위 개념을 부여한다. 예를 들어 '궁궐'이라는 개념을 활용해야하는데 해당 개념이 온톨로지 내에 없는 경우 상위 개념인 '건물'이라는 개념을 활용하거나 '건물'의 하위 개념으로 '궁궐'을 추가 구축하여 활용할 수 있다.

[그림 8-3] 온톨로지 연계를 통한 의미 확장 과정



9. 데이터 품질관리

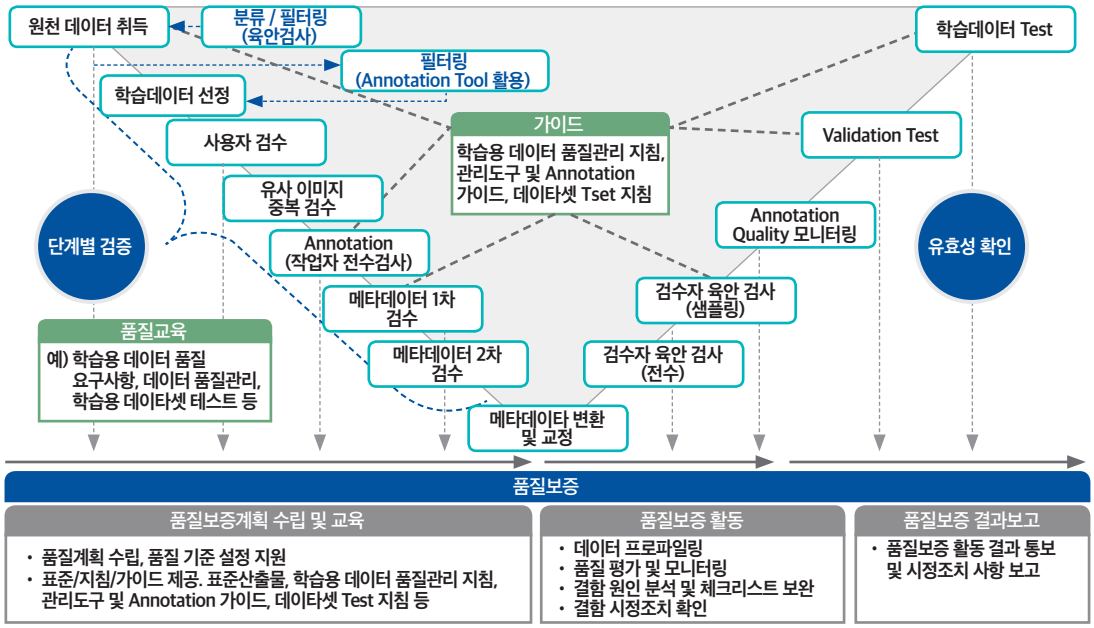
● 9.1 데이터 품질관리 방안

○ 9.1.1 품질관리 체계 개요

- 데이터셋의 품질을 보장하기 위하여 구축 단계별 검증, 데이터 셋 유효성 확인, 품질보증 활동의 3가지 활동을 수행. 품질진단 관리기법, 품질진단 도구, 분야별 업무전문가 등을 통한 품질관리 체계를 구축하고 수행

- 품질진단 관리기법, 품질진단 도구, 분야별 업무전문가 등을 통한 품질관리 체계를 구축하고 수행

[그림 9-1] 품질관리 체계 개요



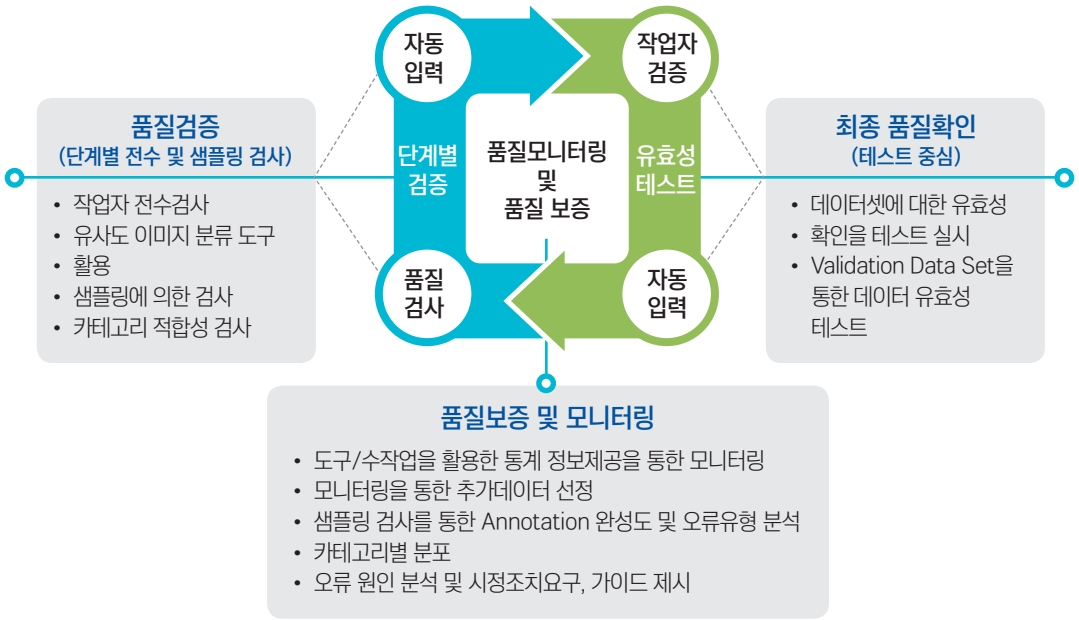
[표 9-1]

관리체계	주요 내용
품질진단 관리기법 적용	• 공공데이터 품질관리 매뉴얼 절차 준수 • 품질관리 정책수립, 계획단계 품질관리, 구축단계 품질관리, 운영단계 품질관리, 활용단계 품질관리 • 단계별 검증 및 피드백 • 데이터 프로파일링 기법 적용
품질진단 도구활용	• 저작도구 및 Annotation 자동화 도구(Annotation Quality 모니터링 도구 포함) • 메타데이터 프로파일링 도구
분야별 업무전문가	• 데이터 프로파일링을 위한 데이터 전문가 • 품질진단을 위한 품질관리 전문가

9.1.2 품질관리 활동 및 대상

- 데이터셋의 품질을 보장하기 위하여 구축 단계별 검증, 데이터 셋 유효성 확인, 품질보증 활동의 3가지 활동을 수행. 각 활동별로 수행하는 작업 및 작업 대상은 다음과 같음

[그림 9-2] 품질관리



9.1.3 학습데이터 적합성 검증

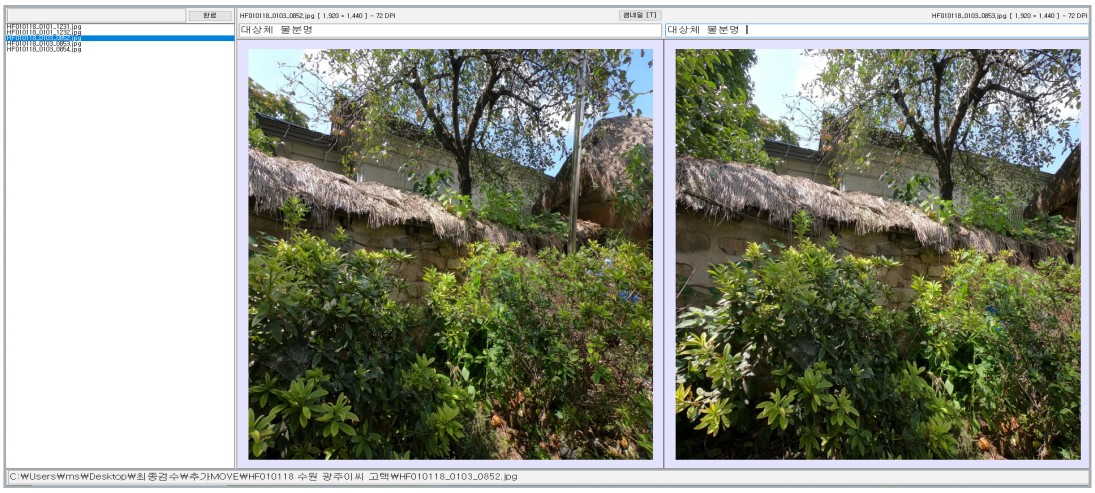
- 한국형 사물 이미지 데이터의 품질 및 활용도를 높이기 위하여 공정별 기준에 따른 검수를 진행

[표 9-2]

공정	주요 내용	비고
대상 목록 선정	• 사물이미지 구축 대상 수량 확인	360만장 이상
촬영/수집	• 촬영 및 수집된 이미지의 수량 확인 • 촬영 및 수집된 대상체 목록과 일치 확인 • 촬영 및 수집된 이미지의 해상도 확인	360만장 이상 수집 장축 800픽셀 이상 촬영 장축 1920픽셀 이상
분류/정제	• 유사 중복 이미지 확인 • 개인정보 비식별화 확인 • 이미지의 분류 정확도 확인	100% 동일이미지 삭제 카테고리 정합성 확인

- 정제 완료 데이터에 대한 적합성 육안검사 이미지 적합성 검사진행

[그림 9-3] 적합성 검증

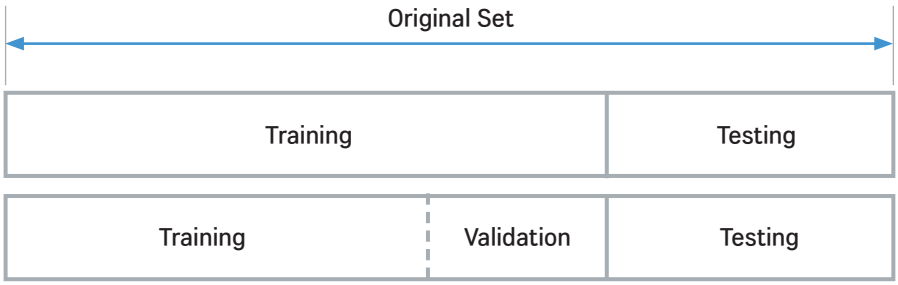


9.1.4 학습데이터 유효성 검증

1) 학습데이터 분할

- 학습데이터는 Training data set, Validation Data set, Testing Data Set으로 구분하여 준비
  - Machine learning 학습데이터의 목적은 unseen data 즉, test data에 대해 좋은 성능을 내는 것이며 모델을 만든 후 이 모델이 unseen data에 대해 얼마나 잘 동작할지에 대해서 반드시 확인이 필요함
  - 하지만 training data를 사용해 성능을 평가하면 안되기 때문에 따로 validation set을 만들어 정확도를 측정하는 것이고 두 번째는 모델을 튜닝하여 모델의 성능을 높일 수 있으며 예를 들어 overfitting 등을 막을 수 있다. training accuracy는 높은데 validation accuracy는 낮다면 데이터가 training set에 overfitting이 일어났을 가능성을 생각해볼 수 있음
  - Overfitting을 막아서 training accuracy를 희생하더라도 validation accuracy와 training accuracy를 비슷하게 맞춰줄 필요가 있기 때문에 Deep Learning을 모델을 구축한다면 regularization 과정을 한다거나 epoch을 줄이는 등의 방식으로 overfitting을 막을 수 있음

[그림 9-4] 학습데이터 분할



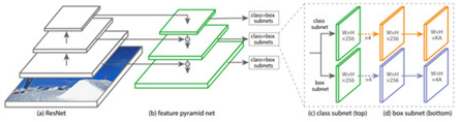
- 데이터셋 구성 목표는 다음과 같음
  - Original Dataset : 전체 학습데이터
  - Training (100%) : Neural Network model의 학습에 활용하고자 하는 실제 데이터셋
  - Validation (15%) : Original 데이터셋중에서 Neural Network model의 비연속성을 고려하여 평가를 위해 사용되는 데이터셋
  - Test (15%) : 최종적인 Neural Network model의 비연속성을 고려하여 평가를 위해 사용되는 샘플 데이터셋
  - Validation Data와 Test data의 샘플링은 고궁, 간판, 상품 등의 각 분류에서 100%를 학습데이터로 활용을 하고 15%에 해당하는 데이터를 Validation으로 활용하며 최종 15%를 TEST용 데이터로 분류하여 학습데이터에서 분리하여 제공함

- 학습데이터의 품질관리 및 유효성 검증 방안
  - 원천데이터 품질 방안
  - 주간단위의 데이터 취득 장소, 환경, 분류 등에 대한 가이드라인 제시
  - 한국형 사물이미지의 데이터셋 검수 : 각 annotation 단계 및 품질검증 단계별로 도구 활용 또는 육안 검사(샘플링 또는 전수검사)
  - 구축 공정별 검수 절차

- 학습 데이터 검수 방안
  - Process와 도구에 의한 품질 관리
  - Ground Truth와 Test 결과의 품질 평가
  - 샘플링에 의한 육안 검수 조사

2) 학습데이터 유효성 검증

- Retinanet 기반 인피닉 개발 커스터마이징 인공지능 엔진 활용
- Training Image Data에 사용하지 않은 Test Image Data를 사용
- 학습이 완료된 Dataset으로 각 Test Image Data 안에서 찾고자 하는 object를 얼마나 정확하게 찾는지 확인
- Test Image Data의 Ground Truth와 결과를 비교, 정확성을 %로 수치화

인피닉 인공지능 엔진 특징	유효성 검증 과정
• Focal Loss, 새로운 Loss Function - 학습에 필요한 오브젝트에 중점. 학습 이미지에서 학습에 불필요한 배경에 대한 가중치는 줄이고, 학습에 필요한 오브젝트의 가중치를 높임 • 빠르고 높은 성능의 Network - One-stage network, Two-stage network - 빠른 속도의 One-stage network와 높은 성능의 Two-stage network의 장점을 모두 만족 	• 메타 데이터 타입 분석 - 분석을 통해 학습을 위한 포맷으로 변경 방식 검토 • 학습 방식 결정 - 데이터와 서버 환경을 고려하여 학습 방식 결정 • 데이터 선별 - 학습용 데이터와 테스트 데이터를 선별 • 학습 진행 • 식별 진행 - 선별된 테스트 데이터로 식별 진행 • 결과 확인 - 식별된 결과를 분석하여 결과 확인

9.1.5 구축 공정별 검수절차, 방법, 목표기준율

• 한국형 사물이미지 데이터셋 구축 공정별로 검수절차와 방법, 목표기준을 다음과 같이 정의함

[표 9-3]

구축 공정	검수절차	방법	목표기준율
원천데이터 취득	• 원천데이터의 분류를 목적으로 육안으로 검수함	• 검수자 육안 검사	• 이미지 데이터의 분류 • 분류기준 일치
학습데이터 선정	• 프레임 연속성을 고려하여 카테고리 기준 수집목표 종에 대하여 선정 • 유효 학습데이터 선정	• 자동추출 도구 활용	• 각 선정프레임 간의 이동 거리 • 각 선정 프레임의 시간간격
사용자 검수	• 협회 및 전문단체를통한수집(FTP)에서 수집된 이미지에 대한 적합성을 검수	• 육안 전수검사	• 이미지 적합성 검수 99% 이상
유사 이미지 중복검수	• 유사도 분류 Tool을 활용하여 중복이미지 검수	• 자동화 전수 검사	• 중복이미지 검수 99% 이상
메타데이터 1차 검수	• 입력된 메타데이타에 대한 20% 샘플링 검수	• 샘플링에 의한 검사	• Annotation 완성도 99%이상
메타데이터 2차 검수	• 데이터베이스 카테고리에 대하여 카테고리 적합성 확인	• 전수검사	• 카테고리 적합도 100%
Training	• 4개월 단위로 제안컨소시엄에서 Test Data를 활용하여 학습데이터의 정확도 평가	• 자동화 CNN model 적용	• 구축된 학습데이터 100% 학습 진행
Validation / Test	• 4개월 단위로 제안컨소시엄에서Training dataset으로 학습하고 Validation Data를 이용하여 학습Data 정확도평가 • 최종단계 유효성 검증 실패를 대비하여 구축초기 단계부터 점진적 검증 진행	• 자동화CNN model 적용 • 초기 학습용 Data 30%로 선 학습 진행 후 10% 단위 검증 진행	• 수행결과와 GTdata를 비교하여 75%이상만족

10. 활용

● 10.1 활용 방안

○ 10.1.1 산업분야 및 연구분야 활용 예정

- 산업분야
  - 활용 ① : 국내 인공지능 사물 인식 카메라 성능 향상과 인식된 사물에 대한 연관정보 제공
  - 활용 ② : 외국인 관광객 대상 관광 안내 및 청소년 현장 교육에 AI 서비스 제공

- 연구분야
  - 활용 : 국내 중소·벤처, 스타트업 등 AI 관련 기업들의 데이터 활용성 증대

- 기타분야
  - 활용 : 확장된 의의정보 기반 이미지 검색 및 정보제공 서비스

수립된 서비스 모델 기반

INPUT	PROCESS	OUTPUT
• 원천 데이터 • 수집 방안 - 수집 방법 - 수집처 - 수집 주기 - 수집도구 등	• 데이터 정제, 변환 규칙 • 데이터 적재 및 분석, 학습 모델 • 데이터 검증 방안 • 데이터 실시간/배치 처리 환경 • 데이터 운영성과 측정 방안 • 비즈니스 효과	• 서비스 데이터 • 데이터 제공 방안 - 제공방법 - 제공처 - 제공 주기

부록 I

Annotation & 검수 자동화 작업도구

● I .1 Annotation 작업도구

○ I .1.1 도구 주요 기능

1) 인공지능 알고리즘의 학습관점에서 학습가능한 객체의 크기/비율 설정 가능
2) 수집된 이미지와 동일한 저장위치에 작업일자 별로 폴더 자동생성
3) 객체 내 다양한 메타데이터 설정 가능 : Config 파일 직접 수정과 편집기를 통한 수정 등 구조편집 제공

[표 1-1] Annotation 도구 메뉴

메뉴	내용
Image Area	선택된 이미지가 표출되는 화면
Draw Style	이미지 형식에 맞게 Style 을 선택하여 작업하며, draw style 로는 Line, Rectangle, Circle, Polyline, Polygon 으로 나뉘어 짐 .
Image state	1) Work : 작업한 경우에 O로 표시, 작업 내역이 없을 경우 빈 공간으로 표출 2) Review : Reviewer 들의 결과 OK, NG 로 표출 3) FileName : 선택된 이미지 이름 표출
Filter by	Annotator or Reviewer 가 이미지 작업 후 나온 결과 값에 대해 Work Flow, NG Image 탭을 선택하여 결과에 맞는 Image List 를 출력
Function	1) 다음과 같은 기회로 되어있으며 (<<,<, O , >, >>) 처음, 이전, 파일열기, 다음, 마지막 화면으로 이미지 이동 가능 2) Follow mouse : 체크 후 이미지 화면 이동 시 Image Area 에 확대되어 표출
Zoom Image area	Image Area 에 선택된 부분이 확대되어 Zoom image Area 에 출력
Object property	Image에 맞는 속성 값 부여
Review & Comment	Annotator 작업한 결과에 대한 OK, NG 값 부여 및 결과 값 작성

[표 1-2] Annotation 지원 Style

Draw Style		내용
	Line	직선의 모양으로 굴곡이 없는 직선 차로의 경우 사용
	Rectangle	직각의 모양으로 도로의 차량 및 표지판 그릴 때 사용
	Circle	원형의 모형으로 도로의 표지판 그릴 때 사용
	Polyline	선을 연속으로 그려 만들어지며 , 주로 커브도로의 경우 사용
	Polygon	3D 물체를 표현할 때 사용 되며 , 주로 표지판 그릴 때 사용

● Ⅰ.2 검수 자동화 도구

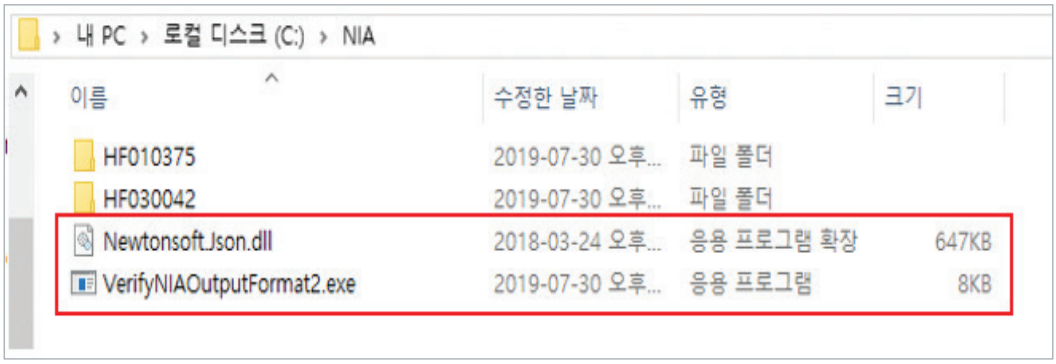
○ Ⅰ.2.1 도구 주요 기능

- .Net framework 기반 검수 스크립트
- 대상 객체가 가져야하는 메타데이터 정보 대상 적합성 여부 확인
- 검수 결과 관련 단순결과(Error 수량) 및 상세결과(Error 유형)에 대한 별도 파일 생성

○ Ⅰ.2.2 도구 사용 방법(절차 예시)

[그림 Ⅰ-1]

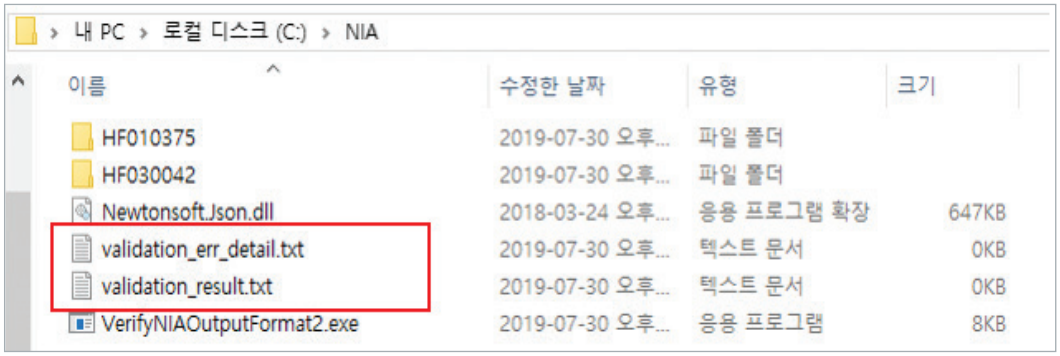
1. 검수 스크립트를 검수 대상 폴더의 상위 레벨에 위치시킴



2. 스크립트 실행

3. 실행 및 결과 화면

4. Output 파일 확인(결과 및 상세결과) : validation_result.txt / validation_err_detail.txt



5. Error 검출 예시

부록 II

인공지능 학습용 데이터 확장

● II.1 온톨로지 연계를 통한 이미지 태깅 정보의 확장

인공지능 분야 이미지 학습데이터 생성 시 이미지 또는 이미지 내 객체를 설명하기 위해 온톨로지를 활용하기 위한 시도가 증가되고 있다. 이미지 태깅 과정에 온톨로지 표현을 활용함으로써 이미지의 분류체계뿐 아니라 각 객체들 간 의미적 관계 표현이 가능하다. 또한, 온톨로지 연계 시 부여되는 고유한 데이터 아이디인 URI를 통해 데이터의 공개와 LOD 연계를 통한 확장 서비스 구현에 유리함이 있다.
본 장에서는 온톨로지 구축의 일반 방법론을 소개하고 1차 태깅된 이미지의 의미 확장을 수행하는 과정에 대한 지침을 제시한다.

● II.2 온톨로지 구축 대표 방법론

기본적인 온톨로지 구축 단계는 아래와 같은 단계를 거친다. 해당 단계는 구축하려는 온톨로지가 충분한 커버리지를 가질 때 까지 프로세스를 반복한다.

[표 II-1] 온톨로지 구축 일반 방법론

단계	설명
적용범위 결정	• 온톨로지 범위가 되는 도메인을 결정 • 온톨로지 사용 목적 이해
기존자원 재활용 검토	• 온톨로지 라이브러리 검토 • 상위 온톨로지 유무 검토 • 범용 온톨로지 활용 검토 • 도메인 온톨로지 유무 검토
용어 열거	• 클래스 혹은 속성(property)의 후보가 될 용어들을 수집해 열거
개념(Class) 정의	• 클래스 정의 및 클래스간 관계 정의 • 텍사노미 관계 정의에 해당
속성 정의	• 각 클래스별 속성 정의 • 하위클래스는 상위 클래스의 모든 속성 상속
제약조건 정의	• 속성의 제약조건을 설명하거나, 하나의 slot 에 가능한 값(value)에 대해 제한 • 요소의 개수(cardinality), 값의 타입, 값의 범위(min/max), 디폴트 값을 정의
인스턴스 생성	• 실제 인스턴스(개체) 생성 • 각 인스턴스별 속성 값 입력

03
한국어 글자체 이미지 AI 데이터 구축 가이드라인

□ Superb AI

	성명	소속	직위
가이드라인 제안	이현동	Superb AI	이사
가이드라인 작성	이현동	Superb AI	이사
	이종혁	Superb AI	이사
	이정권	Superb AI	이사
전문가 검토 위원	구영현	세종대학교	교수
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 키니앤파트너스/ 참여: 슈퍼브에이아이

서 문

1. 가이드라인의 목적

이 가이드라인의 목적은 한글 관련 인공지능 기술개발을 위해, 한글을 Labeling하기 위한 절차와 세부지침에 제공하기 위해 작성되었다. 이를 통해 구축Task 진행하는 과정에서 발생되는 비용(커뮤니케이션 코스트 등)을 축소하고 향후 연구 개발자간에 실제 기술 개발에 필요한 가이드라인을 제공하기 위해 작성되었다.

한글 글자체 이미지 데이터는 자연어 처리 분야에서 가장 기초적인 언어 자원으로 중요성이 높아지고 있다. 반면, 우리나라에는 인공지능 기술 벤처들의 수도 많지 않으며, 관련 기술 개발을 위한 데이터 기반들도 부족한 상황이다.

2. 주요 내용 요약

한국정보화진흥원의 한국어 글자체 이미지 AI데이터의 경우, 손글씨 및 인쇄체 총 500만자 및 Text in the wild 10만장의 이미지로 구축되었다. 본 구축가이드에서는 해당 데이터 구축에 활용된 기준과 수집, 가공 및 검수절차까지의 내용을 요약한다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

한국어 글자체 이미지 AI데이터 구축

Guideline for Korean Character Dataset

03

1. 적용 범위

4차 산업혁명의 핵심 기반인 산업별 실제 데이터, 인공지능(AI) 학습데이터를 전방위적으로 구축하고, 공공데이터의 원칙적 개방 등 공공·민간 데이터의 획기적 개방과 데이터 거래 활성화가 실행되는 상황임.

그럼에도 불구하고 한글에 대한 대규모 학습용 데이터는 부재한 상황으로, 이에 당사는 한국 글자체 인식 처리 기술의 표준화 및 교환·공유할 수 있는 체계를 마련하고 한글 인식기술 개발의 토대를 구축하기 위해 한글형 글씨체 DB 확보하고자 함

본 표준은 한국어 글자체 이미지 AI데이터 구축을 위한 절차와 Labeling 지침을 제시하며, 지식베이스 구축 및 효용성 검증으로 제한함. 효과적인 한글 기반 인공지능 기술개발을 위해 활용 가능한 이미지 및 Labeling 정보 등을 지식베이스로 구축·공개하고자 함

※ 본 과제에서 활용되는 국어 표준은 국립국어원이 지정한 기준을 차용해 사용됨

2. 인용 표준

- 해당 사항 없음

3. 용어 정의

● 3.1 구축 한글 글자체 정의

— 음절(syllable) : 자음과 모음으로 합성된 하나의 말소리 단위. 음절은 몇 개의 음소로 이루어지며, 모음은 단독으로 음절이 될 수 있음

— 단어(word) : 분리하여 자립적으로 쓸 수 있는 말 혹은 이에 준하는 말

— 어절(syntactic word) : 문장을 구성하고 있는 각각의 마디로 문장 성분의 최소 단위로서 띄어쓰기 단위

— 문장(sentence) : 생각 및 감정을 말로 표현할 때 완결되는 내용을 나타내는 최소 단위

● 3.2 Labeling 용어 정의

– Annotation : 해당 데이터 내의 이미지 및 텍스트, 숫자 등의 정보 값을 입력하는 과정을 뜻하며, 해당 정보 값은 프로그래밍 언어로 구성되어 있음
– bounding box : 좌상단, 우하단의 좌표로 구성되어 있는 4point polygon 직사각형
– Rectangle : 4point polygon으로 Center, Width, height, angle의 자유도 5를 가지는 직사각형

● 3.2 기술 용어 정의

– 데이터 증강 (augmentation) : Overfitting의 문제를 해결하기 위해 필요한 학습용 데이터의 양을 늘리는 방법으로 구축된 데이터에 대한 특징을 인위적으로 변경하여 학습용 데이터의 양을 증대하는 기술

4. 디지털 파일 명명 규칙

● 4.1 파일구조

① 구축데이터의 원본 이미지는 아래의 폴더로 구성

가) db1_hw : 손글씨체
나) db1_pr : 인쇄체
다) db2 : Text-in-the-wild

② 각각의 폴더는 아래의 형태로 저장

가) Image의 경우, 다운로드가 용이하도록 10gb 단위로 분할 압축 예정
나) 압축형식 : zip, tar

```
- dataset_info.json
- images
        - *.png or  *.jpg
```

③ 이미지 파일에 대한 정보와 각 이미지 파일에 대한 annotation이 담겨있는 JSON 형식의 파일 구조 : 내부 구조의 경우, 각 데이터 별 “5. Labeling 정보 구조” 부분 참고

가) Image 파일 형식 : PNG (혹은 JPG)의 이미지 파일로 구성되며 해상도는 일정하지 않음

5. 손글씨 데이터 구축

● 5.1 구축 체계

구축단계	세부절차	필수여부	세부설명
수집	수집 대상 선정	필수	수집이 필요한 글자, 단어, 문장 정의 - 글자 : 현대식 한글 11,172자 - 음절(단어) : 국립국어원 한국어 학습용 어휘 6,000 낱말 활용 (약 한글 14,000자) - 문장 : AI Hub내, 기계독해 데이터 활용 (뉴스기사로 구성, 약 한글 30,000자)
	수집 Material 제작	필수	글자/단어/문장으로 구성된 각각의 Work sheet (작업 Material) 구성 후 출력
	데이터수집자 채용	필수	17세 이상~60세 이하, 남여 성비 1:1로 구성
	데이터 수집	필수	Work sheet 작성
정제	데이터 변환	필수	Work sheet의 Text를 Digital file형태로 변환
	데이터 정제	필수	Image file 형태 데이터 유효성 자체 검증
가공	추출 (Cropping)	필수	전체 Image 파일에서 한글 영역 추출
	가공 (Labeling 입력)	필수	추출된 개별 Image 파일의 한글 정보 입력
검수	전수 검사	필수	Image 와 입력된 Label 값의 일치 여부 확인
	품질 인증	필수	외부 자문단을 통한 품질 인증

① 수집

가) 수집 대상 선정

현대 한글의 모든 글자인 11,172자를 기본으로 데이터를 수집,
이후, 실생활 활용도가 높은 단어(혹은 어절), 문장 등을 선별하여 한글 데이터 수집 진행

1) 글자(음절) : 국립국어원에서 공개한 "현대 한글 11,172 글자" 활용
2) 단어(어절) : 국립국어원에서 공개한 "한국어 학습용 어휘 6,000 낱말" 활용
3) 문장 : 말뭉치 자료 활용 예정 (AI Hub내 AI데이터-기계독해 자료 활용)
4) 수집 양 : 전체 구축량 500만자 중, 손글씨 데이터 250만자 확보*

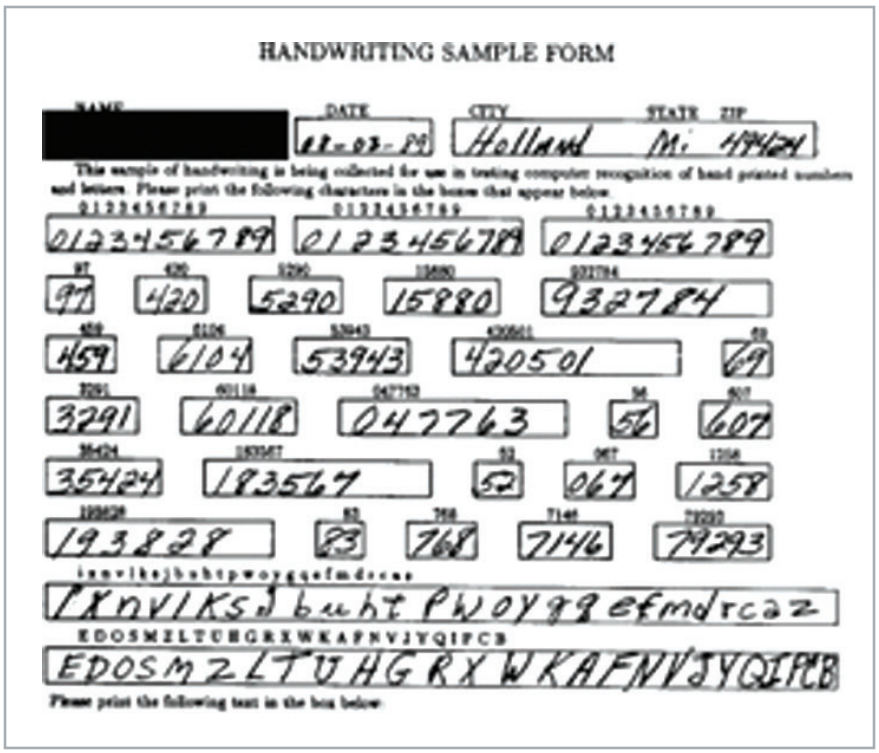
* 한국어 인식기술 관련 개발사 Interview 결과, 손글씨 데이터 양 확보 필요성에 대한 의견이 많았음 이에, 손글씨 데이터를 250만자로 확대하여 수집

나) 수집 Material 제작(Worksheet) : 손글씨 데이터 확보를 위해 글자, 단어, 문장이 포함된 Worksheet 제작

1) 지정된 글자(음절), 단어(어절), 문장으로 구성된 Worksheet*를 데이터 입력원(수집자)가 직접 손글씨로 작성하여 손글씨체 데이터를 구축

* 해당 Worksheet를 기반으로 인쇄체 구축 글자, 단어, 문장도 제작

[그림 5-1] 미 NIST에서 시행한 손글씨/단어별 데이터구축 Worksheet 예시



2) Worksheet 구성

구분	구성 글자 양	비고
글자 (음절)	11,172자	
단어 (어절)	최대 14,000자	필요 시, 추가 단어(어절 편성)
문장	약 30,000자	필요 시, 추가 문장 편성

다) 데이터 수집자 채용

1) 선정기준 : 다양한 손글씨를 확보하기 위해 성비, 나이대에 대한 기준을 수립하여 데이터 입력원 선발

2) 입력원 구성

· 다양한 성별/연령대별 고려하여 Worksheet 작성자 선발

· 약 200여명의 표본 집단을 선정하여 채용

라) 데이터 수집 : 데이터 수집자에게 수집 Material(Work Sheet) 배포 후, 작성 요청

1) 아래 형태로, 안내 문구를 아래칸에 그대로 작성하는 형태로 진행

2) 데이터 수집에 활용되는 필기구는 실생활에서 가장 많이 사용되는 볼펜으로 제한

[그림 5-2] 예시1 : 제시된 글자(음절)을 하단의 박스 안에 기입하도록 안내

가	각	갂	갃	간	갆	갇
갈	갉	갊	갋	갌	갍	갎
갏	감	감	갑	값	갓	갔
강	갖	갗	갘	같	갚	갛
개	객	갞	갟	갠	갡	갢
갣	갤	갥	갦	갧	갨	갩
갪	갫	갬	갭	갮	갯	갰
갱	갲	갳	갴	갵	갶	갷
갸	갹	갺	갻	갼	갽	갾

[그림 5-3] 예시1 : 제시된 단어(어절)을 하단의 박스 안에 기입하도록 안내

참다	크기	고기	남기다	서양
주요	냄새	여기다	공연	남녀
내놓다	떼다	속다	준비	구월
맑다	소년	소식	유월	작용
허리	놓치다	댜기	독립	또다시
머릿속	쇠고기	위반	카드	평생
간부	관념	굉장히	단어	덮다
몰다	배우	비추다	신발	앞서다
자격	통제	계단	김치	낯설다

3) 수집 Guide (예시)

Work Sheet 작성 Guide

1. 활동 소개

본 활동는 한국정보화진흥원주관사업인, 인공지능 학습용 데이터 구축사업의 한국어 글자체 이미지 데이터 구축을 위한 데이터 수집 과제의 일환으로 수행되는 데이터 수집 활동입니다. 향후 연구용 데이터로 공개 예정입니다.

2. 작성 Guide

- 본 Work sheet는 글자 / 단어(어절) / 문장으로 구성되어 있습니다.
- 구성은 아래와 같습니다.
  1) 글자 : 11,172자
  2) 단어 : 약 14,000자 : 평소처럼 연속적으로 작성해도 됨
  3) 문장 : 약 30,000자 : 평소처럼 연속적으로 작성해도 됨
    - 각 페이지에 작성된 글을 아래 지정된 칸에 기입해주시면 됩니다.
    - 오탈자가 발생할 경우, 화이트를 통해 수정 후 작업을 해주시기 바랍니다.

3. 제출 기한

- 본 과제 제출은 과제 발행일로부터 2주입니다.

4. 저작권 판매/양도 규정

- 작업 후, 본 데이터는 Superb AI에 양도/귀속됩니다.

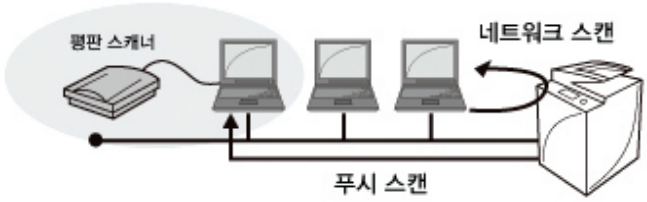
② 정제

가) 데이터 변환 : 수집된 Work Sheet는 수집하여 대규모 네트워크 스캔 작업을 통해 디지털 데이터로 변환

1) 수집 Work Sheet 디지털화 시행
· 파일 형식 : jpg, jpeg, png 등

2) 디지털화된 이미지 데이터 분류 : 글자(음절), 단어(어절), 문장

[그림 5-4] 대규모 네트워크 스캔 시행 방법



나) 데이터 정제 : 고품질 Raw Data 확보를 위해 디지털화된 Work Sheet의 품질검사

– 검증 작업자를 통한 품질 검사 시행 : 디지털 파일 내, 글자 이미지가 계획대로 확보되었는지 육안으로 판단

1) 스캔품질 확인 : 흔들림(육안으로 식별 불가한 흔들림) 및 기타 노이즈(오탈자, 글자 외 오염물질, 구겨짐 등) 확인
2) 오탈자 검수 : 제시된 음절/어절/문장과 수집자의 작성분 일치 여부 확인

③ 가공

가공자 선발

1) 그래픽/디자인 전공 혹은 관련 경력 2년 이상의 후보자 모집
2) 입사 Test (QA이해도) 및 2주간 업무 Test(작업정확도) 시행
3) 작업정확도 기반 우수 작업자 선발 후 가공 작업 수행

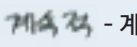
가) 추출 : 1차 가공 output인 글자별 데이터 확보를 위한 이미지 cropping

구분	수집 양	형태
글자 (음절)	11,172자	글자 단위로 11,172개 이미지 추출
단어 (어절)	최대 14,000자(약 6,000 단어)	단어 단위로 약 6,000개 이미지 추출
문장	최대 30,000자 (약 1,000 문장)	문장 단위로 약 1,000개 이미지 추출

나) 가공 : 각 이미지에 표시된 글자를 json 파일 형태로 입력 Annotation 정보 기입

1) Cropping된 이미지는 Annotation 도구에 1개씩 가공됨
2) 작업자가 이미지의 글자를 입력할 경우, 사전에 입력된 형식에 맞춰 Annotation 정보가 생성됨

[그림 5-5] 참고 : 가공 이미지

 - 세 / - 련 / - 된	 - 계속적
글자 단위의 가공	단어 단위의 가공

④ 검증

가) 전수검사 Process : 구축해야 할 Rawdata 및 Annotation 정보와 정확하게 일치하는지 확인

1) 시스템 셋팅 : 서버에 구축된 데이터 “Rawdata, Json파일” 업로드
2) 파일 불러오기 : 데이터 검수인원이 Paring 된 Rawdata, Json 파일 로드
3) 검수하기 : 데이터 검수인원은 아래 2가지 확인

· Rawdata가 유효한가 : 흔들리거나 깨지지 않았는가?

· Json 파일 정보가 정확한가 : 위의 지시문과 Annotation 정보가 일치한가

4) 저장하기 : 검수가 완료되어, 정확하게 작업된 데이터는 저장, 외 데이터는 폐기

5) A/S진행 : 구축목표량에 미달할 경우, 추가 작업 계획 수립 후 실행

나) 품질 인증* : 외부 자문단을 통한 데이터 품질 검증

* 첨부자료#3 품질인증 참조

다) 법률 이슈 검증

1) 지적재산권 법률 검토

· 손글씨의 경우, 데이터 입력원(수집자)을 통해 데이터 활용 및 공개에 대한 동의를 득하여 구입

· 저작권협회를 통해 확인한 결과, "개인 동의" 및 "구입"을 통해 확보한 데이터는 지적재산권 이슈가 없는 것으로 확인

● 5.2 Labeling 정보 구조

① Labeling Structure

```
JSON 형식
{
   "info": info,
   "images": [image],
   "annotations": [annotation],
   "licenses": [license]
}

info{
   "name": str, // 데이터 셋 이름
   "description": str, // 데이터 셋 설명 (optional)
   "url": str, // 데이터 셋 홈페이지 (optional)
   "date_created": datetime, // JSON 파일이 만들어진 날짜/시간
}

images{
   "id": str, // 이미지 아이디
   "width": int, // 이미지 width
   "height": int, // 이미지 height
   "file_name": str, // 이미지 파일 이름
   "license": str, // 라이센스 이름 (optional)
   "date_captured": datetime // 사진 찍힌  날짜/시간
(optional)
}

annotations{
   "id": str, // annotation 아이디
   "image_id": str, // 이미지 아이디
   "text": str, // 이미지에서 보이는 문장  내용
   "attributes": object // annotation 특이사항 (optional, 손글
씨의 경우 작성자의 성별, 연령대 등의 정보가 들어가며 인쇄체 데이
터의 경우 font의  종류와 크기 등 생성에 사용된 변수들이 기록됨)
}

licenses{
   "name": str, // 라이센스 이름 (license안에서 unique 해야 함)
```

```
// JSON 예시
{
   "info": {
     "name": "한글  OCR용 손글씨 데이터 셋",
     "date_created": "2000-01-01T00:00:00+00:00"
  },
   "images": [
   {
      "id": "12345678-1234-5678-1234-567812345678",
      "width": 1920,
      "height": 1080,
      "file_name": "test.jpg",
      "license": "Attribution License",
      "date_captured": "2000-01-01T00:00:00+00:00"
   }
  ],
   "annotations": [
   {
      "id": "87654321-8765-4321-8765-432187654321",
      "image_id":  "12345678-1234-5678-1234-
567812345678",
      "text": "안녕"
   }
  ],
   "licenses": [
   {
      "name": "Attribution License",
      "url":  "http://creativecommons.org/licenses/
by/2.0/"
   }
  ]
}
```

Label Structure

프로젝트	한국어 글자체 이미지 AI데이터 구축	단계	설계
문서번호	EDK_AM_A_13 버전 1.2	작성일자	6/28/2019

인터페이스ID		Structure Name	구축DB1 Label Structure
Label 구조	JSON	개발유형	

NO	항목명		길이	타입	코드 여부	필수 여부	Name/Value	비고
	한글명	영문명						
	데이터셋정보	info						JSON Object
1	데이터셋명	name	100	String		Y		
2	데이터셋상세설명	description	1000	String				
3	데이터셋URL	url	200	String				
4	데이터셋생성일자	date_created	100	String		Y		
	이미지정보	images		List				List of JSON Object
5-1	이미지식별자	id	100	String		Y		
5-2	이미지너비	width	4	Number		Y		
5-3	이미지높이	height	4	Number		Y		
5-4	이미지파일명	file_name	100	String		Y		
5-5	이미지라이선스	license	100	String				
5-6	이미지촬영일자	date_captured	100	String				
	어노테이션정보	annotations		List				List of JSON Object
6-1	어노테이션식별자	id	100	String		Y		
6-2	연관이미지식별자	image_id	100	String		Y		
6-3	어노테이션텍스트	text	1000	String		Y		
	어노테이션속성	attributes		Object				JSON Object
6-4-1	종류	type	50	String				공통 : 글자(음절), 단어(어절), 문장
6-4-2	나이	age	3	String				손글씨 : 정보 수집자의 나이
6-4-3	성별	gender	1	String				손글씨 : 남, 여
6-4-4	직업	job	50	String				손글씨 : 정보 수집자의 직업
6-4-5	글꼴	font	50	String				인쇄체 : 수집된 글꼴
	라이선스	licenses		List				List of JSON Object
7-1	라이선스명	name	100	String		Y		
7-2	라이선스URL	url	200	String		Y		
참고사항								

② 고려사항

가) 모든 이미지의 한글 글자 정보를 하나의 파일(JSON 형식)에 저장하는 구조

나) “Text’에 들어가는 내용은 최소 한 글자에서 최대 한 줄의 문장

6. 인쇄체 데이터 구축

● 6.1 구축 체계

구축단계	세부절차	필수여부	세부설명
수집	수집 대상 선정	필수	수집이 필요한 글자, 단어, 문장 정의 - 글자 : 현대식 한글 11,172자 - 음절(단어) : 국립국어원 한국어 학습용 어휘 6,000 낱말 활용 (약 한글 14,000자) - 문장 : AI Hub내, 기계독해 데이터 활용 (뉴스기사로 구성, 약 한글 30,000자)
	수집 Material 정의	필수	한글/단어/문장으로 구성된 약 5.5만장의 Work sheet 를 하나의 구축 set로 지정 후, 다른 폰트로 구축가공
	Material 다양성 확보	선택	글자 폰트, 배경색, 노이즈 등 case 적용
	데이터 수집	필수	화면 캡쳐
정제	데이터 정제	필수	Image file 형태 데이터 유효성 자체 검증
가공	추출 (Cropping)	필수	전체 Image 파일에서 한글 영역 추출
	가공 (Labeling 입력)	필수	추출된 개별 Image 파일의 한글 정보 입력
검수	전수 검사	필수	Image 와 입력된 Label 값의 일치 여부 확인
	품질 인증	필수	외부 자문단을 통한 품질 인증

① 수집

가) 수집 대상 선정 : "손글씨 데이터 구축"의 1.1. 가 항과 동일

나) 수집 Material 정의 : "손글씨 데이터 구축" 1.1. 나 항과 동일 (글자수 250만자)

다) Material 다양성 확보

1) 다양한 폰트 확보*: 시장 수요 기반으로 활용 빈도수가 높은 폰트 선정

* 첨부자료#2 구축 폰트 참조

라) 데이터 수집 : 인쇄물 생성 프로그램을 통한 인쇄체 Material 생성

1) 인쇄물 자동 생성 프로그램 개발 : 손글씨 데이터 구축과 동일한 데이터를 다양한 변수를 반영하여 출력할 수 있는 시스템 개발

※ 글씨체마다 적용되어 있는 고유한 자간 값을 활용

2) 선정된 인쇄체 폰트를 활용하여 자동으로 글자, 단어, 문자 생성 진행

② 정제 : “손글씨 데이터 구축” 1.2 항과 동일

③ 가공 : “손글씨 데이터 구축” 1.3 항과 동일

④ 검수 : “손글씨 데이터 구축” 1.4 항과 동일

※ 데이터 증강 : 기본 구축량 외 기타 변수를 적용하여 증강 데이터 구축 (선택)

- 데이터 증강 변수 (랜덤 적용하여 3배수 구축)

	노이즈	선명도	왜곡	배경 변경	Flip	크기 변경	회전	글자 색상변경
내용	배경 얼룩	blur 10%	구겨짐	회색, 노랑색	상하, 좌우	12pt, 16pt	정각도, 좌우 20	검정, 빨강, 파랑, 흰색 등

- 변형 가이드 : AI모형 활용을 위한 변형 가이드 안내
  - 노이즈 적용 : Gaussian Noise 및 Salt & Pepper 노이즈 적용
  - 선명도 변경 : 3x3 Blur kernel 및 Gaussian blur kernel을 이용하여 blur 효과 적용
  - 왜곡 : Radial distortion 및 Homography를 이용하여 이미지 왜곡
  - 배경 변경 : 배경 영역을 추출하여 임의의 이미지로 대치 Flip: 상하 및 좌우
  - Flip을 임의로 적용 크기변경 : Area resize method를 적용하여 2배 이내에서 크기 변경
  - 회전 : 텍스트의 맨 왼쪽 지점을 기준으로 90도 이내에서 회전 적용
  - 색상 변경 : 전체 이미지의 색상을 임의의 색상으로 변경 (이미지 내 동일한 color transform 적용)

6.2 Labeling 정보 구조 : 손글씨 데이터 구축 : 2항과 동일

7. Text-in-the-wild 데이터 구축

● 7.1 구축 체계

구축단계	세부절차	필수여부	세부설명
수집	수집 대상 선정	필수	수집 데이터 정의 - 도로교통표지판 / 상품 / 간판 / 도서
	수집 Guide-line 작성	필수	유효한 데이터 수집을 위한 가이드 제시
	수집 담당자/기관 선정	필수	수집전문 인력 채용 및 크라우드 소싱 업체
	데이터 수집	필수	이미지 데이터 촬영/수집
정제	데이터 정제	필수	촬영된 데이터의 유효성 자체 검증
가공	1차 가공 (영역 검출)	필수	이미지 데이터 내, 한글 글자 영역 표시
	2차 가공 (한글 입력)	필수	표시된 영역 내, 한글정보 입력
검수	전수 검사	필수	Image 와 입력된 Label 값의 일치 여부 확인
	품질 인증	필수	외부 자문단을 통한 품질 인증

① 수집

가) 수집 대상 : 한글인식기술 사용빈도가 높은 일상생활의 한글 데이터 확보를 원칙으로 함 (관련 기술 개발사 의견 수렴)

1) 도로교통표지판 : 자율주행 관련 산업 적용
2) 상품 : 이커머스 관련 산업 적용
3) 간판 : 자율주행 관련 산업 적용
4) 도서 : 출판 및 이커머스 관련 산업 적용
5) 수집 양 : 총 10만장

구분	수집 양	비고
도로교통표지판	동일한 비율을 목표로 구축 * 단, 수집 상황에 맞춰 수집양이 변동 될 수 있음	도로교통표지판 및 이정표 등
상품		상표, 특정 브랜드 상품 등
간판		상시 외부 노출되어 있는 자영업자 및 기업 간판
도서		책 외부 표지

나) 수집 Guide-line 작성 : 아래 내용을 수행할 데이터 수집자 모집 (예시)

1) 과제 소개 (수집목적 전달)

일상생활에서 발견할 수 있는 간판/광고판 내에 있는 한글, 영문, 숫자 사진을 촬영하여 업로드 해주시는 공모전입니다.

※ 본 공모전은 한국정보화진흥원 주관 사업인, 인공지능 학습용 데이터 구축사업의 한국어 글자체 이미지 데이터 구축을 위한 데이터 수집 공모전입니다. 향후 연구용 데이터로 공개 예정입니다.

2) 촬영 가이드

- 한글이 사진의 10~80%를 차지해야 함
- 촬영된 이미지 내 문자는 최대 7자를 넘지 않아야 함(작은 글자는 촬영되면 안 됨)
- 가로사진 O / 세로사진 X
- 권장 해상도 : FHD 이상 (1920 X 1080)
- 파일 형식 : JPG / JPEG
- 올려보거나 내려보며 촬영한 사진이 아닌 정면 시선으로 촬영한 사진이어야 함 (가능 시선 각도 상하 15도)
- 피사체가 명확해야 함 (흔들린 사진 X / 노이즈가 심한 사진 X)

※ 목표 피사체 외의 한글, 영문, 숫자 등이 최대한 나오지 않게, 목표 피사체 앵글을 크게 촬영하여 주시기 바랍니다.

3) 저작권 판매/양도 규정

- 공모 요강에 맞는 사진은 실시간으로 판매되며, 판매 완료된 사진의 저작권(2차적 저작물 작성권 포함)은 Superb AI에 양도/귀속됩니다.

다) 수집 담당자/기관 선정 :

1) 전문 사진 촬영 기사 채용 : 전문적인 사진 촬영을 위한 프리랜서 작가 고용
2) 크라우드소싱 업체 선정 : 다양한 환경의 이미지 수집을 위한 크라우드소싱 방법 선택

라) 데이터 수집 : 지정된 수집 Guide에 맞춰 데이터 수집

② 정제 : "손글씨 데이터 구축" 1.2 항과 동일

③ 가공 : Image 당 최대 15자로 제한

가) 1차 가공 (영역 추출) : 1차 가공 output인 이미지 내 Text 영역 검출

1) Text가 있는 위치의 좌표값을 Bounding Box 형태로 좌측 상단의 좌표 (x1,y1)값과 우측 하단의 좌표 (x2,y2)의 값을 json 파일 형태로 기록

[그림 7-1] 예시 : 이미지 내 한글 글자 영역 표시

2) 2차 가공(한글정보 입력) : bounding box가 표시된 영역의 글자를 받아쓰기(Transcription) 하여 기입

[그림 7-2] 예시 : box로 표시된 영역의 Text를 입력

찾을 수 있는 항목

1. PROFESSIONAL
2. FOR
3. HAND
4. DRIP
5. COFFEE
6. DELPINO
7. GOOD
8. CAFE
9. ROASTERY
10. 핸드드립
11. 로스터리
12. 카페
13. 델피노

④ 검수

가) 전수 검사 : 구축해야 할 Rawdata 및 Annotation 정보와 일치하는지 확인

나) 품질 인증* : 외부 자문단을 통한 데이터 품질 검증

* 첨부파일 3 참조

● 7.2 Labeling 정보 구조

① Labeling Structure

```
// JSON 형식
{
   "info": info,
   "images": [image],
   "annotations": [annotation],
   "licenses": [license]
}

info{
   "name": str, // 데이터 셋 이름
   "description": str, // 데이터 셋 설명 (optional)
   "url": str, // 데이터 셋 홈페이지 (optional)
   "date_created": datetime, // JSON 파일이 만들어진 날짜/시간
}

images{
   "id": str, // 이미지 아이디
   "width": int, // 이미지 width
   "height": int, // 이미지 height
   "file_name": str, // 이미지 파일 이름
   "license": str, // 라이센스 이름 (optional)
  "date_captured":   datetime // 사진 찍힌 날짜/시간 (optional)
}

annotations{
   "id": str, // annotation 아이디
   "image_id": str, // 이미지 아이디
   "bbox": [x, y, width, height], // text의 위치 정보
   "text": str, // 이미지에서 보이는 문장   내용
   "attributes": object // annotation 특이사항
(optional, 글자가 기울어지거나 뒤집어졌는지, 난이도 등이 기입될 수 있음)
}

licenses{
   "name": str, // 라이센스 이름 (license안에서 unique 해야 함)
   "url": str // 라이센스 내용이 들어   있는 링크
}
```

```
// JSON 예시
{
   "info": {
     "name": "한글 OCR용 text-in-the-wild 데이터 셋",
     "date_created": "2000-01-01T00:00:00+00:00"
  },
   "images": [
   {
      "id": "12345678-1234-5678-1234-567812345678",
      "width": 1920,
      "height": 1080,
      "file_name": "test.jpg",
      "license": "Attribution License",
      "date_captured": "2000-01-01T00:00:00+00:00"
   }
  ],
   "annotations": [
   {
      "id": "87654321-8765-4321-8765-432187654321",
      "image_id":   "12345678-1234-5678-1234-
567812345678",
      "bbox": [5, 7, 20, 10],
      "text": "안녕"
   }
  ],
   "licenses": [
   {
      "name": "Attribution License",
      "url":   "http://creativecommons.org/licenses/by/2.0/"
   }
  ]
}
```

Label Structure

프로젝트	한국어 글자체 이미지 AI데이터 구축	단계	설계
문서번호	EDK_AM_A_13 버전 1.2	작성일자	6/28/2019

인터페이스ID		Structure Name	구축DB2 Label Structure
Label 구조	JSON	개발유형	

NO	항목명		길이	타입	코드 여부	필수 여부	Name/Value	비고
	한글명	영문명						
	데이터셋정보	info		Object				JSON Object
1	데이터셋명	name	100	String		Y		
2	데이터셋상세설명	description	1000	String				
3	데이터셋URL	url	200	String				
4	데이터셋생성일자	date_created	100	String		Y		
	이미지정보	images		List				List of JSON Object
5-1	이미지식별자	id	100	String		Y		
5-2	이미지너비	width	4	Number		Y		
5-3	이미지높이	height	4	Number		Y		
5-4	이미지파일명	file_name	100	String		Y		
5-5	이미지라이선스	license	100	String				
5-6	이미지촬영일자	date_captured	100	String				
	어노테이션정보	annotations		List				List of JSON Object
6-1	어노테이션식별자	id	100	String		Y		
6-2	연관이미지식별자	image_id	100	String		Y		
6-3	어노테이션텍스트	text	1000	String		Y		
6-4	어노테이션속성	attributes		Object				
6-5	어노테이션바운딩박스	bbox	4	List		Y		List of [x, y, width, height]
	라이선스	licenses		List				List of JSON Object
7-1	라이선스명	name	100	String		Y		
7-2	라이선스URL	url	200	String		Y		
참고사항								

② 고려사항

가) 모든 이미지의 한글 글자 정보를 하나의 파일(JSON 형식)에 저장하는 구조

나) bbox(bounding box)는 해당 이미지를 벗어나지 않음

8. 구축 데이터셋 정보

● 8.1 구축 데이터셋 정보

구분	글자수	데이터 출처	비고
손글씨	2,500,000	글자(음절) : 국립국어원 “현대 한글 11,172자” 단어(어절) : 국립국어원 “한국어 학습용 어휘 6,000낱말”, 약 14,000자 문장 : AI Hub 내 기계독해 데이터, 약 30,000자	-
인쇄체	2,500,000		증강변수를 적용후, 추가 구축
Text-in-the wild	이미지 10만장	서울 시내 간판, 도서, 상품	-

● 8.2 구축목표 수량

구분	구축 내용			운영 인력 (명)	운영 기간	비고
	분류	소분류	구축량			
최종 500만자 + 약 50만자 (이미지10만장)	손글씨 250만자	글자	60만자	50	2019.07.01 ~ 2019.11.30	주간 단위 수집량 약 5,000자 수집과 병렬로 정제, 가공 진행
		단어	80만자	50 (글자 40면 포함)	2019.07.01 ~ 2019.11.30	주간 단위 수집량 약 10,000자 수집과 병렬로 정제, 가공 진행
		문장	110만자	50 (단어 70명 포함)	2019.07.01 ~ 2019.11.30	주간 단위 수집량 약 10,000자 수집과 병렬로 정제, 가공 진행
	인쇄체 250만자	글자	60만자	3	2019.07.01 ~ 2019.11.30	프로그래밍을 통한 작업 수행 후 출력/스캔&가공 진행
		단어	80만자	3	2019.07.01 ~ 2019.11.30	
		문장	110만자	3	2019.07.01 ~ 2019.11.30	
	Text in the Wild (약 50만자)	이미지	10만장	10	2019.06.17 ~ 2019.11.30	크라우드 소싱 병행, 동일한 비율을 목표로 구축하되, 수집 상황에 맞춰 수집양이 변동 될 수 있음

8.3 Ground Truth 정의

① 손글씨 / 인쇄체 / Text-in-the-wild의 Ground Truths 정의

가) Multi Voting 후, 응답자 80%가 같은 판단을 한 경우, Ground Truths로 판정

② 시행방법

가) 제시된 Task와, 데이터 수집자가 만든 이미지 준비 (수집자가 작성한 글자)

나) 5명 이상의 인원이 해당 데이터를 보고 바르게 작성했는지 여부에 대한 true / false를 판단

다) 응답자 중 80%이상이 ture를 응답할 경우, Ground Truths로 판정

부속서 A

구축 폰트

A.1 인쇄체 폰트 선정 기준

- 시장 수요조사 : 설문 진행 (총 272명 참여)
- 대한민국 정부 블로그 : 기업 및 공공기관에서 사용하는 무료 글꼴 (링크)
- 실증기관 interview 조사

A.2 폰트 선정 목록

- Survey 결과, 2회 이상 언급된 폰트 46종 (특정 기관 내에서만 사용되는 폰트 제외) 공공기관에서 사용되는 폰트 11종
- 실증기관의견 반영 3종 선정
- 이 중, 저작권 이슈가 발생가능한 폰트 및 현대한글 11,172자를 구현하지 못하는 글자는 제외

번호	폰트명	소유사	선정근거		
			설문	공공기관	실증기관
1	바탕	-	●		
2	돋움	-	●		
3	굴림	-	●		
4	맑은고딕	-	●		
5	궁서	-	●		
6	휴먼명조	휴먼컴퓨터	●		
7	휴먼고딕	휴먼컴퓨터	●		
8	중고딕	-	●		
9	나눔고딕	네이버	●		
10	나눔바른고딕	네이버	●		
11	나눔고딕 코딩	네이버	●		
12	나눔명조	네이버	●		
13	나눔손글씨펜	네이버	●		
14	나눔손글씨붓	네이버	●		
15	노토산스	구글&어도비	●		
16	본고딕	구글&어도비	●		
17	노토세리프	구글&어도비	●		
18	대한민국정부 상징	문체부	●	●	
19	호국	국방부	●	●	
20	서울남산	서울시청	●	●	

번호	폰트명	소유사	선정근거		
			설문	공공기관	실증기관
21	서울한강	서울시청	●	●	
22	제주 한라산	제주시		●	
23	제주 명조	제주시		●	
24	제주 고딕	제주시	●	●	
25	부산	부산시	●	●	
26	고양	고양시	●	●	
27	전북	전북시		●	
28	푸른전남	전라남도		●	
29	아리따부리	아모레퍼시픽	●		
30	SKT 뫼비우스	SKT	●		
31	빙그레따옴	빙그레	●		
32	티몬소리	티몬	●		
33	미생	다음	●		
34	스포카한산스	스포카	●		
35	함초롱돋움	한글과컴퓨터	●		
36	함초롱바탕	한글과컴퓨터	●		
37	농협희망	농협	●		●
38	하나	하나금융그룹	●		●
39	조선명조	조선일보	●		●
40	한겨레	한겨레	●		
41	대한	윤디자인	●		
42	가는안상수체	-	●		
43	수화명조	정수화폰트연구소	●		
44	태룡고딕	정수화폰트연구소	●		
45	EBS훈민정음새론	EBS	●		
46	EBS훈민정음	EBS	●		
47	KBIZ 한마음고딕	중소기업중앙회	●		
48	KBIZ 한마음명조	중소기업중앙회	●		
49	만화진흥원	(재)한국만화영상진흥원	●		
50	이롭게바탕	이롭게	●		

※ 폰트의 "비영리 사용여부" 기준 지속 확인 및 업데이트 예정. 저작권 이슈가 발생할 경우 해당 폰트는 삭제 혹은 교체 예정

A.3 폰트 선정 근거

- 폰트관련 전문기관 Survey를 통해 선정
- 응답자 분포

구분	공공기관	기업	교육기관	기타
분포	16.54%	58.82%	8.82%	15.81%

- 설문 결과 요약

구분	맑은고딕	나눔고딕	돋움체
분포	37.04%	9.63%	3.70%

A.4 수요조사 및 실증기관 Survey를 통해 추가 선정

- 폰트다운로드 기관 참고 (네이버/다음)
- 한국전자문서협회 114개 회원사를 대상으로 2019년 06월 10일부터 2019년 06월 14일까지 온라인 설문으로 총 7개 항목의 설문으로 진행
  (실증기관 또는 실증기관의 고객사에서 주로 사용하는 한국어 글자체는 맑은체가 52%, 나눔체 16%, 굴림체 12%, 중고딕 4%, 돋움체 4%, 기타로는 기본 윈도우 폰트 외에 기관 고유의 폰트 사용이 4%로 응답)

A.5 고려 사항

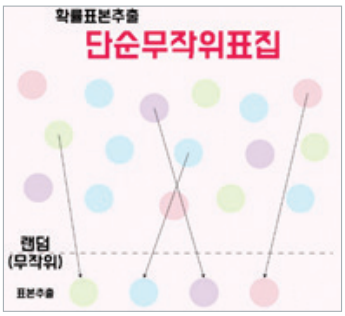
- 폰트 관련 지적재산권 해결 방안 고려 : 비영리 목적의 배포 가능여부 확인
- 다양한 글자 자간에 대한 데이터 분류 제시 : 폰트 고유의 자간을 활용할 예정
- Material다양성 확보를 위해 다양한 폰트 확보, 구겨짐, 오염환경 이외에 글자 크기, 회전, Random Crop, 커브, Low Contrast 등의 효과 추가
  (https://github.com/Sanster/text_renderer 참조)
- 저작권 관련 확인 결과, 민간/공공기관에서 활용되는 폰트의 경우, “비영리” 목적에 한해 대외공개 관련 지적재산권 이슈가 없는 것으로 파악됨
  (현재 추가 확인 중이며, 문제가 예상될 경우 유료 구매 및 폰트 대체 작업 진행)

부속서 B

품질검사 (외부자문단을 통한 품질 검사)

● B.1 무작위 표본 추출 (simple Random Sampling)

본 검증 방법은 통계학에서 사용하는, 모집단(population)의 각각의 요소 또는 사례들이 표본(sample)으로 선택될 가능성이 같게 되는 표본 추출법임. 유한모집단에서 n개의 추출단위로 구성된 모든 부분 집합이 표본으로 선택될 확률이 같도록 설계된 표본추출방법을 뜻함



● B.2 외부 자문단을 통한 데이터 품질 검사 계획

가) 슈퍼브에이아이 구축 데이터셋에서 sample 5,000개를 추출 (전체 구축량의 0.1%수준)

나) 복수의 자문단(추후 확정)이 해당 데이터셋의 정확도를 검증한 후, 의견을 제출

다) Multivoting을 통해 제출된 의견이 “정확하게 구축되었음” 로 일치할 경우, 목표 수준에 충족된 데이터셋으로 판단

● B.3 검사 기준은 아래와 같음 (평가기준 : 일치율 95%이상)

가) 손글씨/인쇄체 :

1) 제시된 음절/어절/문장의 구축 수량이 일치하는가
2) 각각의 이미지는 흔들림 및 노이즈에 대한 문제가 없는가
3) 각각의 한글은 오탈자 없이 기입되었는가

나) Text-in-the-wild

1) 제시된 구축 수량과 일치하는가
2) 각각의 이미지는 흔들림 및 노이즈에 대한 문제가 없는가
3) 각각의 한글은 오탈자 없이 기입되었는가
4)이미지 내 한글이 정확하게 boundingbox로 표시되었는가?
5) 표시된 box 안의 한글이 정확하게 기입되었는가?

다) 추후 추가 문항 개발 필요성 확인 및 추가

1) 자문단 의견 반영 및 타 데이터셋 구축 Guide 참고

부속서 C

Label Structure

● C.1 손글씨 및 인쇄체 데이터 Structure

Label Structure

프로젝트	한국어 글자체 이미지 AI데이터 구축	단계	설계
문서번호	EDK_AM_A_13 버전 1.2	작성일자	6/28/2019

인터페이스ID		Structure Name	구축DB1 Label Structure
Label 구조	JSON	개발유형	

NO	항목명		길이	타입	코드 여부	필수 여부	Name/Value	비고
	한글명	영문명						
	데이터셋정보	info						JSON Object
1	데이터셋명	name	100	String		Y		
2	데이터셋상세설명	description	1000	String				
3	데이터셋URL	url	200	String				
4	데이터셋생성일자	date_created	100	String		Y		
	이미지정보	images		List				List of JSON Object
5-1	이미지식별자	id	100	String		Y		
5-2	이미지너비	width	4	Number		Y		
5-3	이미지높이	height	4	Number		Y		
5-4	이미지파일명	file_name	100	String		Y		
5-5	이미지라이선스	license	100	String				
5-6	이미지촬영일자	date_captured	100	String				
	어노테이션정보	annotations		List				List of JSON Object
6-1	어노테이션식별자	id	100	String		Y		
6-2	연관이미지식별자	image_id	100	String		Y		
6-3	어노테이션텍스트	text	1000	String		Y		
	어노테이션속성	attributes		Object				JSON Object
6-4-1	종류	type	50	String				공통 : 글자(음절), 단어(어절), 문장
6-4-2	나이	age	3	String				손글씨 : 정보 수집자의 나이
6-4-3	성별	gender	1	String				손글씨 : 남, 여
6-4-4	직업	job	50	String				손글씨 : 정보 수집자의 직업
6-4-5	글꼴	font	50	String				인쇄체 : 수집된 글꼴
	라이선스	licenses		List				List of JSON Object
7-1	라이선스명	name	100	String		Y		
7-2	라이선스URL	url	200	String		Y		
참고사항								

● C.2 Text-in-the-wild 데이터 Structure

Label Structure

프로젝트	한국어 글자체 이미지 AI데이터 구축	단계	설계
문서번호	EDK_AM_A_13 버전 1.2	작성일자	6/28/2019

인터페이스ID		Structure Name	구축DB2 Label Structure
Label 구조	JSON	개발유형	

NO	항목명		길이	타입	코드 여부	필수 여부	Name/Value	비고
	한글명	영문명						
	데이터셋정보	info		Object				JSON Object
1	데이터셋명	name	100	String		Y		
2	데이터셋상세설명	description	1000	String				
3	데이터셋URL	url	200	String				
4	데이터셋생성일자	date_created	100	String		Y		
	이미지정보	images		List				List of JSON Object
5-1	이미지식별자	id	100	String		Y		
5-2	이미지너비	width	4	Number		Y		
5-3	이미지높이	height	4	Number		Y		
5-4	이미지파일명	file_name	100	String		Y		
5-5	이미지라이선스	license	100	String				
5-6	이미지촬영일자	date_captured	100	String				
	어노테이션정보	annotations		List				List of JSON Object
6-1	어노테이션식별자	id	100	String		Y		
6-2	연관이미지식별자	image_id	100	String		Y		
6-3	어노테이션텍스트	text	1000	String		Y		
6-4	어노테이션속성	attributes		Object				
6-5	어노테이션바운딩박스	bbox	4	List		Y		List of [x, y, width, height]
	라이선스	licenses		List				List of JSON Object
7-1	라이선스명	name	100	String		Y		
7-2	라이선스URL	url	200	String		Y		
참고사항								

04

인도(人道) 보행 영상 AI 데이터 구축 가이드라인

TestWorks

	성명	소속	직위
가이드라인 제안	금효영	테스트웍스	수석연구원
가이드라인 작성	금효영	테스트웍스	수석연구원
	김효경	테스트웍스	수석연구원
	전지윤	테스트웍스	연구원
	주경돈	KAIST RCV연구실	박사 후 연구원
	박기백	KAIST RCV연구실	박사과정
	신호욱	셀렉트스타	대표
	이정희	디투리소스	이사
전문가 검토 위원	민동보	이화여자대학교	교수
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 테스트웍스/ 참여: 셀렉트스타, 척수장애인협회, RCV, 디투리소스

서 문

1. 인도보행 데이터셋 구축의 목적

이동을 위해 휠체어와 같은 이동보조기구를 사용해야 하는 장애인들에게는 가까운 거리의 일상적인 외출조차 신체적 상해까지 발생할 수 있는 위험한 일로, 조력자 없는 외출은 그들에게 큰 모험이 될 수 있다. 다행히도 로봇과 인공지능 분야의 급속한 발전으로 인해 이들의 안전한 이동을 지원하는 기술들이 빠르게 개발되고 있어 머지않아 이들도 비장애인들처럼 자유롭게 이동할 수 있을 것으로 기대된다.

자율이동 보조기술의 발전을 위해서는 장애인의 주요 이동 통로인 '인도'에 대한 다각적이고 방대한 데이터가 필요하지만, 아직까지 국내외에서 이런 데이터들이 대규모로 구축된 사례는 없다. 인도 보행 데이터셋 구축 사업은 장애인들의 자유롭고 안전한 이동을 위한 기술뿐만 아니라 향후 '인도'에서 이루어지는 일반적인 이동과 관련된 기술 개발에 확장 적용할 수 있는 범용 데이터셋 구축을 목적으로 한다. 본 문서는 인도 보행 영상 데이터셋 구축에 필요한 구체적인 절차와 세부지침을 규정한다.

2. 주요 내용 요약

인도보행 영상 데이터셋은 데이터 설계 → 원천 영상 데이터 수집 → 데이터 가공 → 데이터 검증의 과정을 거쳐 구축한다. 장애인협회와의 인터뷰 및 자료 조사를 통해 장애인의 자유롭고 안전한 이동을 위해 해결해야 하는 문제들을 정확하게 규정하고 이를 기반으로 데이터 구조를 설계한다. 휴대폰과 스테레오 카메라로 '인도' 영상 데이터를 수집한 후 의도하지 않게 수집된 개인정보(초상권, 차량번호판 등)에 대해 비식별화 작업을 수행한다. 이후, 해당 영상 데이터를 대상으로 다양한 가공 작업을 체계적인 프로세스 상에서 수행하여 사용 목적에 적합한 데이터셋을 구축하고, 객체인식 알고리즘, 노면인식 알고리즘을 이용한 성능 평가를 통해 데이터를 검증한다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

인도(人道) 보행 영상 데이터셋 구축

Korea Sidewalk Image Dataset

04

1. 적용 범위

본 가이드라인에서 규정된 용어, 개념과 정의, 작업 절차, 작업 실행 방법과 각종 작업 세부지침들은 인도보행 영상 데이터셋 구축하기 위해 수행하는 작업들에 적용된다.

2. 인용 표준

- 해당 사항 없음

3. 용어 정의

3.1 태깅(tagging)

영상 데이터를 분석해서, 영상에 표시할 객체에 객체클래스를 부여하는 행위를 말한다.

3.2 어노테이션(annotation)

영상 데이터에서 기계학습에 필요한 객체를 표시하기 위해 도형 그리기를 수행하고 객체클래스를 부여하는 일련의 과정을 말한다.

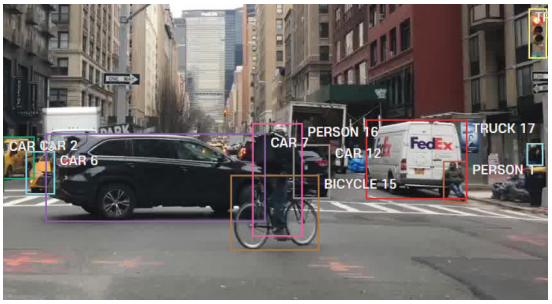
3.3 레이블(label)

어노테이션된 객체의 종류, 즉 객체의 클래스를 말한다.

3.4 바운딩 박스(Bounding Box)

영상 데이터 위에 추출할 객체의 위치를 표시하기 위해 그리는, 네 변이 이미지 상에서 수직/수평 방향을 향하는 직사각형 모양의 상자를 지칭한다.

[그림 3-1] 예시 : box로 표시된 영역의 Text를 입력



3.5 세그멘테이션(Segmentation)

영상 데이터 위에 객체가 어느 위치에 포함되어 있는지 '픽셀 단위로' 분할하기 위해 객체의 모양대로 객체를 추출하는 것을 지칭한다.

[그림 3-2] 세그멘테이션 예시

3.6 오검출

가공된 데이터를 검수하는 과정에서 아래와 같은 오류를 범하여 태깅한 객체가 발견되는 경우를 말한다.

가. 잘못된 레이블을 붙여 태깅한 경우
나. 레이블에 포함되지 않은 객체를 태깅한 경우

3.7 미검출

가공된 데이터를 검수하는 과정에서 태깅해야 할 객체를 태깅하지 않은 오류가 발견된 경우를 말한다.

3.8 과검출

가공된 데이터를 검수하는 과정에서 정의된 객체에 해당하는 객체이지만 객체의 크기, 선명도 등에 의해 태깅에서 제외해야 할 객체를 태깅하는 오류가 발견된 경우를 말한다.

3.9 어노테이션 범위불량

객체를 어노테이션할 때 가이드라인에서 규정한 범위에 못미치거나 벗어나서 어노테이션하는 오류가 발견되는 경우를 말한다.

4. 약어

- 해당 사항 없음

5. 인도보행 영상 데이터셋 구축

5.1 인도보행 영상 데이터셋 구축 절차

인도보행 영상 데이터셋을 구축하기 위한 절차는 아래의 표와 같다.

[표 5-1] 인도보행 영상 데이터셋 구축 절차

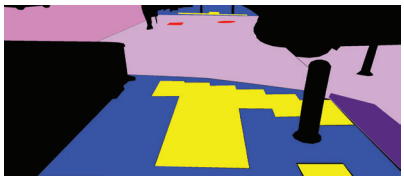
단계	세부 절차	설명	필수 여부
1.설계	1.1 데이터 설계	장애인협회와의 인터뷰 및 자료 조사를 기반으로 다양한 객체클래스 설계 및 데이터 구조 설계	필수
2. 수집	2.1 원천 데이터 선정	원천 데이터 수집 방법 및 규모 설정	필수
	2.2 원천 데이터 수집 교육	데이터 수집 및 정제 기준 교육	필수
	2.3 원천 데이터 수집	모바일폰과 ZED Stereo 카메라로 촬영	필수
3. 정제	3.1 원천 데이터 검수	영상의 흔들림, 촬영 상황 등을 지정된 숙련 검수자가 검수	필수
	3.2 원천 데이터 비식별	수집된 영상 데이터에 의도하지 않게 포함된 개인정보 비식별화	필수
4. 가공	4.1 가공 인력 양성 교육 시행	인공지능 학습용 데이터 가공에 필요한 작업 교육과 훈련	필수
	4.2 데이터 가공	수행할 가공 작업 - 바운딩 박스(Bounding Box) - 폴리곤 세그멘테이션(Polygon Segmentation) - 서피스 마스킹(Surface Masking) - Depth Prediction	필수
5. 검수	데이터 검수	체계적인 3단계 검수 과정으로 99% 데이터 품질* 확보 - 1단계: 작업자 간 검수 - 2단계: 매니저 검수 - 3단계: 최종 검수	필수

* 데이터 품질의 자세한 내용은 '10.1 인도보행 영상 데이터셋 품질 정의' 참고

5.2 인도보행 영상 데이터셋 산출물

인도보행 영상 데이터셋 구축 절차에 따른 단계별 산출물은 다음과 같다.

[표 5-2] 인도보행 영상 데이터셋 산출물

	입력물	산출물
설계	—	—
수집	—	인도보행 원천 영상 이미지
정제	인도보행 원천 영상 이미지 [비식별화 전] 	비식별화된 인도보행 영상 이미지 [비식별화 후]* 
가공 I	비식별화된 인도보행 영상 이미지	바운딩 박스  폴리곤 세그멘테이션  서피스 마스킹 
가공 II		Depth Prediction 
검수	—	—

상기 표의 모든 이미지는 개인정보 보호를 위하여 비식별화되었음.

* 이해를 돕기 위해 비식별화된 범위를 빨간 사각형으로 표시.

6. 인도보행 영상 데이터셋 설계

'장애인 인도보행을 위한 이동권/접근권' 향상에 기여하는 데이터셋 구축을 위해 장애인협회 관계자와의 심층 인터뷰 및 자료 조사를 통해 장애인들의 인도보행에 있어서의 어려움과 문제점을 파악하여 이를 데이터셋 설계에 반영한다. 특히, 인도보행의 어려움으로서 인도상의 다양한 장애물, 노면의 재질, 파손, 특수성 등을 고려하여 이를 반영한 다양한 객체클래스를 설계한다.

● 6.1 인도보행 영상 데이터셋 객체클래스 정의

인도보행 영상 데이터셋 객체클래스는 아래의 표와 같다.

[표 5-1] 인도보행 영상 데이터셋 구축 절차

대분류	소분류	객체클래스
장애물	이동체	barricade, bicycle, bus, car, carrier, cat, dog, motorcycle, movable_signage, person, scooter, stroller, truck, wheelchair
	고정체	bench, bollard, chair, fire_hydrant, kiosk, parking_meter, pole, power_controller, potted_plant, stop, table, traffic_light_controller, traffic_light, traffic_sign, tree_trunk

대분류	레이블명(뜻)	속성
노면	sidewalk (인도)	blocks(보도블럭), cement(시멘트), urethane(우레탄), asphalt(아스팔트), soil_stone(흙, 돌, 비포장도로), damaged(인도의 파손), other(정의된 인도의 속성 외의 경우)
	braille_guiding_block (점자블럭)	normal, damaged(점자블럭 파손)
	roadway(차도)	normal, crosswalk(횡단보도)
	alley(이면도로)	normal, crosswalk(횡단보도), speed_bump(과속방지턱), damaged(이면도로 파손)
	bike_lane (자전거도로)	- (속성값 없음)
	caution_zone (주의구역)	stairs(계단), manhole(맨홀), tree_zone(가로수 한그루 식재를 위한 바닥면 구역), grating(그레이팅), repair_zone(보수구역)

* 각 객체클래스의 자세한 내용은 '9.2.2 바운딩박스 레이블 리스트' 및 '9.4.2 서피스 마스킹 레이블 리스트' 참고

** 일부 객체는 상황에 따라 위 분류에서 벗어날 수 있음

7. 원천 영상 데이터 수집

원천 영상 데이터는 보행이 가능한 구역인 인도를 중심으로 그 주변 환경을 수집자가 휴대폰이나 ZED Stereo 카메라로 직접 촬영한 영상 데이터를 의미한다. 원천 영상 데이터는 크라우드 소싱으로 수집하거나 해당 사업 참여자가 직접 촬영하여 수집한다.

향후 활용 가능성, 시범 모델 및 서비스 개발에 필요한 학습량, 정제과정의 유효데이터 수 등을 고려해 수집할 원천 데이터의 양을 결정한다. 장소, 날씨, 시간 등 다양한 촬영 조건, 촬영 해상도, 촬영 방식을 정확하게 규정된 수집 가이드를 작성하고 수집자들이 유효한 인도보행 영상을 촬영하도록 교육한다.

● 7.1 원천 데이터 수집

셀렉트스타가 운영하는 모바일 앱 서비스 기반 크라우드 소싱과 수요자로 사업에 참여하는 한국척수장애인협회가 선정한 수집자들의 직접 촬영으로 원천 데이터를 수집한다.

[표 7-1] 촬영 방식 및 촬영 해상도

촬영방식	촬영도구	규격
크라우드소싱	휴대폰 앱	1920 * 1080 (px)
직접촬영	ZED Stereo 카메라	1920 * 1080 (px, 좌우 각각)

원천 영상 데이터는 장소, 날씨, 시간 등 다양한 촬영 조건을 규정하여 수집한다.

[표 7-2] 다양한 촬영 상황

항목	종류
장소	인도, 이면도로, 횡단보도 등
날씨	맑음 흐림(구름), 비 - 환경정보로는 비 옴/비 오지 않음으로만 제공
시간	오전, 오후, 밤 - 환경정보로는 낮/밤으로만 제공

한국척수장애인협회 선정 수집자의 직접 촬영 상황은 다음과 같다.

[표 7-3] 다양한 촬영 상황

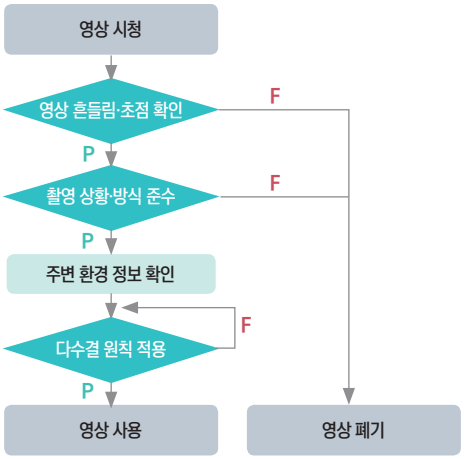
수집자	휠체어 이동 속도	촬영 매크로	촬영지역
수집자 1	5~10km/h	1초 단위	대전광역시 및 세종시 전역
수집자 2	5~10km/h	1초 단위	서울특별시 전역
수집자 3	5~10km/h	1초 단위	김포시 및 고양시 전역
수집자 4	5~10km/h	1초 단위	부산광역시 전역

8. 원천 영상 데이터 정제

● 8.1 원천 영상 데이터 정제 절차

인도보행 영상 데이터셋 정제 절차는 아래와 같다.

[표 8-1] 인도보행 영상 데이터셋 정제 절차

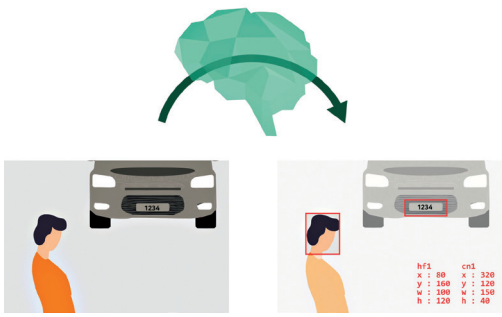
① 영상시청: 숙련 검수자가 앱을 통해 영상 확인 ② 영상 흔들림·초점 확인: 영상이 흔들리지 않았는지, 초점이 맞았는지 확인 ③ 촬영 상황·방식 준수: 영상이 규정된 촬영 상 조건과 촬영 방식을 준수하였는지 확인 ④ 주변 환경 정보 확인: 수집자가 입력한 영상의 촬영 정보(날씨, 유의미한 객체 존재 유무 등) 확인 ⑤ 다수결 원칙 적용: 우선 2명의 검수자가 참여하여 상기 과정으로 검수. 검수 결과 2명의 검수자 간 의견이 다를 땐 제3의 검수자를 투입하여 다수 의견으로 영상 사용/폐기 여부 판정	

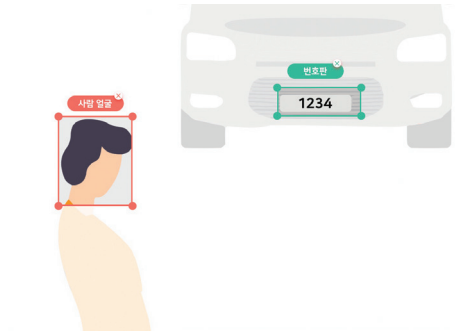
● 8.2 원천 영상 데이터 비식별화

원천 데이터 수집 과정에서 의도하지 않은 개인 정보가 수집될 수 있고 이는 차후 공개데이터로 사용할 때 초상권이나 기타 개인 권리를 침해하는 법적 분쟁 소지가 있어 비식별화 기술을 활용하여 사용 개인의 권리를 보호하고 법적 분쟁을 미연에 방지한다.

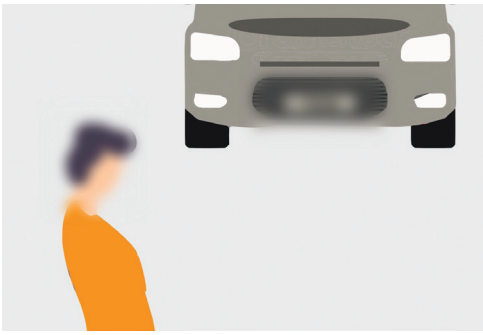
비식별화 작업은 아래 <표8-2> 영상 데이터 비식별화 작업의 순서대로 딥러닝에 기반하여 개발한 비식별화 알고리즘을 학습시켜 자동으로 수행한다.

[표 8-2] 영상 데이터 비식별화 작업

 [1] 딥러닝 기반 얼굴 및 차량 번호판 검출 알고리즘 개발	 [2] 수집된 영상 내의 얼굴 및 차량번호 판의 바운딩 박스 정보 추출



[3] 추출된 얼굴과 차량 번호판의 바운딩 박스 정보의 오검출, 미검출, 과검출, 어노테이션 범위불량에 대해 어노테이션 도구를 사용하여 수정 및 삭제 등 보완 작업 진행



[4] 영상 내 얼굴과 차량 번호판의 바운딩 박스에 해당하는 부분에 대해 OpenCV Library를 활용한 Blur 처리 프로그램으로 비식별화 완료

○ 8.2.1 얼굴 비식별화 알고리즘

얼굴 비식별화 알고리즘은 TensorFlow를 이용한 python기반의 오픈소스 라이브러리 (https://github.com/yeephycho/tensorflow-face-detection) 를 이용하여 개발한다. 해당 오픈소스는 Apache 2.0.License로 인도보행 영상 데이터셋 맞춤 비식별화 작업을 위하여 라이센스 범위 내 수정하여 사용한다. 해당 소스코드에 추가적인 옵션을 적용하여 추가적인 옵션을 적용할 수 있도록 개선되었다. 첫번째로 비식별화 작업에 그려지는 박스의 크기를 조절 가능하다. 프로그램을 실행할 때 커맨드라인에서 – width_ratio와 – height_ratio를 0~1로 조절 할 수 있다. 두번째로 박스가 그려진 이미지 파일을 생성할지 결정할 수 있다. 프로그램을 실행할 때 커맨드라인에서 – save_result_img 를 True로 설정하면 해당 이미지를 생성가능하다. 미검출/오검출된 얼굴에 대해서는 수작업을 통하여 보정한다.
중간 산출물로 제출된 영상 데이터들에 나타나는 사람 얼굴은 모두 어노테이션 툴 상에서 수작업으로 어노테이션한 후 blurring하여 비식별화되었다.

○ 8.2.1.1 얼굴 비식별화 기술 적용 가이드라인

기준	세부 내용
100x100px 이상 얼굴만을 대상으로, IoU Threshold 0.3 기준 평균 정밀도가 0.3 이상인가?	100x100px 이상 얼굴만을 대상으로, IoU(Intersection over Union) Threshold를 0.3을 기준으로 잡았을 때, 평균 정밀도(Average Precision)이 0.3 이상인가?

- 정밀도(Precision): 물체를 검출했을 때 맞게 검출한 비율.
- 재현율(Recall): 찾아야할 물체들 중 검출한 비율.
- 평균 정밀도(Average Precision): 정밀도와 재현율은 Confidence Score Threshold에 따라 달라지는데, Confidence Score Threshold가 높을수록, 즉 확실한 것들만 검출할수록 정밀도는 올라가지만 재현율은 낮아진다. 반대로, Confidence Score Threshold가 낮을수록, 가능한 많은 것들을 찾으려하기 때문에

재현율은 올라가나 정밀도는 낮아진다. 따라서, Confidence Score Threshold에 따라 바뀌는 정밀도를 평균한 평균 정밀도는 Object Detection 분야에서 주로 쓰이는 Metric이다.

- Confidence Score Threshold : 얼굴/번호판 검출 모델은 검출 결과 박스마다 검출 신뢰도(Confidence Score) 또한 예측함. 따라서 이 신뢰도 값이 높을수록 정확한 검출일 확률이 올라가기 때문에, 신뢰도 값이 일정 값(Threshold) 이상인 결과들만 최종 검출 결과에 포함시킴.
- IoU : Intesection over Union의 약자로, 두 박스의 교집합 면적 / 합집합 면적 을 의미함.

○ 8.2.2 차량 번호판 비식별화 알고리즘

차량 비식별화 알고리즘은 PyTorch를 이용한 python기반의 오픈소스 라이브러리 (https://github.com/eriklindernoren/PyTorch-YOLOv3) 를 이용하여 개발한다.
해당 소스코드에 추가적인 옵션을 적용하여 추가적인 옵션을 적용할 수 있도록 개선되었다. 첫번째로 비식별화 작업에 그려지는 박스의 크기를 조절 가능하다. 프로그램을 실행할 때 커맨드라인에서 – width_ratio와 – height_ratio를 0~1로 조절 할 수 있다. 두번째로 박스가 그려진 이미지 파일을 생성할지 결정할 수 있다. 프로그램을 실행할 때 커맨드라인에서 – save_result_img 를 True로 설정하면 해당 이미지를 생성가능하다. 미검출/오검출된 얼굴에 대해서는 수작업을 통하여 보정한다.
중간 산출물로 제출된 영상 데이터들에 나타나는 자동차 번호판은 모두 어노테이션 툴 상에서 수작업으로 어노테이션한 후 blurring하여 비식별화되었다.

○ 8. 2. 2.1 차량 번호판 비식별화 기술 적용 가이드라인

기준	세부 내용
장축 100px 이상 번호판들을 대상으로, IoU Threshold 0.3 기준 평균 정밀도가 0.3 이상인가?	장축 100px 이상 번호판만을 대상으로, IoU(Intersection over Union) Threshold를 0.3을 기준으로 잡았을 때, 평균 정밀도(Average Precision)가 0.3 이상인가?

9. 인도보행 영상 데이터셋 가공

● 9.1 인도보행 영상 데이터셋 가공 종류

인도보행 영상 데이터셋 가공 종류는 다음과 같다.

[표 9-1] 인도보행 영상 데이터셋 가공 종류

가공 종류	가공 도구	목적
바운딩 박스(Bounding Box)	AAT * (학습용 데이터 제작도구)	인도보행 장애물 어노테이션
폴리곤 세그멘테이션 (Polygon Segmentation)	AAT	
서피스 마스킹(Surface Masking)	AAT & 자체 마스킹 도구	노면 마스킹
Depth Prediction	ZED Stereo Camera 촬영 후 후처리 프로그래밍	깊이 정보

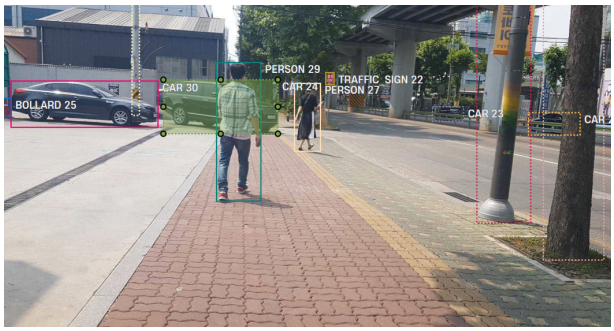
* AAT는 '부록 I'를 참고

● 9.2 바운딩 박스(Bounding Box)

○ 9.2.1 바운딩 박스 가이드라인

[표 9-2] 바운딩 박스 가이드라인

기준	세부 내용
대상	*<표 9-3> 바운딩 박스 레이블 리스트*에 정의되어 있는 객체는 모두 어노테이션함을 원칙으로 하되 아래의 경우는 어노테이션하지 않음 • 객체의 식별이 불가능한 경우(너무 흐리거나 차나 버스 같은 객체가 너무 빨리 이동해서 위치 파악이 어려운 경우 등) • 객체의 가로, 세로의 너비 중 긴 면이 64px 이하인 경우 (단, 자동차(car), 버스(bus), 트럭(truck), 사람(person), 볼라드(bollard), 신호등(traffic_light), 교통 표지판(traffic_sign)의 경우 32px 이하) • 사진을 확대하지 않은 원본상태에서, 태깅 리스트에 포함되며 뚜렷하게 구분되는 가장 원거리 객체를 기준으로, 이 기준보다 먼 원거리의(뒷부분의) 객체는 태깅하지 않음. [어노테이션 대상 선별 기준선 예시] 

기준	세부 내용
사각형 박스 그리기	객체의 보이는 부분이 모두 포함되도록 사각형을 그리되 사각형의 경계와 객체의 경계에 빈틈이 없도록 최대한 꽉차게 그리는 것을 원칙으로 함. 단, 아래와 같은 상황일 때는 상황에 맞게 사각형 박스를 그려야 함. • 아래 예시 그림의 오른쪽에 촬영된 자동차처럼 해당 객체가 잘려서 촬영된 때에는 이미지 가장자리까지 사각형을 그려야 함. 자세한 내용은 아래에 있는 '가림' 항목을 참고 [잘림 현상이 발생한 경우 사각형 박스 그리기 예시]  • 해당 객체가 다른 객체에 의해 겹쳐 촬영된 때에는 객체의 온전한 모양을 추정하여 사각형을 그려야 함. 자세한 내용은 아래에 있는 "겹침" 항목을 참고 [겹침 현상이 발생한 경우 사각형 박스 그리기 예시]  • 해당 객체가 군집을 이루고 있을 경우, 객체 하나씩 모두 사각형을 그려야 함(단, 자전거는 제외, 자전거는 '<표 9-3> 바운딩 박스 레이블 리스트'의 'bicycle' 항목 참고) [군집 현상이 발생한 경우 사각형 박스 그리기 예시] 

<table>
<tr><th>기준</th><th>세부 내용</th></tr>
<tr><td>가림</td><td>
객체끼리 겹치거나 화각으로 인해 객체가 잘려서 촬영된 현상을 '가림 현상'이라 함. '가림 현상'이 나타난 객체는 객체의 온전한 모양을 추정하여 사각형을 그려야 하고, '가림'을 표시해주어야 함. 단, 객체가 잘려 온전한 모양을 추정할 수 없는 경우 객체가 잘려진 사진의 가장자리까지 사각형을 그린 후 '가림'을 표시해주어야 함.
객체가 가려진 정도에 따라 어노테이션 방식이 달라지므로, 다음과 같은 기준에 따라 사각형을 그려야 함
<table>
<tr><th>기준</th><th>어노테이션 방법</th></tr>
<tr><td>객체가 95% 보임</td><td>가려진 부분까지 사각형을 그림</td></tr>
<tr><td>객체가 20~95% 보임</td><td>가려진 부분까지 사각형을 그리고 가림 처리</td></tr>
<tr><td>객체가 20% 이하로 보임</td><td>객체에 사각형을 그리지 않음</td></tr>
</table>
가림 처리는 ai Annotation Tool의 기능을 사용하여 처리하고, 바운딩 박스 산출물(XML)에서 다음과 같이 표시됨.
<pre>...
<box z_order="9" ybr="509.20" xbr="1166.20" ytl="462.50" xtl="1092.01" occluded="1" label="barricade"></box >
<box z_order="0" ybr="485.60" xbr="1459.40" ytl="450.02" xtl="1412.42" occluded="1" label="car"></box >
<box z_order="0" ybr="762.90" xbr="1750.30" ytl="636.30" xtl="1594.40" occluded="0" label="potted_plant"></box >
...</pre>

occluded="1" : 가림 현상이 발생하여 표시된 경우
occluded="0" : 가림 현상이 발생하지 않은 경우

(xtl, ytl) : 사각형의 좌측 상단의 꼭짓점의 위치(x, y)
(xbr, ybr) : 사각형의 우측 하단의 꼭짓점의 위치(x, y)

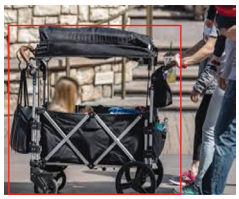
</td></tr>
<tr><td>이미지 품질</td><td>이미지 크기: 1920*1080(px)
가로, 세로의 길이가 1920*1080(px)이 아닌 경우 해당 이미지는 사용하지 않음</td></tr>
<tr><td>반사</td><td>거울과 같은 물체에 반사된 객체는 어노테이션하지 않음</td></tr>
<tr><td>광고물</td><td>인쇄된 객체는 어노테이션하지 않음. 광고물에 있는 객체가 실사인 경우라도 어노테이션하지 않음</td></tr>
</table>

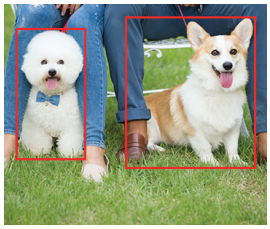
○ 9.2.2 바운딩 박스 레이블 리스트

바운딩 박스 가공 작업에서 어노테이션하는 레이블은 아래와 같다.

[표 9-3] 바운딩 박스 레이블 리스트

레이블	세부 내용
person	남녀노소, 동작/행위 여부와 상관없이 person에 포함, 라이더 역시 포함 **[주의사항]** 사람이 들고 있거나 착용하고 있는 물건들(휴대폰, 핸드백, 백팩 등)은 어노테이션하지 않음  
bicycle	자전거, 전동 자전거, 세발 자전거, 외발 자전거 포함 **[주의사항]** 자전거가 군집을 이루고 있을 경우 원칙적으로 어노테이션하지 않음. 단, 모여있는 자전거 중 외형이 가려지지 않아 뚜렷이 파악되고 온전히 추출할 수 있을 때는 2~3대까지만 어노테이션함 
motorcycle	오토바이, 스쿠터(550cc 이하도 포함)  
scooter	변속 기어가 없는 2륜차 혹은 두 발로 탈 수 있는 기구  

레이블	세부 내용
	 
bus	미니버스, 마을버스, 2층버스, 고속버스, 관광버스 **[주의사항]** 어린이집차와 같이 15인승 이하 승합차(스타렉스, 봉고, 카니발 등)는 car로 분류    
truck	1톤 트럭 이상, 레미콘, 포크레인, 탱크로리, 사다리차, 푸드트럭, 라보 포함     
car	승용차, 승합차 **[주의사항]** 열려있는 트렁크 문이나 차문도 모두 포함하여 어노테이션함. 단, 야구르트 차는 어노테이션하지 않음  
stroller	유모차, 노인용 유모차(senior stroller) 포함 **[주의사항]** 유모차의 내부가 보이지 않으면 유모차만 어노테이션함. 유모차를 탄 아이가 명확히 보일 경우에는 유모차와 아이 각각 어노테이션함.   
wheelchair	휠체어 (예 : 수동 휠체어 , 전동 휠체어) **[주의사항]** 사람이 타고 있을 경우, 휠체어와 사람을 각각 어노테이션함  

레이블	세부 내용
dog	개  
cat	고양이 
movable signage	1. 고정되지 않아 움직일 수 있는 홍보 전시물 (예; 입간판, 에어간판, 스탠드형 라이트 패널 등등) 2. 이동 안내판 포함(예: 주차금지 안내판) **[주의사항]** 게시판과 같이 한 곳에 부착되거나 고정된 안내판은 어노테이션하지 않음 
carrier	리어카, 핸드카트, 핸드카, 마트 손수레 포함. 단, 마트의 대용량 물류 이동 수레는 포함하지 않음 

레이블	세부 내용
tree trunk	가로수 (나무)의 주 기둥 (이미지에서 눈으로 확인되는 기둥) **[주의사항]** 기둥 굵기가 16px 이상이어야 하고 주 기둥에서 가지가 뻗어 나오기 전까지 어노테이션함  
potted plant	화분 (예 : 방부목 화분 , 길거리 화분) **[주의사항]** 화단은 어노테이션 하지 않음    
traffic light	신호등 (예 : 보행자용 , 차량용 신호등) / 32px 이상 육안으로 확인이 되는 경우 태깅 **[주의사항]** 뒷면이 찍혔을 때도 어노테이션함  
traffic sign	교통 표지판, 교통 보조 표지판, 장애인 주차장 표지판 포함(이정표, 도로명 등의 안내판은 제외) / 32px 이상 육안으로 확인될 경우 어노테이션함 **[주의사항]** 뒷면이 찍혔을 때도 어노테이션함. 여러 개 나열된 표지판의 경우, 각각 어노테이션함       

레이블	세부 내용
pole	대/기둥 (예: 신호등 대, 전봇대 대, 가로등 대, 도로표지판 대, 햇빛 파라솔 대, 현수막 대) **[주의사항]** 두께가 16px 이상이어야 함. 그 이상이어도 희미한 것은 어노테이션하지 않음
bench	2인 이상이 앉을 수 있는 기구 (의자) **[주의사항]** 꺾인 벤치의 경우: 가로와 세로를 따로 어노테이션함    
chair	1인이 앉을 수 있는 기구 (의자)    
table	탁자(원형 , 네모 , 플라스틱 탁자 등). 편의점 탁자에 설치된 파라솔은 어노테이션하지 않음 **[주의사항]** 벤치가 붙어 있는 테이블은 한꺼번에 테이블로 어노테이션함    

레이블	세부 내용
stop	버스정류장, 택시정류장 **[주의사항]** 앉을 수 있는 벤치나 의자가 설치된 정류장이어야 하고, 정류장 표지판만 있는 정류장은 제외함    [예제] - 왼쪽의 정류장 표지판은 제외 - 오른쪽의 정류장만 어노테이션함
kiosk	물품 및 서비스를 제공하는 조그만 점포와 독립된 단말기, ATM기기 **[주의사항]** 단, 좌판, 매대, 포장마차, 주차관리소 포함하지 않음    
fire hydrant	소화전     
parking meter	주차요금정산기 (예 : 무인요금 정산기 등)     

레이블	세부 내용
bollard	차량 진입 금지 설치물 (교통 원뿔, PE드럼 포함) 
barricade	이동 설치가 가능한 출입 금지 장애물(안전펜스, 가림막 펜스 포함) **[주의사항]** 여러 개의 바리케이트를 연속으로 세우거나 조립하여 설치한 경우, 각각의 바리케이트를 어노테이션함 
power controller	전력 제어함 
traffic light controller	신호 제어기 **[주의사항]** 신호등에 매달려있는 신호 제어기의 경우, 교통 경찰이 조작 가능한 위치에 있는 경우만 신호 제어기로 어노테이션함 

● 9.3 폴리곤 세그멘테이션(Polygon Segmentation)

○ 9.3.1 폴리곤 세그멘테이션 가이드라인

[표 9-4] 폴리곤 세그멘테이션 가이드라인

기준	세부 내용
대상	*9.3.2 폴리곤 세그멘테이션 레이블 리스트*에 정의되어 있는 객체는 모두 세그멘테이션함을 원칙으로 한다. 단, 아래와 같은 상황일 경우는 예외로 하여 객체를 세그멘테이션하지 않는다. • 객체가 무엇인지 확신할 수 없는 경우 • 객체의 가로, 세로의 너비 중 긴 면이 128px 이하인 경우 • 객체의 가로, 세로의 너비 중 긴 면이 64px 이하인 경우 (단, 자동차(car), 버스(bus), 트럭(truck), 사람(person), 볼라드(bollard), 신호등(traffic_light), 교통 표지판(traffic_sign)의 경우 32px 이하) • 객체의 70% 이상이 보이지 않는 경우(*단, 가림 또는 잘림의 현상으로 인해 객체가 30% 이상 보이지 않아도 객체의 클래스를 구분할 수 있다면 세그멘테이션을 함*) • 사진을 확대하지 않은 원본상태에서, 태깅 리스트에 포함되며 뚜렷하게 구분되는 가장 원거리 객체를 기준으로, 이 기준보다 먼 원거리의(뒷부분의) 객체는 태깅하지 않음 [어노테이션 대상 선별 기준선 예시] 
다각형 그리기	객체의 가시 면적에 대해 다각형을 표시한다. 다각형은 객체의 외곽선을 기준으로 10px 내외에 그려야 한다. 단, 아래와 같은 상황일 경우 해당 상황에 맞게 세그멘테이션을 해야 한다. • 객체 잘림 현상이 발생한 경우 잘림 현상이 발생한 면까지 포함하여 세그멘테이션을 해야 한다. • 객체 가림 현상이 발생한 경우 가림과 상관없이 객체의 가시영역만 세그멘테이션을 해야 한다.
이미지 데이터 품질	이미지 크기는 1920×1080(px)이여야 한다. 해상도는 동일하나 가로, 세로의 길이가 1920×1080(px)이 아닌 경우 해당 이미지는 세그멘테이션 작업을 할 수 없다.

기준	세부 내용
가는 영역 (Thin Structure)	아래와 같이 1px 내외의 영역일 경우는 정밀한 세그멘테이션이 어려운 경우로 판단하여 세그멘테이션 하지 않는다. • 전선 • 연결 장치(고리) • 수염 [가는 영역 다각형 그리기 예시] 
반사	거울에 비친 객체는 세그멘테이션하면 안 된다.
광고물	광고물에 있는 객체가 실사(*일러스트나 아이콘이 아닌 그림*)인 경우라도 세그멘테이션을 하지 않는다.
집단(그룹)	객체가 너무 밀집되어 있어 정확한 세그멘테이션을 하기 힘든 경우 집단에서 객체가 가장 선명하고, 외곽선이 뚜렷한 객체를 우선적으로 하며, 그 외는 세그멘테이션하지 않는다.

○ 9.3.2 폴리곤 세그멘테이션 레이블 리스트

폴리곤 세그멘테이션 가공 작업에서 어노테이션하는 레이블은 <표 9-3> 바운딩 박스 레이블 리스트와 동일하다.

● 9.4 서피스 마스킹(Surface Masking)

○ 9.4.1 서피스 마스킹 가이드라인

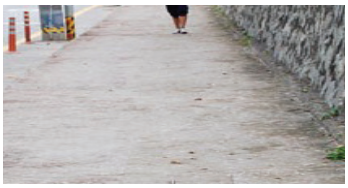
[표 9-5] 서피스 마스킹 가이드라인

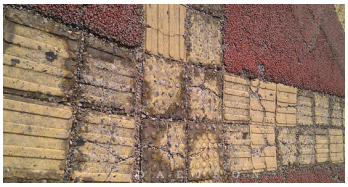
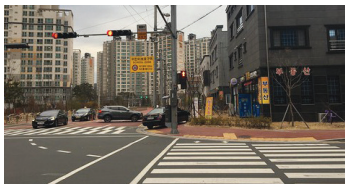
기준	세부 내용
대상	*9.4 Surface Masking*의 *<표 9.4.2-1> Surface Masking Label List*에 정의되어 있는 노면는 모두 마스킹함을 원칙으로 한다. 모든 노면은 노면 종류와 특수성/장애물 중 선택하여야 하며, 특수성과 장애물의 경우 노면 위 이중 마스킹을 해야 한다. • 특수성/장애물 : 주의 구역(caution_zone) • 특수성/장애물의 경우 타 객체를 먼저 그린 후 마지막에 그려야 한다.
마스킹 칠하기	노면의 가시 면적에 대해 마스킹 작업을 한다. 마스킹은 노면의 외곽선을 기준으로 2px 내외에 그려야 한다. 단, 아래와 같은 상황일 경우 해당 상황에 맞게 마스킹을 해야 한다. • 노면 잘림 현상이 발생한 경우 잘림 현상이 발생한 면까지 포함하여 마스킹을 해야 한다. • 특수성/장애요소를 마스킹해야 하는 경우 재질별/목적별 중 적합한 노면으로 선 마스킹 후 특수성/장애요소에 해당되는 요소를 마스킹해야 한다. 특수성/장애요소를 마스킹해야 하는 경우 노면 종류와 특수성/장애물 중 적합한 노면으로 선 마스킹 후 특수성/장애요소에 해당되는 요소를 마스킹해야 한다.
이미지 데이터 품질	이미지 크기는 1920×1080(px)이여야 한다. 해상도는 동일하나 가로, 세로의 길이가 1920×1080(px)이 아닌 경우 해당 이미지는 마스킹 작업을 할 수 없다.
구멍 (hole)	장애요소 중 구멍이 존재하는 객체일 경우 구멍까지 포함하여 마스킹을 해야 한다. • 맨홀의 구멍 • 그레이팅의 구멍 [가는 영역 다각형 그리기 예시] 
잡초	노면 위 잡초가 존재하는 경우 보행에 영향을 미치지 않는 선의 풀이라면 객체가 없다 가정하고 노면을 그린다.
사유지	사유지로 명확히 판단되는 노면은 그리지 않는다.

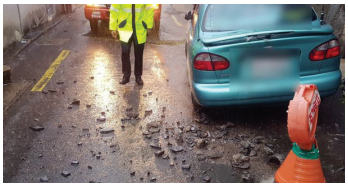
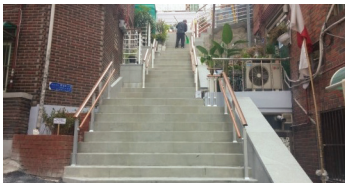
○ 9.4.2 서피스 마스킹 레이블 리스트

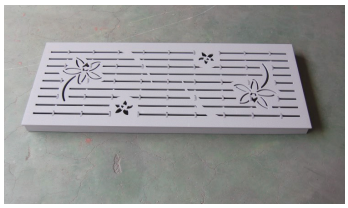
서피스 마스킹 가공 작업에서 어노테이션하는 레이블은 아래와 같다.

[표 9-6] 서피스 마스킹 레이블 리스트

레이블 이름	속성	세부 내용
sidewalk	blocks	인도의 보도블럭      
	cement	인도의 시멘트  
	urethane	인도의 우레탄  
	asphalt	인도의 아스팔트/아스콘  

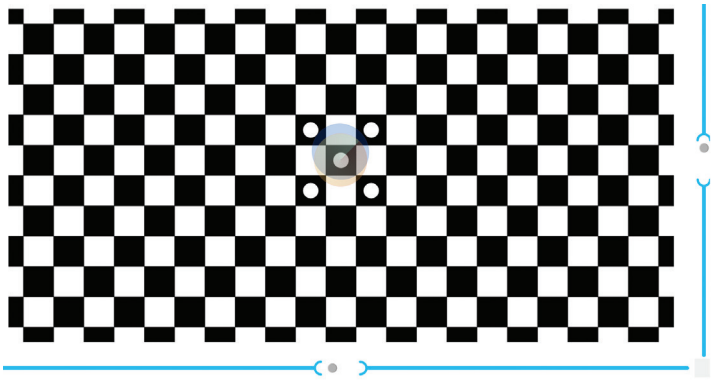
레이블 이름	속성	세부 내용
	soil_stone	흙, 돌, 비포장 도로    
	damaged	인도의 파손영역  
	other	위에 언급한 인도의 재질 외 모든 인도
braille_guiding_block	normal	점자블럭  
	damaged	점자블럭 파손  
roadway	normal	차도  
	crosswalk	차도의 횡단보도  

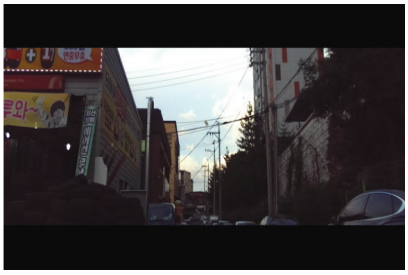
레이블 이름	속성	세부 내용
alley		이면도로  
	crosswalk	이면도로의 횡단보도  
	speed_bump	이면도로의 과속방지턱  
	damaged	이면도로 파손영역  
bike_lane		자전거 도로  
caution_zone	stairs	계단  

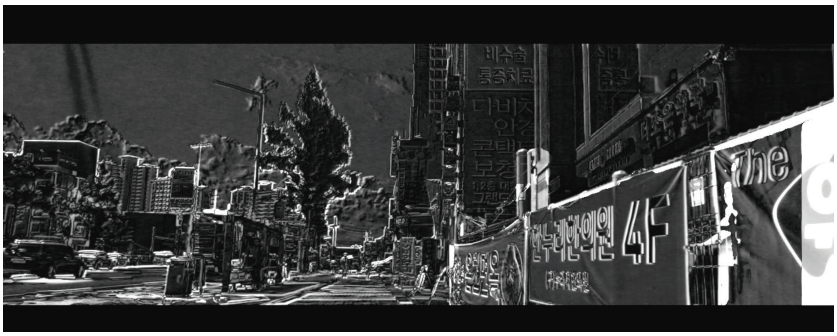
레이블 이름	속성	세부 내용
caution_zone	manhole	맨홀  
	tree_zone	나무 한그루가 식재되어 있는 바닥면의 구역 **[주의사항]** 보수가 필요한 가로수 영역 포함  
	grating	그레이팅 **[주의사항]** 디자인 그레이팅 포함  
	repair_zone	보수구역(가림막이 쳐져 있지 않은 보도(단, 이면도로 포함)를 수리하고 있는 영역)  

9.5 Depth Prediction

9.5.1 Depth Prediction 가이드라인

기준	세부 내용
촬영 도구	STEREOLABS ZED Camera(ZED Stereo Camera), GPU 탑재한 노트북(*https://www.nvidia.co.kr/object/cuda_learn_products_kr_old.html*), 3축 짐벌
Calibration	Stereo camera 영상을 교정하여 disparity를 구하기 위해 촬영 전 calibration 작업을 진행해야 한다. ZED Calibration 앱*을 사용해서 직접 교정한다. 교정 앱은 모니터에서 ZED 인식패턴을 인식하게 함으로써 작동한다. 밝은 조명을 끄고 화면을 반사될 수 있는 창문이 있다면 블라인드를 닫는다.  중앙에 빨간색 점이 있는 격자를 볼 수 있고, 모니터에서 ZED를 조준하면 조준하고 있는 곳에 파란 점이 보인다. 해당 작업의 목표는 파란 점을 빨간점과 일치시키는 것이다. 해당 작업은 매일 촬영 전 recalibration을 통해 스테레오 카메라를 보정한다. * ZED Calibration 앱 출처: https://support.stereolabs.com/hc/en-us/articles/360011828773-Can-I-recalibrate-my-ZED-camera-
Depth Image 추출	스테레오 카메라를 사용해서, 촬영을 하면 프레임이 캡처되는 즉시 카메라가 비디오 데이터를 전송할 수 있다. 전송된 데이터 중 왼쪽 렌즈로부터 추출된 이미지 L_Image 와 오른쪽 렌즈를 통해 추출된 이미지 R_Image를 가지고 Depth Image를 생성 추출하는 작업을 진행한다.   1차 추출된 RGB 이미지를 아래와 같은 기준으로 품질 검증을 진행해 Depth 추출에 적합하지 않은 이미지를 삭제하는 작업을 진행한다.

기준	세부 내용
이미지 검수 기준	빛 번짐, 흔들림, 좌 또는 우 이미지에 휠체어 부품이나, 촬영자의 신체가 찍힌 사진, Y축 Rectification의 Disparity의 오류를 통해 이미지를 1차 검수한다. • Y축 Rectification의 Disparity   • 빛 번짐/직사광선  • 휠체어 또는 신체 일부가 보이는 사진 
유효 데이터 추출 과정	이미지 검수가 끝난 좌,우 두 개의 이미지는 Depth Prediction 이미지 생성을 위해, GA-net을 통해 Depth 이미지를 추출한다. GA-net의 학습에 쓰인 Ground truth와의 Resolution을 맞추기 위해 1920*592 사이즈로 cropping 한다. 최종 선택된 데이터 셋을 기준으로 개인정보를 비식별 한 후 GA-net을 통해 다시 한번 Depth 이미지를 추출하여 데이터 셋을 구축한다.

기준	세부 내용
Depth 산출물	GA-net을 통해 생성된 거리 인식을 위한 학습데이터는 아래와 같이 6개의 파일을 포함한다. • disp16 : disparity의 data(좌우 이미지에서 대응점의 x축 상 pixel 차이, 거리감을 나타내는 data)가 저장된 파일  • disp : disparity의 시각화(visualization)한 파일  • confidence_save : 각 픽셀에 대해 좌우 이미지의 밝기 차이(0-255)에 대한 값이 들어있는 파일  • confidence : confidence_save의 값을 분석하여 2진화 변환한 파일 : 전체 pixel 개수 중 1의 개수의 합이 20% 이상인 것을 취함 - 1 : true confidence(백색) - 0 : false confidence(흑색)

기준	세부 내용
	 • crop_L : GA-net의 입력 이미지으로 쓰이는 Stereo Camera의 왼쪽 이미지  • crop_R : GA-net의 입력 이미지으로 쓰이는 Stereo Camera의 오른쪽 이미지 

10. 인도보행 영상 데이터셋 검수

● 10.1 인도보행 영상 데이터셋 품질 정의

$$(\text{데이터품질}) = \frac{x-(y_1+y_2+y_3+y_4)}{x} \times 100$$

x = 이미지 내 어노테이션해야 할 객체의 수

y_1 = 오검출된 객체의 수

y_2 = 미검출된 객체의 수

y_3 = 과검출된 객체의 수

y_4 = 어노테이션 범위 불량인 객체의 수

인도보행 영상 데이터셋 가이드라인을 기준으로 어노테이션 작업을 수행하여 가공된 데이터셋의 '데이터 품질'은 검수를 통해 얻은 오검출, 미검출, 과검출, 어노테이션 범위불량의 빈도를 위의 식으로 산출하여 결정한다. 해당 데이터셋에 오검출, 미검출, 과검출, 어노테이션 범위불량인 객체의 빈도가 낮을수록 데이터 품질 지표가 높아진다.

바운딩 박스는 인도보행 영상 데이터셋 품질 정의를 동일하게 적용하여 데이터 품질로 정의한다.
폴리곤 어노테이션은 인도보행 영상 데이터셋 품질 정의를 토대로 아래와 같이 데이터 품질을 정의한다.

$$(\text{데이터품질}) = \frac{x-(y_1+y_2+y_3+y_4+y_5)}{x} \times 100$$

x = 이미지 내 어노테이션해야 할 객체의 수

y_1 = 오검출된 객체의 수

y_2 = 미검출된 객체의 수

y_3 = 과검출된 객체의 수

y_4 = 어노테이션 범위 불량인 객체의 수

y_5 = 객체 간 그룹화 오류의 수

서피스 마스킹은 인도보행 영상 데이터셋 품질 정의를 토대로 아래와 같이 데이터 품질을 정의한다.

$$(\text{데이터품질}) = \frac{x-(y_1+y_2+y_3+y_4+y_5)}{x} \times 100$$

x = 이미지 내 어노테이션해야 할 객체의 수

y_1 = 오검출된 객체의 수

y_2 = 미검출된 객체의 수

y_3 = 과검출된 객체의 수

y_4 = 어노테이션 범위 불량인 객체의 수

y_5 = 어노테이션 순서 불량인 객체의 수

뎁스 프리딕션은 인도보행 영상 데이터셋 검수 기준과 동일하게 적용한다.

● 10.2 인도보행 영상 데이터셋 검수 절차

인도보행 영상 데이터셋은 아래와 같이 3단계의 검수 절차를 거쳐 오검출, 미검출, 과검출, 어노테이션 범위불량인 객체가 없는지 검수하여 데이터 품질을 관리한다.

[표 10-1] 인도보행 영상 데이터셋 3단계 검수 절차

단계	세부 내용
Peer Review	• 작업자 간 리뷰를 진행한다.
Manager Review	• 작업자-매니저 간 리뷰를 진행한다. • 품질 기준에 미달한 데이터셋은 재작업을 수행한다.
Final Audit	• 최종 전달(delievery)이전에 2단계 검수 절차에서 품질 기준을 충족한 데이터 중 샘플을 추출하여 최종 검수한다.

11. 인도보행 영상 데이터셋 검증

● 11.1 '객체/거리 인식 모델' 프로토타입 모델

○ 11.1.1 '객체/거리 인식 모델' 프로토타입 모델

'객체/거리 인식 모델' 프로토타입 모델은 본 과제에서 생성한 데이터셋을 검증하기 위해 이용한 객체인식/깊이정보인식 네트워크를 융합한 AI 시범 서비스 모델이다.
프로토타입 모델은 ZED 카메라의 좌/우 칼라 영상(이미지)를 입력으로 하며, 객체인식 네트워크는 좌측 영상에서 객체를 바운딩 박스 형태로 검출하고, 깊이정보인식 네트워크는 좌/우 영상에서부터 깊이 정보를 예측한다. 두 네트워크의 출력을 활용하여 인식된 객체까지의 거리를 계산한다.

[표 11-1] '객체/거리 인식 모델' 프로토타입 모델 개념

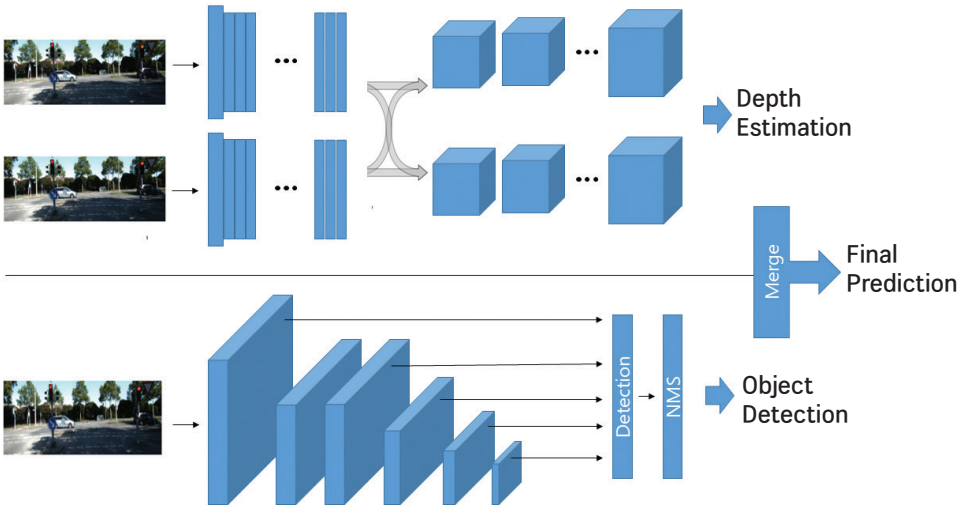
종류	세부 설명
객체 인식 (Object Detection)	생성한 데이터셋(보행에 관찰되는 객체에 대한 데이터셋)에 대한 객체 인식 네트워크(Faster-RCNN)을 학습하고, 시야 범위(FoV) 내의 객체를 바운딩 박스 형태(이미지에서 박스의 좌측 상단 x, y좌표와 박스의 높이와 너비)로 검출한다.
깊이 인식 (Depth Prediction)	생성한 데이터셋(좌/우 영상 및 깊이 정보)로부터 깊이 인식 네트워크를 학습하고, 서비스 시에는 입력된 좌/우 영상을 입력으로 좌측 영상 기준의 깊이 정보를 인식한다.

○ 11.1.2 '객체/거리 인식 모델' 프로토타입 구현 방안

	세부 내용
딥러닝 연산 가능한 고성능 GPU 활용	• 각 픽셀마다 담겨있는 깊이(depth) 정보를 정확하게 추출, 분석 및 연산하기 위해 고성능의 GPU를 모델 개발에 활용 예정 • 또한, 많은 양의 연산을 처리하기 위해 노트북이 아닌 데스크탑 이용
원천 데이터 확보 및 활용방안	• 휠체어에 부착된 스테레오 카메라(예, 상용 ZED 카메라)를 이용해 취득한 영상 확보, 확보 영상 범위 내의 객체 인식 및 거리(깊이)정보를 추정 및 개발 활용
모델 구현 방안	• ZED 깊이 인식 카메라 촬영 이미지 데이터 활용, 거리(Depth) 예측한 딥러닝 네트워크를 차용 • 객체 인식 관련 YOLO 네트워크를 학습하여 활용 • 깊이 정보 인식 관련 PSM 네트워크(혹은 그 변형버전)를 학습하여 활용 • 이미지상의 인식된 객체 위치와 해당 영역의 깊이 정보를 융합

○ 11.1.3 '객체/거리 인식 모델' 프로토타입 모델 산출물

객체 인식 네트워크와 깊이 인식 네트워크에서 예측한 정보를 가지고, 현재 기준 위치(왼쪽 영상 기준)로부터 어느 방향(각도), 어느 거리에 어떤 객체가 있는지 계산한다.



● 11.2 '인도 노면 안정성 모니터링 시험 서비스' 프로토타입 모델

○ 11.2.1 '인도 노면 안정성 모니터링 시험 서비스' 프로토타입 모델

'인도 노면 안전성 모니터링 시험서비스'는 본 과제에서 생성한 데이터셋을 활용하여, 학습한 모델, 네트워크를 통하여, 인도 노면 안전성에 대한 판단을 하고 지도에 매핑하여, AI에 기반한 '인도 노면 안전성' 지도를 만들어가고, 모니터링할 수 있는 시범 서비스 프로토타입이다.

유저는 서비스 모바일 앱 또는 서비스 웹사이트를 통하여, 현장에서 발생한 인도 노면의 영상 또는 사진을 등록할 수 있으며, 클라우드 플랫폼에 원천데이터(영상, 사진)가 도달하면, AI 서비스 플랫폼이 분석 후, 해당 결과를 지도에 필요한 방식으로 표현한다.

유저들이 참여하여, 인도의 안전성을 AI 서비스 플랫폼이 분석 표현하며, 보다 안전한 길을 쉽게 확인할 수 있고, 많은 참여가 이뤄질 경우, 보다 정밀하고 자주 업데이트 되는 안전한 길 안내 서비스를 목표할 수 있다.

○ 11.2.2 '인도 노면 안전성 모니터링 시험 서비스' 프로토타입 모델 구현 방안

	세부 내용
서비스 모델 설계 및 개발	• 모바일 앱, 웹 서비스 기반으로서, 누구나 손쉽게 참여가 가능하도록 서비스 모델을 설계. • 유저가 시작 단계에서 불편함없이 바로 현장 영상, 사진을 쉽게 등록 가능하도록 하고, 유저에게 부담을 주지 않는 설계. • 웹표준 기반으로 개발하여, 확장 가능성을 높인다.
원천 데이터 확보 및 활용방안	• 크라우드 소싱 방식으로 학습용 데이터를 확보 • 확보된 학습용 데이터를 통해, 목표 탐지 대상에 대해서, 딥러닝을 통한 학습
서비스 모델 구현 방안	• 모바일 앱: 유저가 손쉽게 현장 영상, 사진을 클라우드 AI 서비스 플랫폼에 등록할 수 있는 모바일 어플리케이션 • 웹 서비스: 웹표준 서비스로서, 별도의 설치 플러그인 없이 유저가 손쉽게 현장 영상, 사진 등록 • 지도 기반 현장 확인: 원천데이터(현장 영상, 사진)으로부터, 위치 정보 파악과 AI에 의해 탐지된 현장 사진을 지도에 매핑하여 노출하여, 육안으로 현장의 현황을 파악할 수 있도록 한다. • 주요 탐지 대상 - 점자블록 존재 유무 - 점자블록 파손 유무 - 보도 균열, 함몰, 파손 탐지

○ 11.2.3 '인도 노면 안전성 모니터링 시험 서비스' 프로토타입 모델 산출물

모바일 앱, 웹서비스를 통하여, 현장 영상/사진을 등록하면, AI 서비스 플랫폼이 분석 및 판단 후, 지도에 현장 사진을 매핑 표현하며, 지도에 기반한 확인이 가능.

- **시범 서비스: 인도 노면 안전성 모니터링 시험 서비스(웹서비스) 1식**
- **모바일 앱: 안드로이드 모바일 앱 1식**

인도 보행 영상 녹화

AI 모델링 프로세스 수행

인도 안정성 모니터링 시스템

- 점자블록 존재 유무 탐지
- 점자블록 파손 유무 탐지
- 보도 균열, 함몰, 파손 탐지

실제 적용 화면

부록 1

학습용 데이터 제작 도구(AAT : aiWorks Annotation Tool)

학습용 데이터 제작 도구(이하 AAT)는 MIT License로 배포된 데이터 가공 도구를 커스터마이징한 어노테이션 도구이다.
사업 종료 후, 도구를 공개해 다양한 데이터 가공 작업에 활용될 수 있도록 한글화 작업 및 영상 데이터에 대한 가공 기능 등 추가 구현하였다.

[부록 Ⅰ – 그림 1] 학습용 데이터 제작 도구 실행 화면

● 주요 도구 기능

가. 4가지 형식의 어노테이션

- 이미지를 박스(Box), 다각형(Polygon), 복선(Polyline), 점(Point) 등의 4가지 형식으로 분류하여 정의된 객체클래스에 따라 어노테이션 가능함

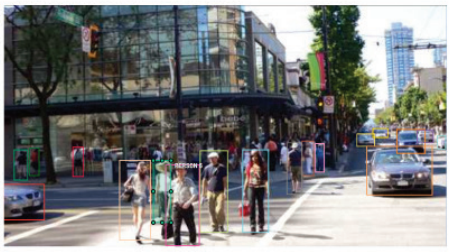
[부록 Ⅰ – 그림 2)] 주요 도구 기능

박스(Box) 다각형(Polygon) 복선 (Polyline) 점(Point)

나. 자동화 기능(Automation)

- 원천 데이터에서 사람이나 자동차와 같이 반복적으로 검출되는 객체들은 미리 학습된 인공지능 모델을 기반으로 자동 검출(detection)하여 어노테이션함

[부록 Ⅰ – 그림 3] **학습용 데이터 제작 도구 작업 단계**

❶ 	❷ 
AAT에 Task 업로드	자동 Annotation 실행
❸ 	❹ 
자동 Annotation을 통해 데이터 품질 80% 확보	보정·검수를 통해 데이터 품질 99% 확보

다. 동영상 가공 기능

- 객체 움직임 시작 시점과 종료 시점(변곡점)을 어노테이션하면 시간과 거리를 자동 연산하여 프레임별로 객체를 추적하여 해당 객체를 어노테이션함

부록 2

참고 문헌

[1] YOLO, Joseph Redmon, Santosh Divvala, Ross Girshick, You Only Look Once: Unified, Real-Time Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[2] MIT lincense,
https://ko.wikipedia.org/wiki/MIT_%ED%97%88%EA%B0%80%EC%84%9C

[3] PSMNet, Chang, Jia-Ren and Chen, Yong-Sheng, Pyramid Stereo Matching, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[4] GANet, Zhang, Feihu and Prisacariu, Victor and Yang, Ruigang and Torr, Philip HS, GA-Net: Guided Aggregation Net for End-to-end Stereo Matching, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019.

[5] Faster RCNN, Shaoqing Ren, Kaiming He Ross Girshick Jian Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, Conference on Neural Information Processing Systems, 2015

※ 상기 기재된 참고 문헌의 발간일이 기재된 경우, 해당 표준(문서)의 해당 버전에 대해서만 유효하며, 연도를 표시하지 않은 경우에는 해당 표준(권고)의 최신 버전을 따름

부록 3

가이드라인의 예시 및 추가 설명

[부록 7.1] 원천데이터 수집에 이용 된 앱 내 사진 수집 메뉴얼

SelectStar Inc.

인도(人道)에서 사진 찍기 설명서

※ 190720 19:06 조건2 주의사항 추가 :
사진이 조금이라도 흔들리면 불통과 입니다.
특히 밤에 주의해주세요!

※ 190815 15:05 조건1 주의사항 추가 :
1. 물체는 1~10개인 사진이 적당합니다.
2. 인도와 평행하지 않은사진,
빛 번짐이 심한사진,
카메라 노후로 인해 흐린 사진 불통과 입니다.
예시를 주의해서 살펴봐 주세요.

본 미션은 사람이 걸어다닐 수 있는 길인 '인도'에서 사진을 찍는 미션입니다.

시각장애인의 원활한 인도보행을 위해 활용될 인도보행영상 인공지능을 만들고자 "인도"에서의 사진을 수집하려합니다.

이 미션의 목적은
일반 인도, 골목길, 공원, 횡단보도 등 사람이 걸을 수 있는 실외 모든 곳에서
"인도 및 생활도로가 포함되도록" 주변을 찍는 것입니다.

조건1. - 조건2. - 조건3. - 부록.

조건1. 어떤 것이 인도사진을 포함할까?

제목을 클릭하면 상세 설명서로 이동합니다

1. 실외의 인도사진을 찍습니다. 또한 인도에서서 촬영된 사진이어야 합니다.
2. 50m이상 길이가 보이는 인도 사진을 찍습니다.
3. 인도지만 물체가 과다하게 포함되어 있거나 너무 적은 경우는 제외합니다.
   - 강가, 들판, 풀밭을 포함한 옆의 산책로의 경우
   - 공원, 주차장, 광터, 놀이터의 경우
   - 물체가 0개면 절대 안됨,(190815 15:05 추가)
   - 10개 이하가 적당(190815 15:05 추가)
4. (190815 15:05 추가) 인도와 평행하지 않은사진, 빛 번짐 심한 사진, 카메라 노후로 인한 흐린 사진 모두 불통과 입니다. 특히 주의해 주세요

조건1. - 조건2. - 조건3. - 부록.

조건2. 카메라 세팅은 어떻게 되어있어야 할까?

제목을 클릭하면 상세 설명서로 이동합니다

. 사진의 방향이 세로이거나, 뒤집혀 있으면 안됩니다.
. 카메라가 정면을 바라보고 있어야 합니다.
. 사진의 해상도가 19:6 비율(1920X1080)이 아니거나, 너무 어두워 물체를 식별할수 없거나, 일부가 손으로 가려져 있으면 안됩니다.

조건1. - 조건2. - 조건3. - 부록.

조건3. 어떻게 통과 사진을 다양하게 찍을 수 있을까?

제목을 클릭하면 상세 설명서로 이동합니다

꼭 멈춰서서 초점이 맞은 상태에서 촬영해주세요

1. 한개의 길을 계속해서 찍는 경우 최소 30m 간격으로 찍으면 됩니다!
2. 시간대, 날씨 등 주변 환경은 상관 없습니다!
3. 같은 자리에서 사람이 걸을 수 있는 곳을 여러 방향으로 찍을 수 있습니다!

조건1. - 조건2. - 조건3. - 부록.

부록. 왜 '인도'사진을 수집하나요?

인도 보행 사진을 수집하는 목적은
인도의 사진을 수집하여 그 길의 어떤 물체가 있는지 파악하는 AI(인공지능)를 제작하는데에 도움을 주기 위함입니다.

인도에서의 많은 위험이 존재하는 시각장애인들에게 소리를 이용하여 길의 모양이나 특징들을 알려줄 수 있다면 어떨까하는 질문에서 시작된 프로젝트입니다.

또한 밤이거나 다양한 날씨 등 많은 상황을 알 수 있게 될수록 더 많은 상황을 인식할 수 있다는 뜻이고 여러 정보를 줄 수 있게 될 것입니다.

인식하려하는 것의 종류는 다음과 같습니다.

움직일 수 있는 물체 : • 사람 (person) • 버스 (Bus) • 트럭 (Truck) • 차 (Car) • 오토바이 (Motorcycle) • 이륜차 (Scooter) • 유모차 (Stroller) • 휠체어 (Wheelchair) • 개 (Dog) • 고양이 (Cat) • 안내표지판(도로의 광고 입간판, 주차금지 표지판 등) (Movable signage) • 리어카(쇼핑 카트, 손수레 등) (Carriage)
움직일 수 없는 물체: • 가로수 (Tree trunk) • 화분 (Potted plant) • 신호등 (Traffic light) • 교통 표지판 (Traffic sign) • 기둥 (Pole) • 벤치 (Bench) • 의자 (Chair) • 탁자 (Table) • 소화전 (Fire hydrant) • 주차요금정산기 (Parking meter) • 버스/택시 정류장 (Stop) • 차량 진입 금지 설치물 (Bollard) • 사람 출입 금지 설치물 (Barricade) • 점포(박스형 점포 또는 가판매, 공중전화박스 등) (Kiosk)

여러분이 수집해주신 사진 한장 한장은 인공지능이 '사람이 걷는 길에 어떤 장애요소가 있는지'를 학습하는 귀중한 자료로 사용됩니다. 사용자분들의 많은 참여 부탁드립니다.

05

멀티모달 영상 AI 데이터 구축 가이드라인

acryl

	성명	소속	직위
가이드라인 제안	오성식	㈜ 아크릴	이사
가이드라인 작성	안형석	㈜ 아크릴	책임 연구원
	홍준기	㈜ 아크릴	주임 연구원
	김진영	㈜ 아크릴	주임 연구원
	송병후	㈜ 아크릴	주임 연구원
전문가 검토 위원	성연식	동국대학교	교수
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 아크릴/ 참여: 테스트웍스

서문

1. 가이드라인의 목적

멀티모달 영상 데이터는 영상, 텍스트, 사운드 등의 다양한 모달 정보가 포함된 영상 데이터로 유니모달 데이터에 비해 다양한 정보를 포함하는 데이터이다. 실생활에 가까운 데이터로 많은 활용도를 인정받아 마케팅, 방송, 의료/노인 복지, 서비스 로봇, 감성 관련 산업과 같은 다양한 분야에서 인공지능 데이터로 활용이 가능하다. 그러나 다양한 모달 정보가 포함된 데이터는 유니모달 데이터에 비해 수집 난이도 및 비용이 높아 고품질의 AI 데이터셋을 확보하기 어렵다. 따라서 본 가이드라인은 인공지능 학습용 멀티모달 영상 데이터를 수집하기 위한 가이드라인을 제공하는 것을 목적으로 한다. 인공지능 학습용 멀티모달 영상 데이터는 크게 원천데이터(영상) 수집과 데이터 어노테이션(태깅) 두 프로세스로 진행되며 각 프로세스는 데이터 품질 향상을 위한 단계가 필수적이다. 해당 문서에서는 고품질 인공지능 학습용 멀티모달 영상 데이터 수집 프로세스에 대한 가이드라인을 제공하기 위해 작성되었다.

2. 주요 내용 요약

인공지능 학습용 멀티모달 영상 데이터 수집을 위해서는 원천데이터 수집과 데이터 어노테이션 프로세스로 진행된다. 원천데이터 수집에서는 영상 데이터 수집에 필요한 시나리오 작성, 영상 촬영 방식에 대한 가이드라인을 제시하며 고품질 원천데이터 수집을 위한 시나리오 작성 - 시나리오 검수 / 피드백 - 영상 촬영 - 영상 촬영 검수 / 피드백 프로세스를 정립한다. 데이터 어노테이션 프로세스에서는 인공지능 학습용 데이터 구성을 위하여 어노테이션 해야 할 항목을 정의하고 이에 대한 기준을 제시한다. 또 고품질 학습 데이터 수집을 위하여 데이터 어노테이션 - 데이터 교차검수 - 어노테이션 피드백 / 원천데이터 수집 피드백 프로세스에 대한 가이드라인을 제시한다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

멀티모달 영상 데이터 구축 가이드라인

Guidelines for multimodal video data construction

05

1. 적용 범위

본 문서는 멀티모달 영상 데이터 구축을 위한 절차와 가공 지침을 제시한다. **영상 데이터의 수집 단계**에서 **데이터 가공, 데이터 검수** 각 절차에 대한 설명을 제공한다. 멀티모달 영상 데이터 및 저작도구의 활용성을 높이기 위해서는 원천데이터 및 가공 데이터의 형태와 구축 절차(6장, 7장 참조)가 표준화되어야 한다. 구축된 데이터의 가치를 보장하기 위해 어노테이션 결과에 대한 검수 기준 및 방법(8장 참조)을 함께 제시하여 고품질의 AI 학습 데이터를 확보할 수 있도록 한다.

2. 인용 표준

- 해당사항 없음

3. 용어 정의

● 어노테이션

원천데이터에서 의미있는 요소들을 형식에 맞게 표시하는 것을 의미한다. 본 가이드에서는 **영상 데이터**(원천데이터)에서 **인물의 위치, 감정, 객체의 위치**(의미있는 요소)를 표시하여 JSON 포맷으로 저장하는 어노테이션 작업을 수행한다.

● JSON[1)]

JSON은 속성-값 쌍 또는 키-값 쌍으로 이루어진 데이터 오브젝트를 전달하기 위해 자연어 텍스트를 사용하는 **개방형 표준 포맷**이다. 비동기 브라우저/서버 통신을 위해, 넓게는 XML을 대체하는 주요 데이터 포맷이다.

1) https://ko.wikipedia.org/wiki/JSON

<JSON 포맷 예시>

```
{
    "이름":"가나다",
    "나이":15,
    "성별":"남",
    "주소":"대구광역시 수성구 신매동",
    "특기":["야구","미술"],
    "가족관계":{"#":2,"아버지":"가두리","어머니":"마바사"},
    "회사":"서울특별시 강남구 청담동"
}
```

4. 약어

- VRD[2)] : Visual Relationship Detection
- CMU-RAPT[3)] : Carnegie Mellon University - Rapport-Aware Peer Tutor

5. 멀티모달 영상 데이터 구축 단계

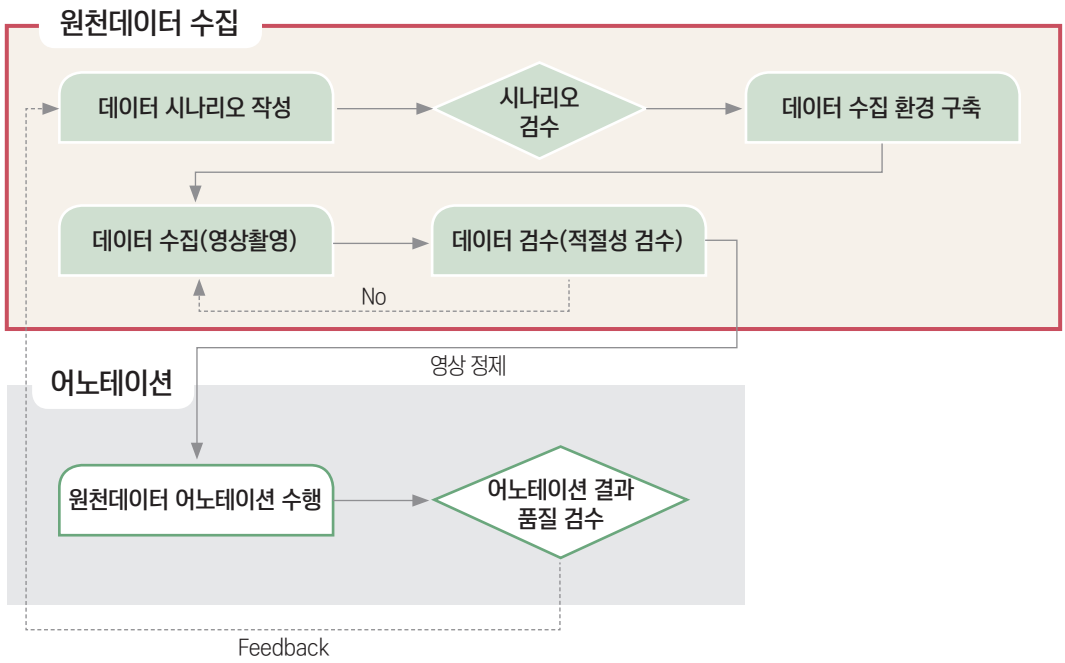
본 항목에서는 멀티모달 영상 데이터 구축 단계를 기술한다. 멀티모달 영상 데이터 구축 단계는 **원천데이터 수집 프로세스**와 **데이터 어노테이션 프로세스**로 구분된다. 원천데이터 수집 프로세스는 멀티모달 정보가 포함된 영상을 수집하는 프로세스이며 데이터 어노테이션 프로세스는 수집된 원천데이터를 기반으로 의미있는 정보를 어노테이션하는 프로세스이다.
원천데이터 수집 프로세스와 데이터 어노테이션 프로세스는 개별적인 검수과정을 거치게 되며 데이터 어노테이션 프로세스가 끝난 후 최종적으로 원천데이터 수집에 대한 피드백을 제공하여 상호보완하는 구조를 가진다. 상기 설명은 멀티모달 영상 데이터를 자체적으로 제작하는 것을 전제로 한다.

6. 원천데이터 수집 프로세스(데이터 수집)

본 항목에서는 멀티모달 정보가 포함된 영상을 제작하는 원천데이터 수집 프로세스를 기술한다. 원천데이터 수집 프로세스는 데이터 시나리오 작성, 데이터 수집 환경 구축, 데이터 수집, 데이터 검수 단계로 이루어진다.

2) Lu, C., Krishna, R., Bernstein, M., & Fei-Fei, L. (2016). Visual relationship detection using language priors. In European conference on computer vision (ECCV).
3) Madaio, M. A., Cassell, J., & Ogan, A. (2017, June). The Impact of Peer Tutors' Use of Indirect Feedback and Instructions. In Proceedings of the Twelfth International Conference of Computer-Supported Collaborative Learning, 2017.

[그림 1] 원천데이터 수집 프로세스



● 6.1 데이터 수집 고려사항

멀티모달 영상 원천데이터를 구축할 때 고려해야 할 사항은 표 2와 같다. (데이터 균형에 대한 가이드라인은 6.3 참조.)

[표 1] 멀티모달 영상 원천데이터 수집 시 고려사항

고려사항	설명	비고
저작권 및 초상권	• 외부 데이터 활용 시 저작권 문제 해결이 필수적 • 인물에 대한 초상권 이용 동의 획득 필요 • 저작권 보유기관과의 협의를 통해 권리를 획득하거나, 자체 데이터 제작 권장	데이터 공개를 위한 필수 사항
데이터 균형	• 등장하는 인물의 다양성 • 개체의 다양성 • 감정 표현의 다양성 • 대화 주제 및 상황의 다양성	활용 분야별로 검토

● 6.2 데이터 시나리오 작성

데이터의 시나리오는 현대 시대를 담고 있는 일상 대화를 기본으로 하며 역사극, 특정 전문 분야에 종속적인 콘텐츠, 대화를 포함하지 않는 콘텐츠는 제외한다. 데이터의 시나리오는 집(침실, 방, 주방)과 강의실, 카페, 사무실 그리고 장소 무관을 포함하여 7가지 상황을 가정하며(촬영 장소를 특정하는 것이 아닌, 시나리오 내의 상황을 가정) 각 상황에 적합한 소품이 시나리오에 포함된다. 데이터 시나리오는 연기자 연령, 상황, 소품의 등장 빈도를 균등하게 맞추도록 작성한다.

적합한 소품이란 시나리오 배경 장소에 어울리며 영상을 통해 소품을 인식하기에 보편적인 형태를 지닌 소품으로 정의한다(적합한 소품 예시: 카페-커피잔, 부적합한 소품 예시: 독특한 디자인으로 인해 커피잔으로 인식하기 어려운 컵).
욕설이나 비속어는 적절한 수위를 유지하며, 은어, 관용적 표현, 일반적 사투리는 허용되며 특정 브랜드의 언급은 불가하다. 상기 언급한 적합한 소품 및 연기자의 인적사항과 성별을 포함하여 시나리오를 작성한다. 단편적인 발화만 주어질 경우 감정 이입이 어렵기 때문에, 상황을 함께 제시하여 그에 맞는 발화를 수행하고 반복되는 문장 및 표현을 지양한다.

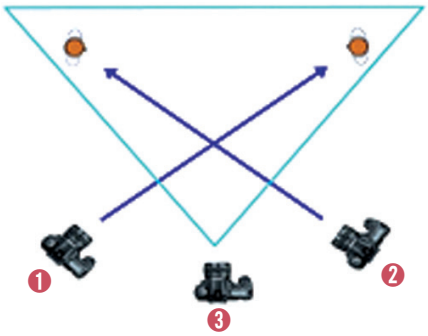
● 6.3 원천데이터 수집 환경 구축

데이터의 수집 환경의 구성 내용은 연기자의 성별, 연령대, 시나리오 상황, 소품으로 구성된다. 연기자의 성별 (남, 여) 및 연령대 (10대 이하, 10대, 20대, 30대, 40대, 50대, 60대 이상)는 원천데이터 내 출연 빈도 균등하게 배분하여 모집한다. 데이터 수집은 스튜디오 등 실내 작업 공간에서 수집하는 것을 기본으로 하며 시나리오의 상황(침실, 방, 주방, 강의실, 카페, 사무실, 장소무관)에 적합한 소품(6.2 '적합한 소품' 정의 참조)을 포함한다.

● 6.4 원천데이터 수집

데이터의 수집은 카메라 3대로 구성되며, Full HD(1080p) / H.264 mp4의 영상을 확보한다. 각 카메라의 구도는 3번 카메라 정면을 기준으로, 좌우에 1, 2번 카메라를 연기자의 얼굴의 정면을 촬영할 수 있도록 배치한다(그림2 참조). 개별 영상의 길이는 2분 내외로 촬영하며 영상의 컨텐츠는 6.3장 수집 환경 구축의 내용을 따른다.
감독 및 관계자의 판단 하에 연기가 제대로 수행되지 않았을 경우 연기 적절성을 검수하여 재촬영을 진행한다. 데이터를 수집할 때, AI 학습을 위해서 일정하게 존재하는 지배적인 소품은 배제하여 수집하며 사용자가 원하는 AI 서비스에 맞추어 시나리오나 환경을 구성하여 수집한다.

[그림 2] 원천데이터 카메라 구도



● 6.5 원천데이터 검수

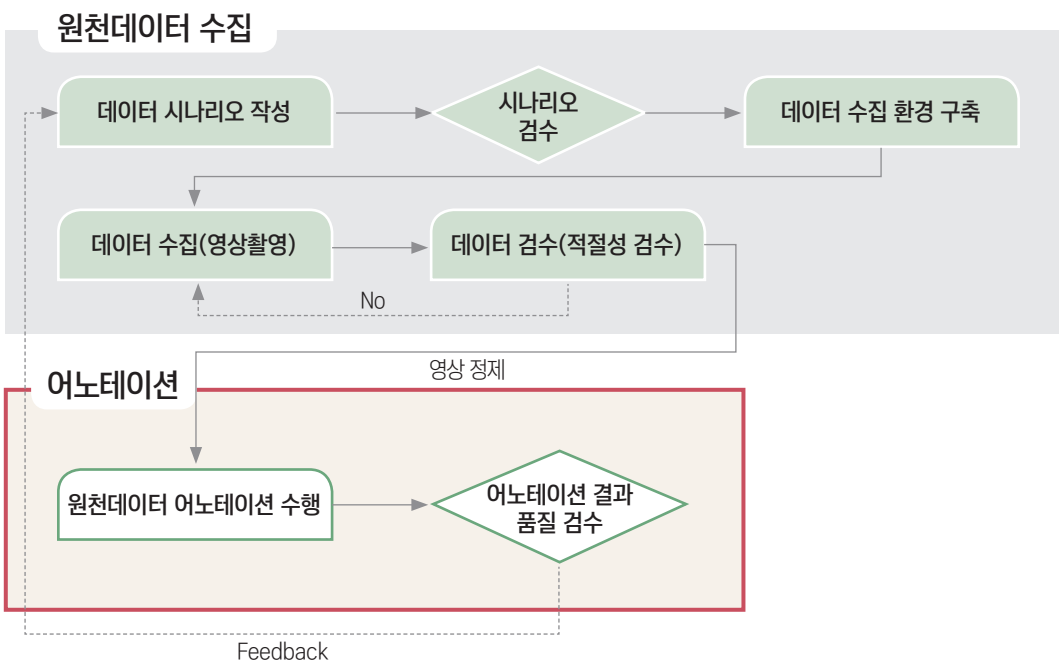
원천데이터의 정제 오류를 방지하기 위한 절차로써 정제 작업자와 다른 작업자가 정제 결과물의 검수를 진행한다. 영상의 데이터 정상 분할 여부, 노이즈 여부, 데이터 내 구축할 원천데이터의 구성을 고려하며 해당 기준에 부합하지 않으면 다시 영상 검수 작업을 시행한다. 데이터 검수는 영상의 정상 분할, 노이즈를 먼저 고려하고 사용자가 사용하는 AI 서비스에 맞추어 데이터의 구성 등을 고려한다.

[표 2] 원천데이터 검수 기준

원천데이터 검수 기준	설명	비고
영상의 정상 분할 여부	• 개별 영상에서의 시나리오 겹침이나 시나리오 잘림의 유무(대본 기준) • 개별 영상의 시작과 끝 부분에서 시나리오와 무관한 영상의 포함 유무	
노이즈 여부	• 영상의 깨짐 유무 • 영상의 번짐 유무 • 음성의 명확성	
원천데이터의 구성	• 시나리오와 무관한 소품의 등장 유무 • 연기자의 성별, 나이, 등장 소품의 균등한 출연 검수	

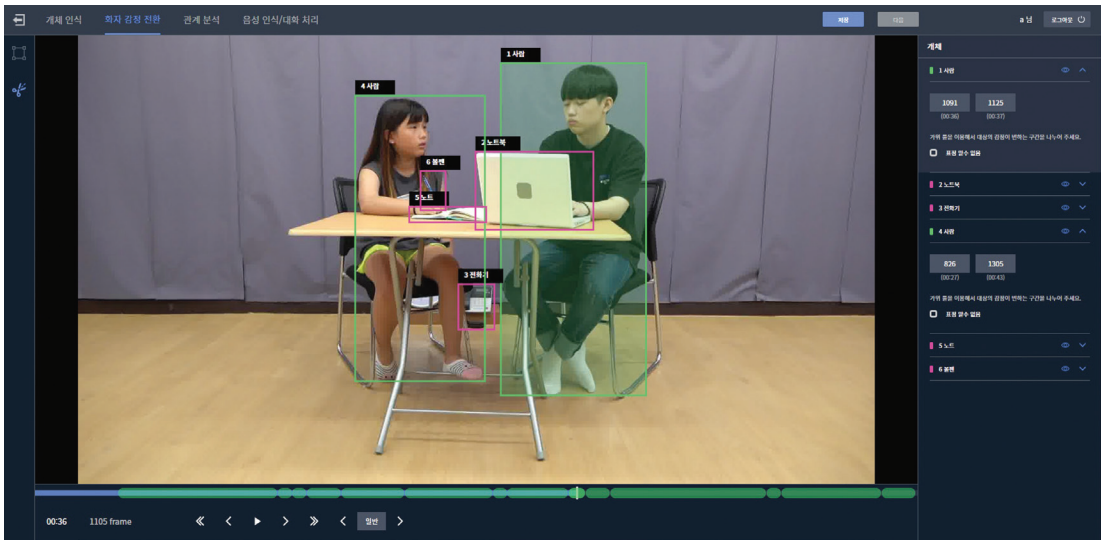
7. 데이터 어노테이션 프로세스(데이터 정제 및 가공)

[그림 2] 데이터 어노테이션 프로세스



본 항목에서는 수집된 원천데이터 영상에 의미요소를 표시하는 어노테이션 프로세스를 기술한다. 어노테이션 프로세스는 인물의 정보 및 감정 어노테이션, 개체 정보 및 관계 어노테이션, 발화 정보 및 의도 어노테이션으로 구성된다. 본 장에서는 의미요소를 정의하는 어노테이션 항목 정의, 일관된 데이터 수집을 위하여 어노테이션 기준을 정의하는 과정을 포함한다. 또한 완성된 데이터셋의 최종 형태(스키마)는 부록에서 확인 할 수 있다.

[그림 3] 어노테이션 수행 화면 예시



● 7.1 어노테이션 항목 정의

멀티모달 영상 데이터 구성을 위하여 필요한 데이터 어노테이션 항목은 표 3과 같다. 총 9개의 어노테이션 항목(인물별 감정, 인물별 성별, 인물별 연령대, 인물별 발화 스크립트, 상황 설명 정보, 발화별 대화 의도, 발화별 대화 전략 분류 정보, 개체 정보, 관계 정보)에 대해 어노테이션을 수행한다.

[표 3] 데이터 어노테이션 항목

어노테이션 대항목	데이터 항목	어노테이션 기준
인물의 정보 및 감정	인물별 감정	• 발화 텍스트, 발화 음성, 얼굴 표정, 멀티모달의 4개 모달리티에 대해 8종의 감정 기입 • 인물별 감정 8종 (수치 1~10, 기본정서단계 분류 : 기쁨, 슬픔, 분노, 놀람, 공포, 경멸, 혐오, 중립)
	인물별 성별	• 2종 이상 (남, 여)
	인물별 연령대	• 7종 (10대 이하, 10대, 20대, 30대, 40대, 50대, 60대 이상)
	인물별 발화 스크립트	• 발화스크립트

어노테이션 대항목	데이터 항목	어노테이션 기준
개체 정보 및 관계	개체 정보	• 개체 위치 • 개체 분류
	관계 정보	• 약 20종의 개체 관계(위치/행동 관계) • 개체에 대한 질문/답변
발화 정보 및 의도	상황 설명 정보	• 대화 전체의 주제를 15종 이상으로 분류
	발화별 대화 의도	• 진술/주장, 질문, 명령/요청, 약속, 표출, 응대/답변, • 인사/부르기/환기, 기타의 8종 태깅
	발화별 대화 전략 분류 정보	• CMU-RAPT의 자기대화, 질문대화, 공감대화, 칭찬대화, 비윤리적대화, 완곡대화, 비언어적대화의 7종 태깅

● 7.2 어노테이션 기준 정의

○ 7.2.1 인물의 정보 및 감정

• 인물 감정

멀티모달 감정 분석을 위해서는 모달별 감정 정보와 복합 감정 정보가 함께 주어져야 하며, 따라서 데이터 구성 시 텍스트, 음성, 이미지 및 멀티모달의 총 4가지 감정값을 기입한다.

[표 4] 감정값 기입 항목

모달리티	어노테이션 기준
텍스트 발화문	대화문을 기준으로 느껴지는 감정을 기입 ※ 음성 및 표정 정보를 고려하지 않음
발화 음성	발화 음성의 빠르기와 높낮이 등 음성 자체에서 느껴지는 감정을 기입 ※ 표정 정보를 고려하지 않음
얼굴표정(이미지)	발화자의 표정을 기준으로 느껴지는 감정을 기입 ※ 텍스트 및 음성 정보를 고려하지 않음
멀티모달	발화자로부터 느껴지는 종합적인 감정을 기입

위와 같은 방식으로 기입된 4개의 감정 정보는 서로 다를 수 있다.
예시) 아이고 이를 어쩐다 (텍스트: 공포, 멀티모달: 기쁨)
정말 잘했다 (텍스트: 기쁨, 멀티모달: 혐오)

멀티모달 감정 분류 (텍스트 발화문)는 하기 표 5와 같으며, 8종 감성을 따른다.

[표 5] 멀티모달 감정 분류 (텍스트 발화문)

감정 이름	예시 문장	비고
기쁨(happy)	오늘 정말 기분이 좋아	
놀람(surprise)	믿을 수가 없어!	
분노(angry)	여기는 동네 단골을 이런 식으로 대하나?	
슬픔(sad)	나 진짜 너무 속상해 엄마.	
혐오(dislike)	아이고, 잘도 그러겠다	
공포(fear)	이러다가 영영 안되는거 아닐까 걱정되어요	
경멸(contempt)	너라는 사람 정말 지긋지긋하다	
중립(neutral)	이거 한번 볼래?	

- 인물 정보

[표 6] 인물정보 어노테이션 기준

항목	어노테이션 기준
인물 위치	• 인물 전신이 가급적이면 완전히 포함될 수 있도록 bounding box 설정 • 인물 끝부분(전체 면적의 5% 이내)을 포함하기 위해 지나치게 영역을 크게 잡아야 하는 경우, 영역을 축소 설정하는 것을 허용
인물 정보	• 인물 ID • 나이, 성별 (가급적 실제 출연 인물의 정보를 바탕으로 작성)

7.2.2 개체 정보 및 관계

- 개체 정보

개체 리스트(예시)는 다음과 같다.

집(침실)						
액자	캔들	화장품	화장도구	거울	베개	각티슈
화분	시계	쿠션	휴지통	디퓨져	스피커	
집(방)						
액자	캔들	거울	책	신문	멀티탭	쿠션
쟁반	볼펜	롤휴지	화분	스피커	시계	휴지통
먼지떨이	디퓨저	화분	리모컨	바구니	우산	쓰레받기
빗자루	각티슈	노트				
집(주방)						
쟁반	각티슈	화분	시계	휴지통	디퓨저	수저
포크	접시	쟁반	롤휴지			

강의실						
리모컨	마이크	지시대	책	보드마카	보드지우개	시계
물통	음료병	커피잔	텀블러	노트	필통	클리어파일
볼펜						
카페						
커피잔	컵받침	멀티탭	물잔	물통	커피잔	음료병
화분	포크	수저	접시	쟁반	볼펜	
사무실						
멀티탭	텀블러	물잔	화분	볼펜	각티슈	키보드
마우스	모니터	연필꽂이	종이컵	책	노트	클리어파일
테이프	스테이플러	포스트잇	가위	달력	우산	전화기
커피잔	액자					

개체에 대한 어노테이션 기준은 다음과 같다.

[표 7] 개체 어노테이션 기준

항목	어노테이션 기준
개체 위치	• 개체가 가급적이면 완전히 포함될 수 있도록 bounding box 설정 • 개체 끝부분(전체 면적의 5% 이내)을 포함하기 위해 지나치게 영역을 크게 잡아야 하는 경우, 영역을 축소 설정하는 것을 허용
개체 분류	• 개체 분류 선택 • 기 정의된 분류에 없을 경우 '기타' 선택

• 개체 관계 분석

시각적 관계 탐지를 벤치마킹하기 위한 스탠포드 VRD에서는 두 개체 간의 관계를 Action, Spatial, Preposition, Comparative, Verb로 구분하고 있다. 본 가이드라인에서는 스탠포드 VRD의 기준을 일부 차용하되, 일상대화 장면에서 등장할 수 있는 개체 관계를 위주로 어노테이션 항목을 정의하였다.

[표 8] 위치관계, 행동관계 어노테이션 항목 리스트

분류	관계	비고
위치관계	~ 위에 있다	
	~ 아래에 있다	
	~ 앞에 있다	
	~ 뒤에 있다	
	가까이 있다	상기 위치관계가 명확하게 대응되지 않는 경우

분류	관계	비고
	닿아 있다	
	~ 안에 있다	
행동관계	잡다	
	들다	
	밀다 / 당기다	
	기대다	
	사용하다 / 착용하다	
	먹다 / 마시다	
	앉다	
	눕다	
	따라가다	
	때리다	

○ 7.2.3 발화 정보 및 의도

• 의도 분류

진술/주장, 질문, 명령/요청, 약속, 표출, 응대/답변, 인사/부르기/환기, 기타의 8종 중에서 문장의 형식이나 의도에 해당되는 것을 선택한다.

[표 9] 발화 의도 분류

	설명	예시문장
진술/주장	과거에 있었던 일을 서술 /S가 O를 해야한다고 주장	이렇게 해서는 장사가 안됩니다
질문	문장부호에 상관없이 발화자의 궁금함이 드러난다면 질문으로 분류	거기까지 가는데 얼마나 걸리는데?
명령/요청	발화자가 아닌 다른 사람이 무언가를 해주기를 바라는 것	오늘 청소 좀 해줬으면 좋겠는데
약속	발화자가 무언가를 하겠다는 의지, 다짐의 표현	이건 내가 내일까지 해올게
표출	발화자의 감정이 드러나는 표현	아이고 억울해 죽겠어
응대/답변	문장을 보고, 앞선 상대방의 발화가 추측되는 경우	내가 안했다니까 그러네
인사/부르기 /환기	인사, 타인을 부르기, 환기 (대화 소재를 바꾸는 형태의 표현)	그건 그렇고, 어제는 무슨 일 있었어?
기타	상술한 분류에 해당하지 않는 것	

각각의 발화문은 문장의 표면적 표현 형식과 내재적 의도가 상이할 수 있다. 표면적 형식은 한 가지를 선택하며 내재적 의도는 복수 선택이 가능하다.

예시) 창문좀 닫아줄래? (표면적 형식은 질문이지만 내재적 의도는 명령/요청)

내가 한 일이 아니라니까! (표면적 형식은 응대/답변, 의도는 응대/답변 및 표출)

• 대화 전략

CMU-RAPT의 7가지 대화 전략을 기준으로 한다.

[표 10] 대화 전략 분류

	설명	예시문장
자기대화	본인의 정보나 측면을 드러내는 회화적 행위	나는 애완동물을 좋아해
질문대화	상대방으로부터 자기대화 응답을 이끌어내는 행위	그래서, 너는 어디에 사니?
공감대화	대화에 있는 두 명의 사람이 동시에 경험한 일이나 공유된 관심사에 대한 언급	ㅇㅇㅇ가 지난주에 페이스북에 업로드한거 대단했지 너도 xbox 게임을 좋아하니
칭찬대화	칭찬하는 행위	이번에 정말 잘했어
비윤리적대화	일반적, 사회적으로 받아들여지는 규범에 반하는 행동. 두 화자가 더 가까워지고 있다는 신호이거나, 더 이상 커뮤니티 규범을 지키려고 하지 않는 상황	(상대방에 대한 비난, 놀림, 모욕 등) 봄에 남자친구도 안만들고 뭐했어 기분 잡치네 진짜
완곡대화	간접적으로 돌려서 말하는 행위로, 모호한 단어로 돌려 말하거나 단어의 강도를 줄이는 어구를 추가하는 등의 표현	어쩌면 그게 아닐수도 있어 내 생각엔 다른 방법을 찾아보는게 좋을 것 같은데
비언어적대화	비언어적 의미 전달로 상대방에게 피드백을 제공하는 것을 제외하고는 자체적 의미 없음	머리 끄덕임 또는 추임새, 웃음 등

각각의 발화문은 두 가지 이상의 대화 전략을 가질 수 있으므로 복수 선택한다.

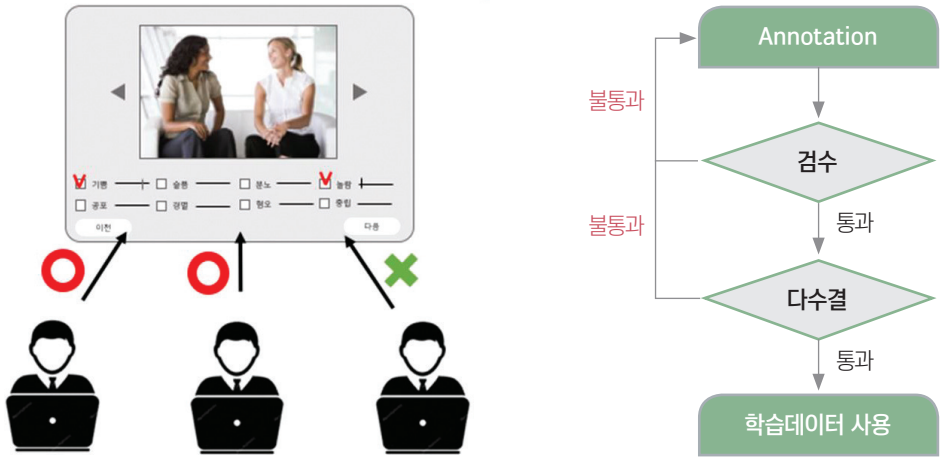
예시) 혹시 다른 방식은 없을까? (질문대화, 완곡대화)

나는 그거 신청하려는데 너는 어때? (자기대화, 질문대화)

비언어적 대화의 경우 발화문에서 선택하지 않는다.

8. 어노테이션 결과 검증(데이터 검증)

[그림 4] 어노테이션 결과 검증 절차



어노테이션 결과에 대한 검증은 양질의 데이터를 얻기 위해 필요한 작업이며 특정 개인의 의사에 치우치지 않아야 한다. 이에 제 3자 결정 등을 통한 일반화 작업을 수행한다. 웹 시스템을 이용하여 어노테이션한 결과를 3자 다수결을 통해 판단하며, 3자 판단 기법으로 판단이 힘들 경우에는 집단 지성을 통한 판단으로 결정한다. 또한 수집하고자 하는 원천데이터의 어노테이션 결과가 8종 Task를 포함하도록 균등화되어야 한다. 검수자의 Feedback을 통하여 데이터의 균등화와 고도화를 도모한다. 실제 어노테이션 품질 검수는 사용자가 필요한 AI 서비스에 맞추어 어노테이션 결과의 적합성과 학습 가능성 여부를 판단한다.

[표 11] 어노테이션 결과 피드백 기준

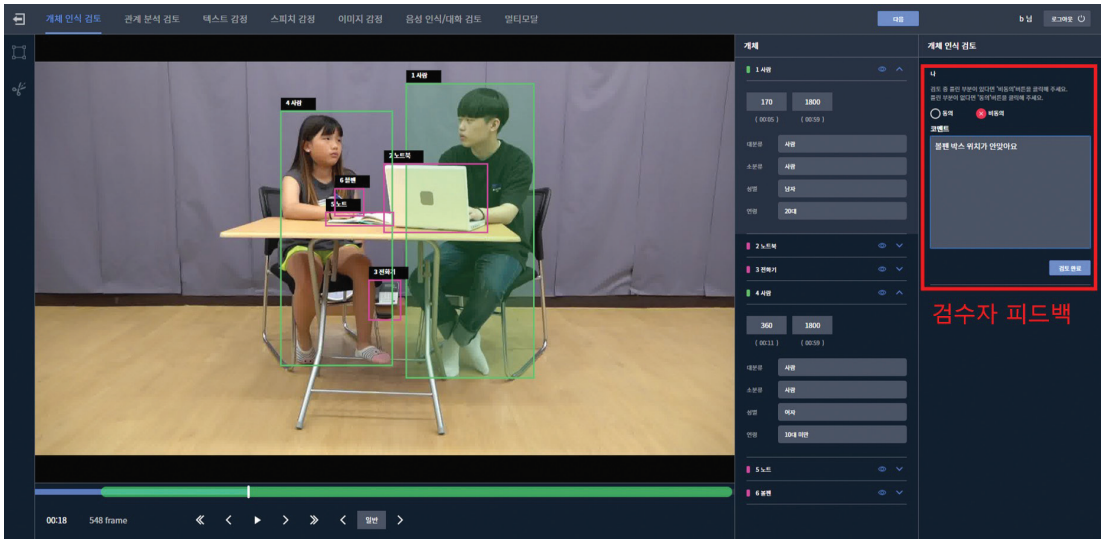
어노테이션 대항목	피드백 목표	비고
인물의 정보 및 감정	• 인물의 텍스트, 이미지, 사운드, 멀티모달 감정 어노테이션 결과의 적절성 여부 • 인물 감정의 치우침 여부(감정 종류의 균등화) • 인물의 고유 ID, 위치, 형상에 맞는 Bounding Box 어노테이션 여부 • 인물 정보의 치우침 여부(나이, 성별의 균등화)	검수자 주관에 맡기며 2인 이상 동의 시 적절함으로 판정
개체 정보 및 관계	• 개체 위치, 형상에 맞는 Bounding Box 어노테이션 여부 • 개체 분류 어노테이션 결과의 적절성 여부 • 개체 분류의 치우침 여부	상기와 동일
발화 정보 및 의도	• 상황정보, 발화별 대화의도, 발화별 대화전략 어노테이션 결과 적절성 여부 • 발화 정보의 치우침 여부(대화의도, 대화전략의 균등화)	상기와 동일

각 어노테이션 데이터 항목에 대한 세부 판정 기준(Ground Truth)은 아래와 같다.

[표 12] 어노테이션 데이터 항목별 세부 판정 기준

어노테이션 대항목	데이터 항목	어노테이션	판정 기준
인물의 정보 및 감정	인물별 감정	• 4개의 모달리티에 대한 8종의 감정 기입	• 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단
	인물 정보	• 인물 위치 • ID, 나이 및 성별	• 인물 위치는 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효 • 나이 및 성별은 등장인물 메타 정보 활용
	인물별 발화 스크립트	• 인물의 발화 내용	• 시나리오 활용을 통해 스크립트 자동 구성
개체 정보 및 관계	개체 정보	• 개체 위치 • 개체 분류	• 개체 리스트에 기반한 개체 위치정보 태깅 • 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단
	관계 정보	• 약 20종의 개체 관계 (위치/행동 관계)	• 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단
발화 정보 및 의도	상황 설명 정보	• 대화 전체의 주제에 대한 15종 기입	• 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단
	발화별 대화 의도	• 진술/주장, 질문, 명령/요청, 약속, 표출, 응대/답변, 인사/부르기/환기, 기타의 8종 태깅	• 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단
	발화별 대화 전략 분류 정보	• CMU RAPT 기반의 자기대화, 질문대화, 공감대화, 칭찬대화, 비윤리적대화, 완곡대화, 비언어적대화의 7종 태깅	• 3인 이상의 검수자 중 과반수 이상이 동일하게 평가하면 유효하다고 판단

[그림 5] 검수자 피드백 화면 예시



부록 I

가이드라인의 예시 및 추가 설명

● 데이터 스키마

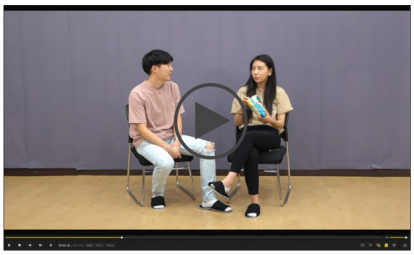
[그림 1] 데이터 구조도

● 구축 데이터 활용안

[표 1] 구축 데이터 활용 AI 서비스 유형

구축 데이터	AI 서비스 유형	비고
• 모달별 감정 정보 (표정, 음성, 대화내용) • 멀티모달 감정 정보	감정 인식	
• 개체별 위치 정보 (bbox) • 개체 분류	개체 인식	
• 인물별 위치 정보 (bbox) • 인물 정보 (나이, 성별)	성별/나이 인식	
• 관계, 능동 대상, 수동 대상 • 개체별 위치 정보 (bbox) • 개체 분류	관계 분석	
• 질문, 답변	멀티 모달 영상 질의응답	
• 발화 의도 정보	단일 발화 의도 분석	
• 인물별 음성 정보, 발화 내용	음성인식	
• 영상 별 상황 정보	복수 발화 의도 분석	

[그림 2] 전체 데이터 스키마 및 최종 JSON 데이터 예시

	예시	데이터 항목
영상		멀티 모달 감정 상황 정보
이미지		인물 성별 및 연령 인물 및 객체 영역 객체 종류 위치 관계 정보 행동 관계 정보 이미지 감정 이미지 기반 질문/답변
텍스트	@ID : 3008 @등장인물 : @등장인물151 - 20대 남자, 연인; @등장인물152 - 20대 여자, 연인; @상황 : 연인이 벤치에 앉아 과자를 먹고 있다. @151 : 먹는 거 너무 귀엽다. @152 : 아 뭐야. @151 : 그 뭔 소리야 좀 그만해. 뭘 맨날 뭔 소리야. ... @152 : 진짜 짜증난다 이거. @151 : 그치? 그럼 그만하고 항복해. @152 : (@상황; 웃음) 알았어. 그만할게 항복.	텍스트 감정 발화 의도 대화 전략 발화 스크립트 문장 품사 상황 분류
음성	00:00 / 01:23	음성 감정

JSON 형식


```
"actor": {
    "151": {
        "age": "20s",
        "gender": "male"
    },
    "152": {
        "age": "20s",
        "gender": "female"
    }
},
"situation": "연인이 벤치에 앉아 과자를 먹고 있다.",
"category": "relationship",
"data": {
    "524": {
        "1": {
            "label": "person",
            "frame": "524",
            "xtl": "954.13",
            "ytl": "210.99",
            "xbr": "1370.44",
            "ybr": "1029.68",
            "position": { "2": "close"},
            "predicate": { "1": "sit"},
            "person_id": "152",
            "text": {
                "script": "아 뭐야. " ,
                "script_start" : 524,
                "script_end" : 597,
                "morpheme" : [["아","IC"],
                              ["뭐","NP"],
                              ["야","JKV"],
                              [".","SF"]],
                "intent": "question",
                "strategy": "QESD"
            },
            "emotion": {
                "multimodal": {
                    "emotion": "dislike",
                    "arousal": 2,
                    "valence": 5,
                    "emotion_sub": "fear"
                },
                "image": {
                    "emotion": "contempt",
                    "arousal": 6,
                    "valence": 3,
                    "emotion_sub": "dislike"
                },
                "sound": {
                    "emotion": "dislike",
                    "arousal": 4,
                    "valence": 4,
                    "emotion_sub": "contempt"
                },
                "text": {
                    "emotion": "dislike",
                    "arousal": 4,
                    "valence": 5,
                    "emotion_sub": "neutral"
                }
            }
        },
        "QA": { "사람이 물건을 들고 있나요?" : "네", }
                    ⋮
                    ⋮
    }
}
```


*위 JSON 구조는 전체 JSON 구조의 일부 예시임

[그림 3] 최종 JSON 데이터 설명


```
{
        "clip_id": "300873",                    영상 ID
        "actor": {                              인물 정보
                "151": {                        인물 ID
                        "age": "20s",                   해당 인물 연령
                        "gender": "male"                해당 인물 성별
                },
                "152": {
                        "age": "20s",
                        "gender": "female"
                }
        },
        "nr_frame": "1830",                     영상 전체 프레임 수
        "video_size": {                         영상 해상도
                "width": "1920",                프레임 너비
                "height": "1080"                프레임 높이
        },
        "situation": "연인이 벤치에 앉아 과자를 먹고 있다.",      상황 정보
        "category": "relationship",             상황 분류
        "data": {
                "524": {                        #번째 프레임
                        "1": {                  객체 ID
                                "label": "person",              객체 종류
                                "frame": "524",                 #번째 프레임
                                "xtl": "954.13",                객체 위치   객체 좌측 끝 좌표
                                "ytl": "210.99",                            객체 상단 끝 좌표
                                "xbr": "1370.44",                           객체 우측 끝 좌표
                                "ybr": "1029.68",                           객체 하단 끝 좌표
                                "position": {                   위치 관계 정보
                                        "2": "close"            관계가 있는 객체 ID : 위치 관계
                                },
                                "predicate": {                  행동 관계 정보
                                        "1": "sit"              관계가 있는 객체 ID : 행동 관계
                                },
                                "person_id": "152",             인물 ID
                                "text": {                       발화 정보
                                        "script": "아 뭐야. ",          발화 스크립트
                                        "script_start" : 524,           발화 시작 프레임
                                        "script_end" : 597,             발화 종료 프레임
                                        "morpheme" : [                  문장 품사
                                                [
                                                        "아",           단어
                                                        "IC"            품사
                                                ],
                                                [
                                                        "뭐",
                                                        "NP"
                                                ],
                                                [
                                                        "야",
                                                        "JKV"
                                                ],
                                                [
                                                        ".",
                                                        "SF"
                                                ]
                                ],
                                        "intent": "question",   발화 의도
                                        "strategy": "QESD"      대화 전략
                                },
                                "emotion": {                    감정
                                        "multimodal": {         멀티 모달 감정 정보
                                                "emotion": "dislike",   멀티 모달 감정
                                                "arousal": 2,           멀티 모달 감정 세기
                                                "valence": 5,           멀티 모달 감정 긍부정도
                                                "emotion_sub": "fear"   멀티 모달 부수 감정
                                        },
                                        "image": {              이미지 감정 정보
                                                "emotion": "contempt",  이미지 감정
                                                "arousal": 6,           이미지 감정 세기
                                                "valence": 3,           이미지 감정 긍부정도
                                                "emotion_sub": "dislike"  이미지 부수 감정
                                        },
                                        "sound": {              음성 감정 정보
                                                "emotion": "dislike",   음성 감정
                                                "arousal": 4,           음성 감정 세기
                                                "valence": 4,           음성 감정 긍부정도
                                                "emotion_sub": "contempt"  음성 부수 감정
                                        },
                                        "text": {               텍스트 감정 정보
                                                "emotion": "dislike",   텍스트 감정
                                                "arousal": 4,           텍스트 감정 세기
                                                "valence": 5,           텍스트 감정 긍부정도
                                                "emotion_sub": "neutral"  텍스트 부수 감정
                                        }
                                }
                        },
                        "QA": { "사람이 물건을 들고 있나요?" : "네", }   질문 & 답
                }
        }
}
```

06

사람동작 영상 AI 데이터 구축 가이드라인

We are always Sweet *

	성명	소속	직위
가이드라인 제안	이준호	스위트케이	연구소장
가이드라인 작성	천제민	스위트케이	이사
	김영광	스위트케이	전임연구원
	박영숙	스위트케이	과장
	권혁민	스위트케이	연구원
	문경식	서울대	연구원
	소영준	케이티	책임연구원
	양기혁	모션테크놀러지	대표
전문가 검토 위원	김도형	ETRI	박사
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 스위트케이/ 참여: 서울대학교, 케이티, 모션테크놀로지

서 문

1. 가이드라인의 목적

사람의 자세추정 연구는 컴퓨터비전(CV, Computer Vision) 분야에서 오래전부터 연구되어온 분야이며 관절 데이터는 해당 연구에서 가장 필수적이며 기초적인 데이터이다. 국내 기초 연구 및 여러 산업 분야에서 활용 가능한 범용적인 내국인 및 국내 환경에 맞는 관절 데이터가 부족한 상황이며, 학습데이터 구축 및 관리에는 많은 비용과 시간이 소요된다. 이에 사람 동작 영상의 범용적인 동작 데이터 셋을 기존 연구되고 있는 데이터 셋과 연속성이 있도록 설계하며, 연구 개발자 사이 공유 가능하도록 학습 데이터 셋을 만들어야 한다. 이 가이드라인의 목적은 사람 자세 추정 연구에 활용되는 범용적인 동작 학습 데이터 셋 구축을 위한 데이터 획득, 가공, 검수 절차의 세부 지침을 제공하기 위하여 작성되었다.

2. 주요 내용 요약

사람 동작 영상은 기초 동작 20종, 도구를 사용하는 등의 의미론적 동작 30종에 대하여 영상별 5~10초 길이, 총 20만 클립의 영상 데이터를 확보함에 있다. 학습 데이터 셋은 FHD 해상도의 동영상, 초당 1~3프레임의 이미지, json기반의 데이터 셋 목록 및 메타 정보로 구성된다. 사람 동작 영상은 보편적인 분야에서 활용될 수 있도록 촬영 단계부터 동작의 종류, 시나리오 및 촬영 각도, 조도 등을 다루고 있으며, 영상에서 얻어지는 수천만 건의 이미지에 대한 어노테이션 가공 절차를 담고 있다.
이 가이드라인은 사람의 동작 데이터 구축을 위한 데이터 획득, 가공, 검수 절차의 세부 지침을 제시한다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

사람 동작 표현을 위한 관절 메타 정보에 대하여 Human3.6M Dataset을 부분 인용하여 작성되었으며 해당 데이터는 관절인식 분야에서 가장 보편적으로 활용되는 데이터 중 하나이다. 본 가이드라인은 기존 데이터와의 연계를 위해 인용문서의 관절 속성 카테고리를 공유하며 동작 분류에 대한 속성은 실제 산업 서비스를 고려하여 새롭게 작성되었다. 인용문서의 관절정보는 총 16개 속성을 가지며 각 속성은 x좌표, y좌표, 가시정보로 구성되며 본 지침은 해당 내용을 따른다.

3.2 인용 가이드라인 또는 문서와 본 가이드라인의 비교표

[표] 가이드라인 비교표

본 가이드라인	Human3.6M	비고
1. 16개 관절 속성	1. 32개 관절 속성	수정 (32개의 관절이 있지만 움직이는 관절은 16개로 이들을 대상으로 함)
2. 동작 분류	2. 동작 분류	수정 (단순 동작 분류를 기본동작과 의미론적 동작으로 분리)
3. 차원 속성	3. 차원 속성	수정 (2차원 데이터에서 2차원, 3차원 데이터 동시)
4. 촬영 장소(실내·실외)	-	추가

사람 동작인식을 위한 동작 데이터 구축 가이드라인

Guidelines for Keypoints dataset for Human Pose Estimation

06

1. 적용 범위

이 가이드라인은 사람의 동작데이터 수집에 관한 속성과 카테고리에 관한 지침을 제시하며 데이터는 관절정보와 동작분류 정보로 구성되어 있다. 본 가이드라인에서는 이후 구축될 데이터들의 연계 및 통합을 위한 구체적인 관절 속성에 대해 제시한다. 특히 동작데이터가 개인 및 기업에서 공유되어 사용될 실질적인 형식을 보이기 위해 JSON 형태의 데이터 포맷을 사례로 제시한다. 동작 분류는 기초 동작과 의미론적 동작으로 구성되며 관절인식을 이용한 서비스를 위한 것이다. 두 분류의 의미적 모호성을 제거하기 위해 각 분류의 정의와 예시를 추가로 제시한다.

2. 인용 표준

- 해당 사항 없음.

3. 용어 정의

3.1 가시정보(visibility flag)

— 관절의 시각적 판별유무를 표현하는 정보를 말한다. 예를 들어 물체에 의하여 추출하려는 관절이 영상 밖에 있어 보이지 않을 때를 0, 영상 내부에 관절이 있으나 다른 물체 등에 의해 가려져서 보이지 않을 때를 1, 정상적으로 검출될 때를 2로 구성할 수 있다.

3.2 데이터세트(training dataset)

— 학습 세트(training dataset) : 모델을 학습시키기 위한 데이터 세트의 일부분

— 검증 세트(validation dataset) : 학습 데이터를 적용한 모델의 결과를 평가하기 위한 데이터 세트의 일부분

— 평가 세트(test dataset) : 모델을 테스트하기 위한 데이터 세트의 일부분

— 평가 세트 두 가지 필요조건으로 통계적으로 유의미한 결과를 도출할 만큼 커야 하고 데이터 세트를 전체적으로 나타내야 하고 평가 세트가 학습 세트와 같은 특징을 가지도록 선별해야 함

3.3 클린업

– 보통 광학식 카메라를 이용한 모션캡쳐 작업에 사용하는 용어로 광학식 카메라의 해상도와 마커의 크기로 인해 데이터의 그래프가 일정한 곡선으로 형태를 띄기는 하나 그래프의 선들이 작은 노이즈를 이루어 곡선을 만들게 된다. 이로 인해서 데이터가 약간의 떨림 현상을 발생할 수 있다. 이와 같은 일들을 방지 하기 위해서 그래프의 경향에 영향을 주지 않으면서 그래프를 깨끗이 다듬는 것을 "클린업"(Clean UP) 이라고 한다.

3.4 트리거

– 한글로는 "방아쇠" 정도로 번역 할 수 있는데 보통 장비를 구종하는 스위치를 가리킨다. 본 작업의 경우 물리적인 트리거 스위치가 아닌 소프트웨어로 이루어진 트리거로 구성되게 되는데, 이 경우 2개 이상의 프로그램을 동시에 구종해서 동시에 트리거를 누루기가 어렵다. 그래서 이 경우 2개의 소프트웨어 트리거를 동시에 눌러수(또는 실행) 할 수 있도록 물리적인 트리거를 만들어 작동 시키게 된다. 이와 같은 기능을 위해서 별도의 프로그램과 하드웨어의 제작이 필요하다.

3.5 타임코드

– 타임코드는 표준 표시 방식인 00:00:00:00 으로 표시되는 시:분:초: 프레임을 뜻한다. 타임코드의 기존의 아날로그 장비에서 다른 2개 이상의 장비를 기록하고 그 기록한 영상의 시점을 알기위해 영상에 시간을 기록하는데서 출발한 방식이다. 본 경우에서도 2개 의 상이한 장비를 녹화하고 기록해 서로 같은 시점을 찾기 위해 타임코드를 사용한다. 이 타임코드의 기록 시간을 보다 동 시간에 촬영한 영상을 찾아 저장 한다.

4. 약어

- **HPE** Human Pose Estimation

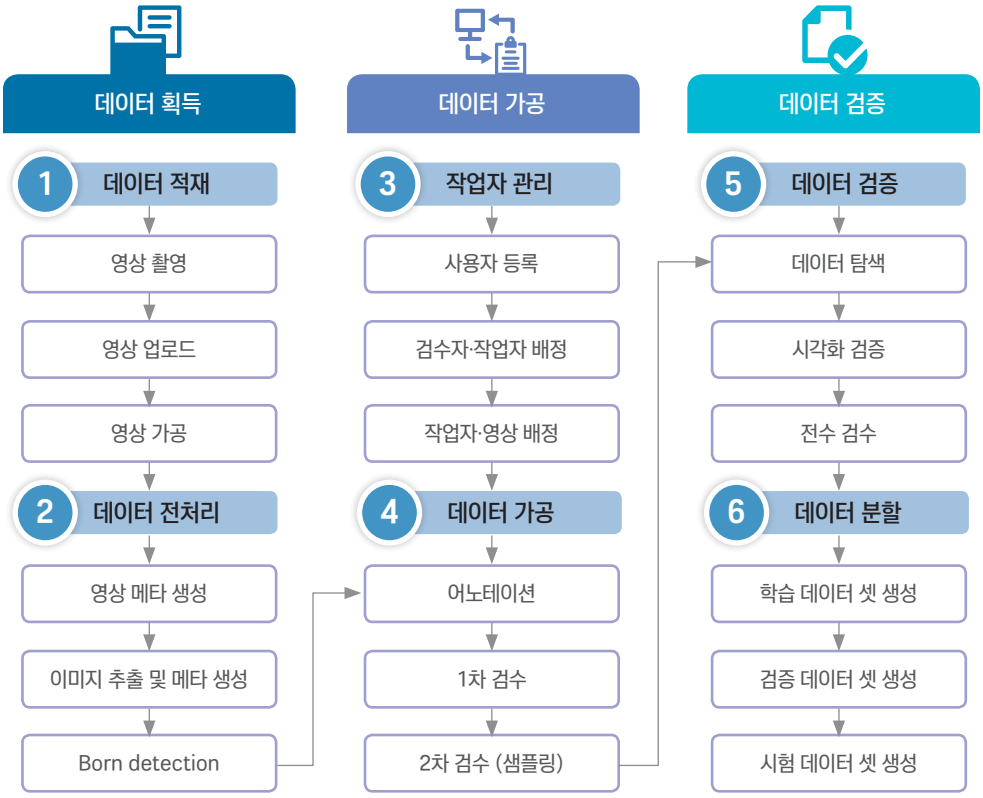
5. 사람 동작 영상 데이터 구축

사람 동작 영상 데이터는 크게 전수 촬영에 의한 데이터 획득, 영상/이미지/어노테이션의 데이터 셋을 만드는 데이터 가공, 데이터 셋의 검증 3단계로 구성됨

- 가공, 검증을 거쳐 산출된 최종결과물은 아래 표와 같음
- 이미지 수량으로 환산 시 약 450만장 dataset(평균 초당 3프레임, 평균 7.5초 동영상 20만 클립 기준)으로 비교대상인 MS COCO(20만)1)보다 20배 가량 많음.
- 동작분류와 세부수량에 대한 내용은 부속서 A참조

1) MSCOC Keypoints Dataset http://cocodataset.org/#keypoints-2019

[그림 1] 데이터 구축 프로세스



[표 1] 구축 데이터셋

영상 종류	비디오 클립 (avi)	이미지 (jpg)	2D annotation (json)
기초, 의미론적 동작 50여종 약 300 가지 시나리오	200,000	약 4,500,000	약 4,500,000 (3D 약 400,000 포함)

※ 전체 20만 클립 중 2만 클립에 대해서는 3D annotation을 포함하고 있음

- Human3.6M 데이터셋의 경우 실제 training에 적용할 경우 모든 프레임을 사용하지 않으며, 5프레임 간격으로 샘플링해서 training set으로 쓰기도하고, 50프레임 간격으로 샘플링해서 training set으로 쓰기도 하나 두 경우, 성능의 차이는 거의 없슴.(50프레임 간격으로 샘플링하면 60,000장의 training set 정도만 실제 사용됨)
- 프레임 갯수보다 데이터의 diversity가 얼마나 넓고, 의미있는 statistics를 지니는지가 중요하며 Human3.6M의 경우 15개 정도의 senario인 반면 본 데이터는 50종의 senario를 대상으로 함에 더 큰 의미가 있음

5.1 데이터 획득

- 사람 동작 영상 데이터는 수집 시 라이센스, 데이터 편향 및 데이터 정제 문제, 범용적인 활용 분야를 만족시킬 수 있는 상/중/하의 다양한 카메라 각도 등의 조건을 고려하여 기존 영상의 수집이 아닌 조건에 맞는 영상 촬영을 통하여 획득함(수집이 아니기에 별도의 정제 과정은 없음)
- 데이터 획득은 정해진 시나리오에 따라 12개의 카메라와 시간동기화를 통하여 동작구간의 영상을 편집 및 획득하며, 일부 데이터에 대하여 3D데이터의 정합성을 확인하는 가공 단계를 거쳐 이루어짐
- 영상 데이터는 데이터 가공 시스템에 등록 과정을 거치며, 최초 등록 시 영상의 분류에 따른 코드 생성, 시간 균등한 초당 3프레임의 이미지 추출, 이미지로부터 joint 추출 단계를 거침

[표 2] 데이터 획득 절차

구축단계	세부 절차	설명	필수여부
1. 데이터 획득	1.1 동작종류	기초동작과 의미론적 동작 50종 선정	필수
	1.2 동작별 촬영 시나리오	350여가지 시나리오 액터수, 의상수, 소품수, 조도 등 세부 설계	필수
	1.3 사람 동작 영상 촬영 시 고려사항	공개된 사람동작에 관한 기존 데이터셋의 한계를 검토	필수
	1.4 촬영 환경 구성	카메라위치, 공간크기, 촬영프로세스 등 실내, 실외 촬영 환경 구성 설계	필수
	1.5 카메라 사양	12대의 카메라 동시 촬영과 3D 데이터 연동을 위한 카메라 선정	필수
	1.6 카메라 구성 및 연결	12대의 카메라 동기화 및 3D 데이터 연동 구성	필수
	1.7 액터 및 촬영 계획	액터 및 촬영 계획	필수

5.1.1 동작 종류

- 동작 분류는 기초 동작과 의미론적 동작으로 구분함
- 기초 동작은 사람의 동작 가운데 가장 기초적인 동작인 쓰러짐, 싸움, 걷기, 뛰기, 앉기, 놀기 등 20종을 대상으로 함
- 의미론적 동작은 크게 일상생활, 홈피트니스, 운동, 악기 연주 등의 동작에 해당되는 종류로 선정하였음.
- 상세 수집 수량은 부속서 A 참조

[표 3] 동작종류

번호	동작 분류	항목	영상 수량
1	기초	걷기	4,128
2	기초	달리기	2,016
3	기초	앉기	2,112
4	기초	인사 (목례+절)	3,840
5	기초	인사 (손흔들기)	3,840
6	기초	인사 (악수)	1,440
7	기초	포옹	2,112
8	기초	쓰러짐	3,360
9	기초	팔짱끼기	2,304
10	기초	식사동작	2,304
11	기초	점프	2,688
12	기초	계단 오르기	3,072
13	기초	푸쉬업	2,304
14	기초	방향 지시	2,304
15	기초	박수치기	1,920
16	기초	윗몸일으키기	2,304
17	기초	포복	5,184
18	기초	태권도	2,304
19	기초	국군도수체조	1,440
20	의미론적	요가	9,600
21	의미론적	댄스 (아이돌) - 여러명	14,400
22	의미론적	댄스 (클럽 외) - 여러명	14,400
23	의미론적	댄스 (스포츠) - 여러명	14,400
24	의미론적	악기연주 (기타)	3,840
25	의미론적	악기연주 (전자피아노)	3,840
26	의미론적	악기연주 (드럼)	3,840
27	의미론적	훌라후프	2,880
28	의미론적	웨이트트레이닝	8,640
29	의미론적	홈트레이닝	5,400
30	의미론적	탁구 (서브)	1,728
31	의미론적	탁구 (리시브)	1,728

번호	동작 분류	항목	영상 수량
32	의미론적	야구 (투구)	2,304
33	의미론적	야구 (배팅)	2,304
34	의미론적	축구 (드리블)	5,616
35	의미론적	축구 (슛팅)	1,728
36	의미론적	골프 (스윙)	2,880
37	의미론적	골프 (퍼팅)	2,400
38	의미론적	배드민턴 (서브)	3,600
39	의미론적	배드민턴 (리시브)	3,600
40	의미론적	줄넘기	1,728
41	의미론적	음료마시기	1,920
42	의미론적	핸드폰 조작	1,728
43	의미론적	통화	3,072
44	의미론적	물건 운반	4,032
45	의미론적	담배 피우기	8,880
46	의미론적	사진 찍기	3,072
47	의미론적	농구	2,688
48	의미론적	역도(물건 들어올리기)	1,728
49	의미론적	펜싱(찌르기)	1,728
50	의미론적	집안일	9,216

- 5~10초 길이의 영상으로 촬영이 이루어지며, 댄스의 경우 여러 구간이 합쳐진 클립 수임
- 프로토타입 모델 개발에 따른 요건 변경 또는 추가 필요 동작 발생 시 일부 동작 종류(20% 이내)에 대하여 변동 가능
- 동작 시나리오에 따른 구체적인 수집 수량과 영상 구성은 부속서A 참조

5.1.2 동작별 촬영 시나리오

- 사람 동작 영상을 얻기 위한 50종의 동작 시나리오는 350여 가지의 상세 시나리오를 구성함
- 각 시나리오에 대하여 액터수, 의상수, 소품수, 조도, 반복촬영 등을 달리하여 촬영 함으로서 100에서 1,000 클립의 영상을 확보
- 각 동작의 상세 시나리오는 “부속서 A 동작별 촬영 시나리오”에서 확인 할 수 있음
- 동일 환경이나 형태의 반복 데이터를 피하기 위해 시나리오를 세분화하였으며 동일 시나리오에 대하여도 액터, 의상, 소품, 조도 등의 항목을 달리하여 촬영하도록 촬영 계획을 구성함 (부속서 A에서 전체 목록 확인 가능)

○ 5.1.3 사람 동작 영상 촬영 시 고려사항

[그림 2] 동작별 촬영 시나리오 예시

번호	시나리오 번호	동작 분류	항목	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수	소품	비고
1	1-1	기초	걷기	실내	1	2	2	1	2	1	12	96	맨손, 핸드폰, 책, 신문, 팜플릿, 큰 쇼핑백, 작은 쇼핑백, 백팩, 한팔백팩, 서류가방	무언가를 보며 걷는다 (ex, 신문, 팜플릿, 책 등) 추가
1	1-2	기초	걷기	실내	1	2	2	1	2	1	12	96		무언가를 들고 걷는다 (ex, 큰/작은 쇼핑백 또는 가방 등) 추가
1	1-3	기초	걷기	실내	1	2	2	1	2	1	12	96		무언가를 메고 걷는다 (백팩, 한팔 백패 메기) 추가
1	1-4	기초	걷기	실내	1	2	2	1	2	1	12	96		
1	1-5	기초	걷기	실외	1	2	2	1	2	1	12	96		1~ - 4명까지 다양하게 촬영 (같이 걷기)
1	1-6	기초	걷기	실외	1	2	2	1	2	1	12	96		
1	1-7	기초	걷기	실내	2	2	2	1	2	1	12	96	쇼핑백, 백팩, 가방	
1	1-8	기초	걷기	실내	2	2	2	1	2	1	12	96		

- 장애물로 인한 보이지 않는 영역, occlusion이 발생하는 경우와, 일정 영역에서만 촬영된 이미지로 인한 학습의 정확도 저하가 우려되는 부분있으므로, 영상 촬영시 상/하/좌/우 여러 방향에서 동시에 촬영하여 occlusion 및 일부 영역 제한의 이슈가 없도록 함. (서비스 적용시 카메라의 각도는 다양함으로 촬영영상 또한 여러 각도에서 촬영)
- 현재 존재하는 3D 인간 자세 추정 데이터 셋(Human3.6M, HumanEva, MPI-INF-HP)은 그림 55와 같이 한 명의 사람만을 포함하고, 스튜디오에서 찍힌 프레임만을 포함하고 있음

[그림 3] Human3.6M 데이터 셋 예시

- 이러한 데이터 셋들은 실생활의 어플리케이션에서 사용할 모델을 학습시키기에 한계가 있음
- 첫 번째 한계점은 한 명의 사람만을 포함하기 때문에 여러 사람이 존재하는 사진을 처리할 수 있는 모델을 개발하여 사용할 수 없다는 것임

- 실생활에서는 입력 영상에 오직 한 사람만 포함되어있는 경우보다 두 명 이상의 여러 사람이 포함되어있는 경우가 훨씬 빈번함
- 따라서 한 사람에 대해서만 모델을 학습시킬 경우 실생활에서 여러 사람이 포함 되어있는 입력 영상을 정확하게 처리하지 못하게 됨
- 두 번째 한계점은 스튜디오에서 촬영된 단조로운 배경과 복장의 연기자에 대한 프레임만 포함하고 있기 때문에 다양한 배경과 복장이 존재하는 이미지를 처리할 모델을 학습시키기에 부적절하다는 것
- 아래 그림에서 볼 수 있듯이, 기존 3D 인간 자세 추정 데이터 셋들은 단조로운 배경과 복장으로 된 프레임으로만 데이터 셋들이 이루어져 있음

[그림 4] HumanEva 데이터 셋 예시

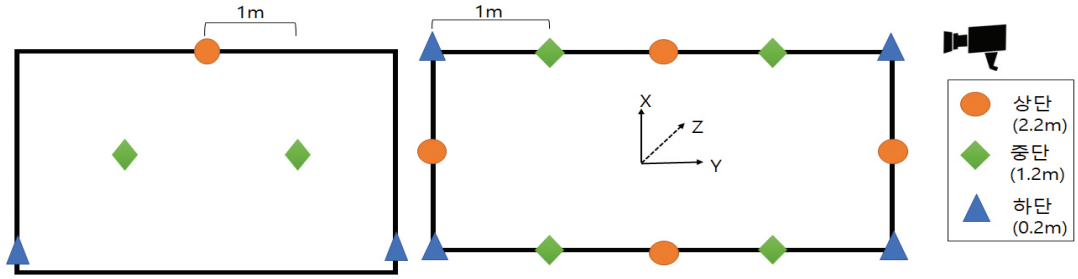
- 이는 3D 데이터를 얻기 쉬운 스튜디오 환경에서만 데이터 셋이 촬영되었기 때문
- 하지만 이 데이터 셋에 학습된 모델은 다양한 배경과 일상복을 입은 사람의 3D 인간 자세 검출을 정확하게 수행하는데 어려움을 겪게 됨
- 이에 촬영 시나리오에는 여러 사람이 야외 환경에서 일상복을 입고 촬영하는 부분을 최소 1만 클립 이상 확보

5.1.4 촬영 환경 구성

- 실내환경
  - 장애물로 인한 보이지 않는 영역, occlusion이 발생하지 않도록 12대의 카메라를 상/하/좌/우 배치하여 시간 동기화 된 영상을 얻을 수 있도록 함.
  - 추가 제안의 모션 캡처 데이터는 각각의 고속 카메라와 동기화 된 데이터를 가지고 있어야 하는데 비디오 클립에서 얻은 인체의 데이터와 모션 캡처 동작의 일치성을 확인하기 위해 필요 함
  - 만일 이 2개의 비디오 데이터와 모션 캡처 데이터가 다를 경우 이를 수정해 데이터의 정밀도를 높임

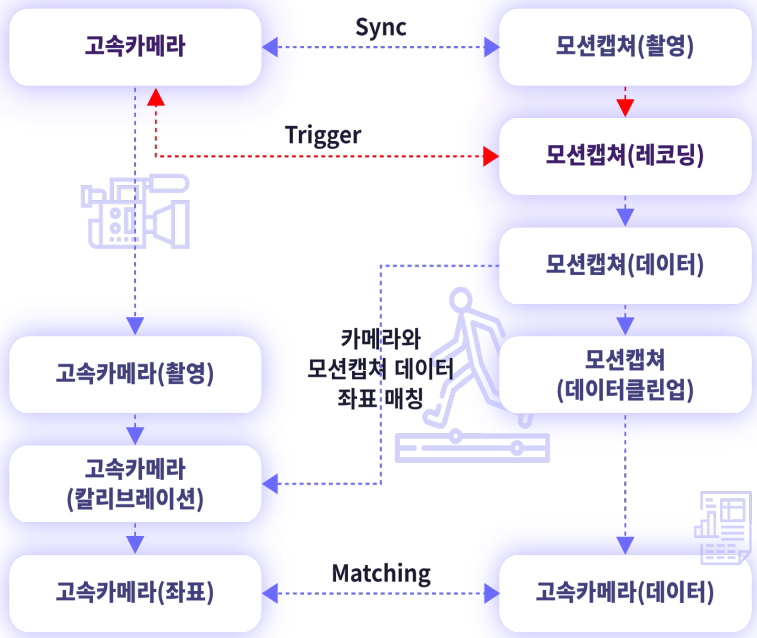
[그림 5] 촬영 방법 예시

[그림 6] 카메라 간격 및 높이(측면도, 평면도)



- 영상을 촬영하고 모션 캡처 촬영을 위한 약 8m x 8m 공간을 12대 정도의 고속 카메라를 동기화 하여 촬영을 위한 공간을 만듦
- 이 공간에서 자이로 슈트를 착용하고 그 위에 평상복을 입은 멀티 연기자의 동작을 캡처함
- 12대의 고속 카메라는 서로 간에 동기화와 트리거를 연결하여, 동시에 촬영을 해 영상 클립의 싱크를 맞추어 타임 동기화를 함

[그림 7] 촬영 환경 프로세스 예시



- 실외 환경 구축
  - 스튜디오 내에서 촬영이 불가한 동작들 경우에는 외부(실외)에서 촬영을 실시함
  - 카메라 설치, 이동 공간 표시 등의 사전 작업이 필요하며 큰 동작을 할 수 있는 체육관과 같은 넓은 공간에서 촬영을 준비
  - 액터 외에 영상에 등장하는 사람이 없도록 작업자는 별도의 공간에서 촬영을 진행함

• 실외 환경 구축

외부(옥상) 촬영

— 외부 환경과 유사한 옥상 또한 외부 촬영지로 적합하다고 판단

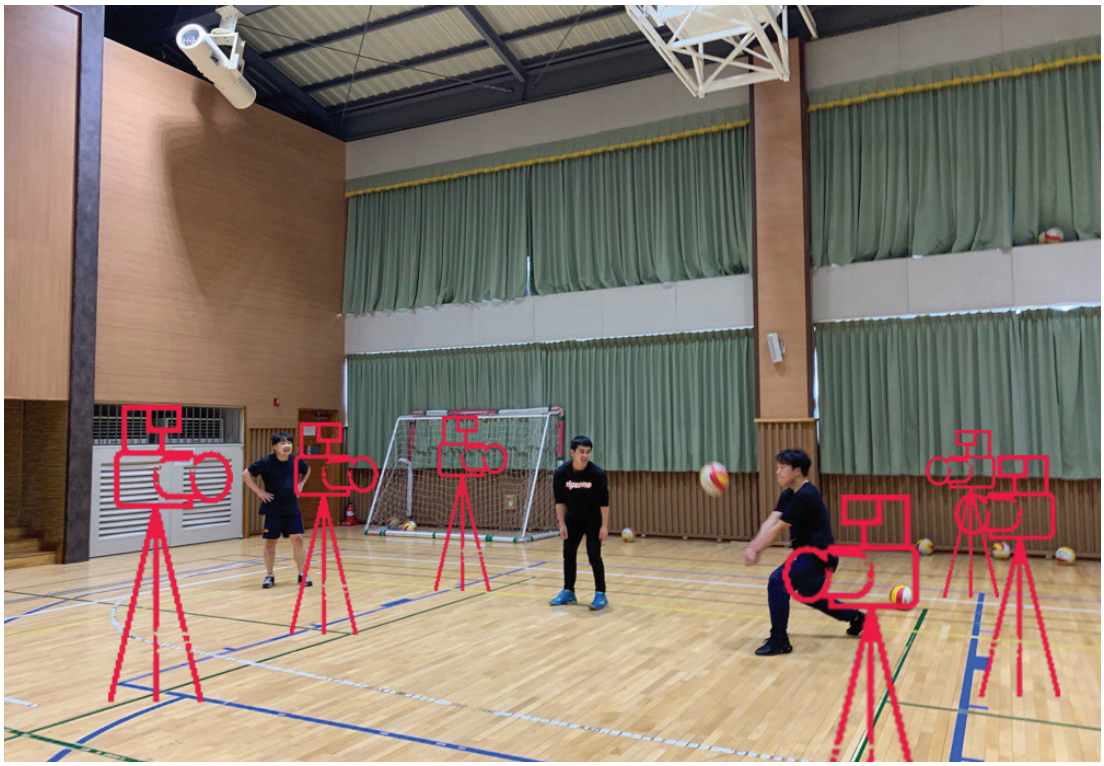
• 옥상 촬영의 경우 날씨의 영향을 받기 때문 촬영 날짜와 시기를 조율함

스튜디오	촬영기간	액터
B1	7월 ~ 11월	각 스튜디오 와 일정을 고려 액터 순환
옥상	10월	
체육관	10월	
창고 및 공터	10월 ~ 11월	

• 동적인 동작에 대하여 넓은 공간이 필요하여 100평 규모의 넓은 창고를 빌려 촬영 장소로 사용(주로 2~4인에 대하여 촬영)

• 창고의 경우 의자, 쇼파, 책상, 커텐 등의 가구를 변경해 가면서 촬영

[그림 8] 실외촬영 예시 1



[그림 9] 실외촬영 예시 2

[그림 10] 실외촬영 예시 3

○ 5.1.5 카메라 사양

HD Camera 1

sensor : 5.5 μm × 5.5 μm pixel size

Global Shutter

10 μs (0.01 ms) minimum shutter speed

1920 × 1080 maximum resolution

3800 μs (3.8 ms) maximum shutter speed

Image Processor :

On-board H.264 video compression (CAVLC 4:4:4 Intra Profile) with Constant Bit Rate (CBR) and Variable Bit Rate (VBR) compression options.

I/O : GigE/PoE (1000BASE-T) Ethernet .

Isolated Sync Output (for driving an eStrobe)

Lens & Filter : 850nm IR bandpass / 650nm IR cut filter switcher

HD Camera 2

Sensor : 13.2mm x 8.8mm CMOS , 3:2

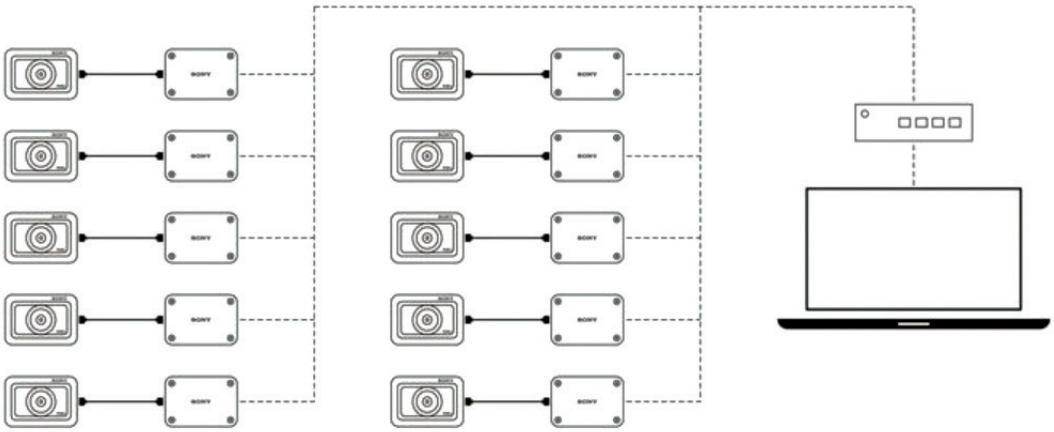
Lens : F4.0, f=7.9mm

FoV : 84도

I/O : Multi Fn USB

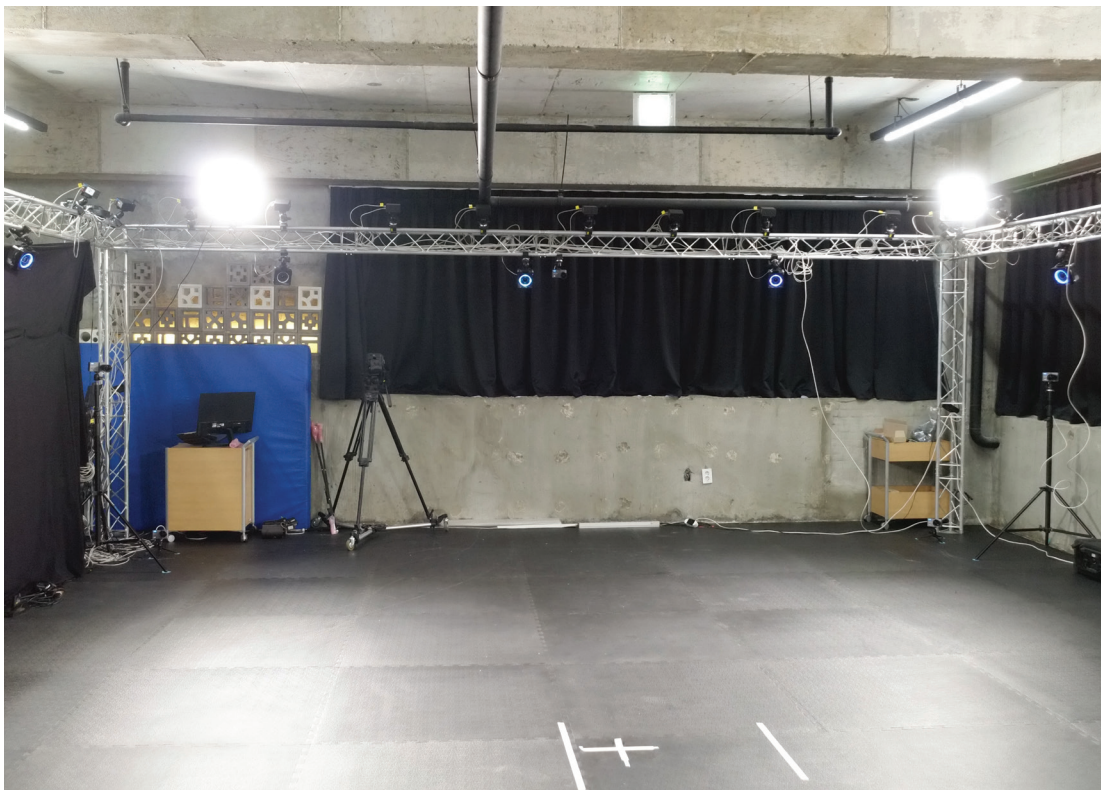
5.1.6 카메라 구성 및 연결

[그림 11] 멀티 콘트롤 및 동기화 구성

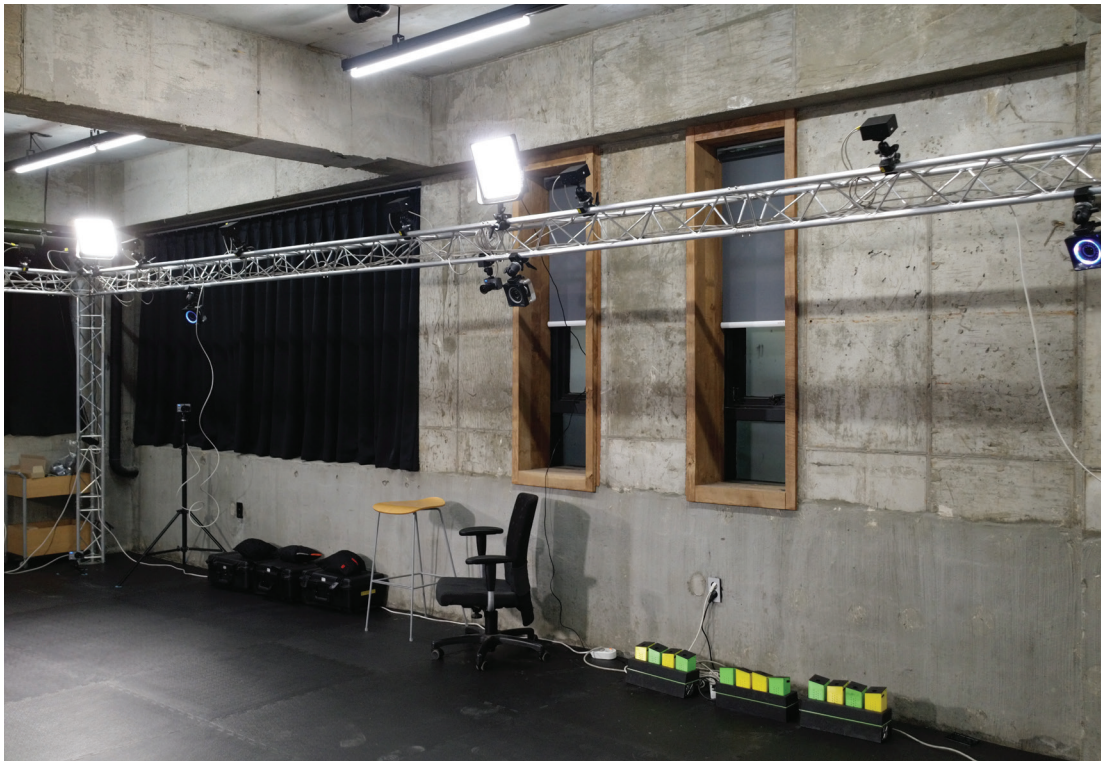


- 이더넷을 통해 스위치 허브와 연결 후 PC를 통한 멀티 콘트롤 및 동기화 구성

[그림 12] 실내 촬영 환경 1



[그림 13] 실내 촬영 환경 2



[그림 14] 실내 촬영 환경 3

○ 5.1.7 액터 및 촬영 계획

- 액터는 1명에서 4명까지 동시에 촬영을 진행함
- 액터의 연령, 성별, 신장은 별도 파일로 기록되며, 촬영 영상에 대하여 액터별 폴더를 생성하여 업로드 후 최종 데이터셋에 같이 반영될 예정임
- 총 2개의 팀으로 나누어 촬영을 진행함
- 각각의 팀은 12대의 카메라를 사용하면 다양한 앵글의 모션을 촬영 및 캡처
- 월별 촬영 수량은 아래와 같음 (단위 : 클립)
- 주변 환경 및 동작 종류, 시나리오 등의 변경에 따라 일부 변경 될 수 있음.

[표 4] 월별 수집 수량 계획

6월	7월	8월	9월	10월	11월	합계
5,000	15,000	46,000	46,000	46,000	42,000	200,000

● 5.2 데이터 가공

- 데이터 가공에는 여러 명의 작업자와 검수자를 동반하게 되며, 이들에 대하여 사전에 가공시스템을 사용할 수 있도록 사용자 등록, 작업배정 및 교육을 실시함
- Keypoint detection으로 추출된 2D joint 정보와 이미지를 웹상에서 확인하면서 잘못 추출된 joint에 대하여 수정하는 검수 단계를 수행함
- 2차례의 검수 단계를 통하여 데이터 셋 생성

[표 5] 데이터 가공 절차

구축단계	세부 절차	설명	필수여부
2. 데이터 가공	2.1 관절정보 정의	3D/2D 관절정보 정의	필수
	2.2 데이터 전처리	3D/2D 데이터 전처리	필수
	2.3 시스템 기반 가공 프로세스	시스템 기반 2D 데이터 가공 절차	필수
	2.4 2차원 관절 데이터 검수	2차 검수자가 DCAMP를 이용하여 어노테이션 데이터 확인 및 수정	필수

○ 5.2.1 관절 정보 정의

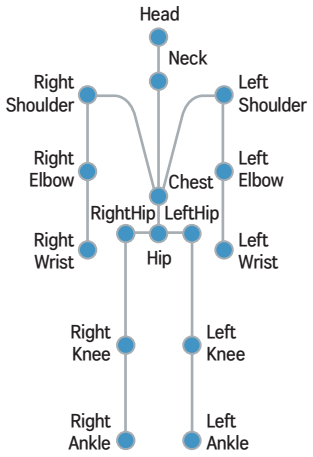
- 2D 관절데이터는 총 16개 속성으로 구성되며 아래 표와 같음

[표 6] 월별 수집 수량 계획

ID	정보	ID	정보
0	오른쪽 발	8	목 상단
1	오른쪽 무릎	9	머리
2	오른쪽 둔부	10	오른쪽 손목
3	왼쪽 둔부	11	오른쪽 팔꿈치
4	왼쪽 무릎	12	오른쪽 어깨
5	왼쪽 발	13	왼쪽 어깨
6	골반	14	왼쪽 팔꿈치
7	흉부	15	왼쪽 손목

- 3D 데이터를 담은 BVH에는 23개의 골격과 18 joint 정보로 되어 있으나 Human3.6M 규격과 동일한 16관절을 2D 학습 데이터셋으로 맞추기 위하여 아래와 같이 16 joint 정보만을 활용함.

[그림 15] 모션 캡처 데이터 포맷



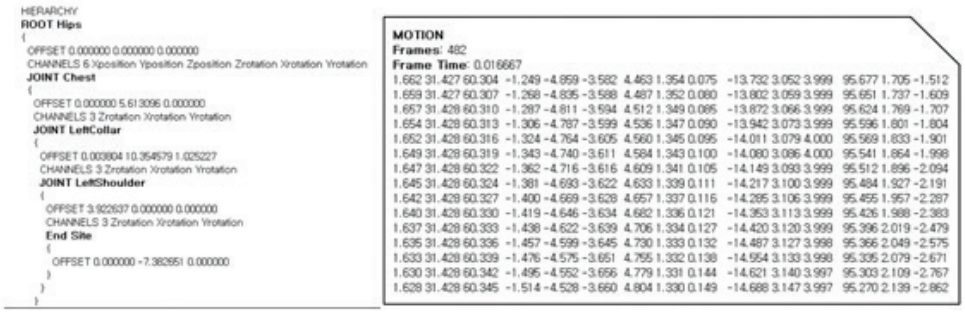
BVH파일 포맷

- BioVision사에 의해 개발되었고 모션 정보와 스켈레톤 정보를 계층적으로 표현하고 있음.
- 캡처한 캐릭터의 골격 구조를 정의하는 HIERARCHY SECTION과 동작을 기술하는 MOTION SECTION으로 구성
- 골격구조는 23개의 골격 세그먼트를 연결하는 18개의 joint로 표현되고 HIP을 루트로 해서 왼쪽하체, 오른쪽하체, 상체 방향의 순서로 세그먼트들을 연결하는 joint에 적용되는 오일러 각들의 나열
- 전체 프레임 수, 프레임 당 시간, 각 프레임에서의 이동, 회전 값들로 구성

[표 7] BVH 포맷 요소

JOINT 명칭	BVH	3D JSON
오른쪽 발	JOINT RightAnkle	right_ankle
오른쪽 무릎	JOINT RightKnee	right_knee
오른쪽 둔부	JOINT RightHip	right_hip
왼쪽 둔부	JOINT LeftHip	left_hip
왼쪽 무릎	JOINT LeftKnee	left_knee
왼쪽 발	JOINT LeftAnkle	left_ankle
골반	JOINT Hips	pelvis
흉부	JOINT Chest	thorax
목 상단	JOINT Neck	neck
머리	JOINT Head	head
오른쪽 손목	JOINT RightWrist	right_wrist
오른쪽 팔꿈치	JOINT RightElbow	right_elbow
오른쪽 어깨	JOINT RightShoulder	right_shoulder
왼쪽 어깨	JOINT LeftShoulder	left_shoulder
왼쪽 팔꿈치	JOINT LeftElbow	left_elbow
왼쪽 손목	JOINT LeftWrist	left_wrist

[그림 16] BVH파일 포맷 예시



5.2.2 데이터 전처리

• 5.2.2.1 3D 데이터 전처리

자이로 방식의 모션캡쳐의 경우 원칙 적으로 데이터의 손실을 일으키는 카메라 가림으로 인한 손실이나, 데이터 드롭 및 스왑등의 현상이 생기지 않아 원칙적으로는 데이터의 가공이 필요하지 않다. 다만, 데이터를 비디오와 동기를 맞추어 편집하는 과정이 필요하다.

- 모션캡처 데이터의 경우 데이터의 그래프를 깨끗하게 다듬는 이른바 클린업(Clean Up) 과정을 거쳐 데이터를 깨끗하게 정리한다. 또한 데이터 포맷을 일반적인 모션캡처 포맷인 *. bvh 포맷으로 변경
- 모션캡처 데이터의 경우 클린업과정에서 오류 탐색이 가능하고 데이터를 보정하는 작업 수행

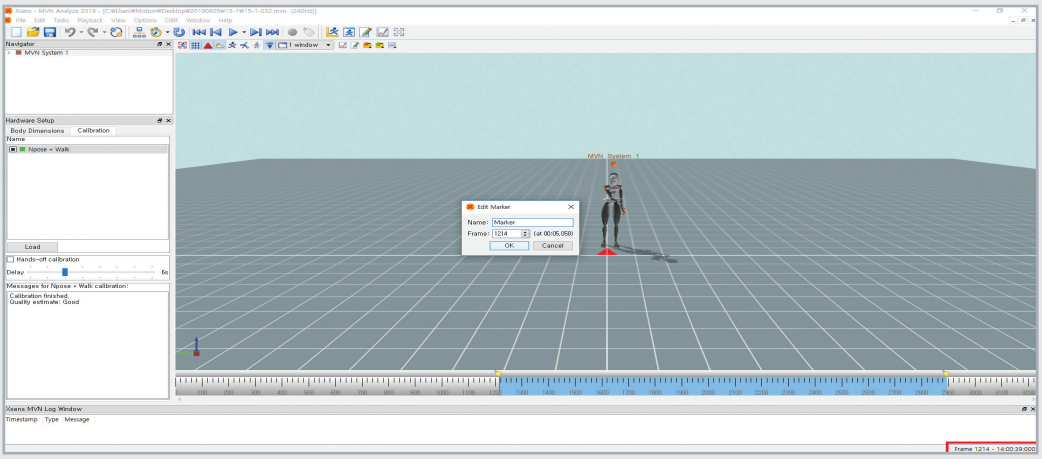
데이터 클린업 과정

- 영상 시스템 과 모션캡처 시스템으로 데이터 취득 후에는 양쪽 시스템에 동기화(Sync) 시키는 작업이 필요
- 동기화는 타임 코드를 활용하여 정확한 Sync 조절

1) Xsens (Motion Capture System_BVH)

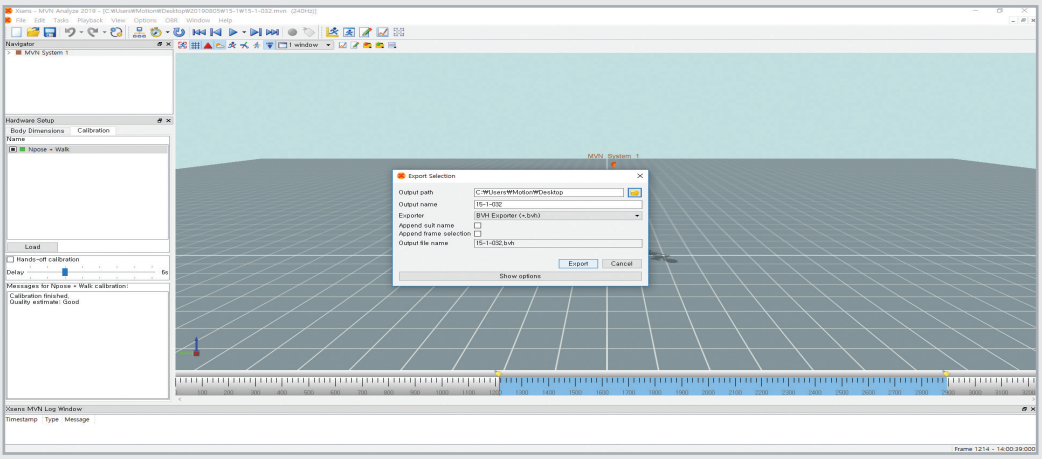
- 자이로 모션캡처로 데이터 획득 시 Audio LTC를 이용해 TimeCode 를 입력
- 입력 된 TimeCode 를 이용해 작업에 필요한 영역을 지정 하여 출력

[그림 17] Xsens 작업 1



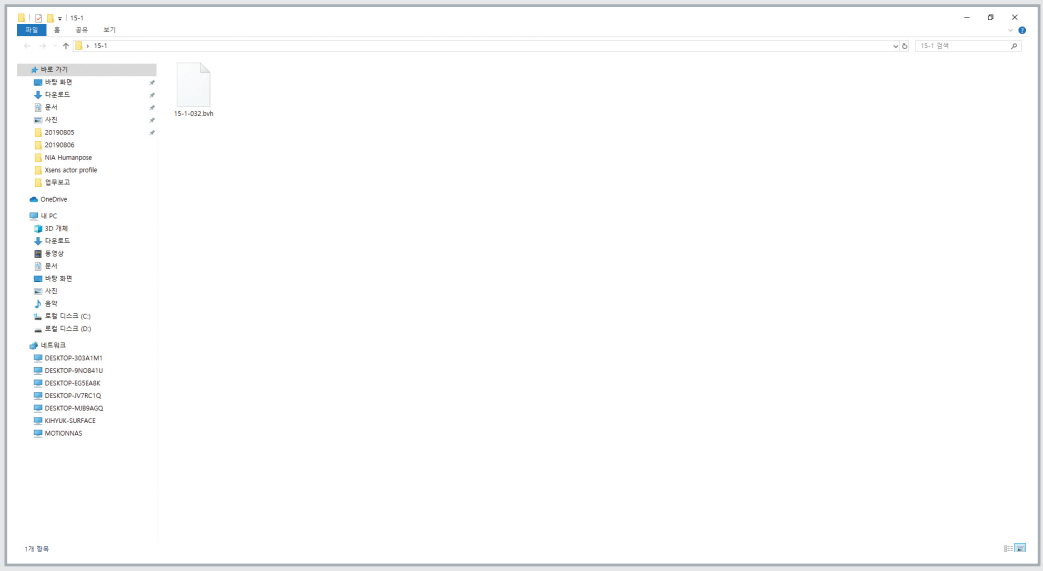
- Time Code 를 확인 하여 필요한 데이터 영역을 지정

[그림 18] Xsens 작업 2



- 지정된 영역을 BVH 파일로 출력

[그림 19] Xsens 출력 파일

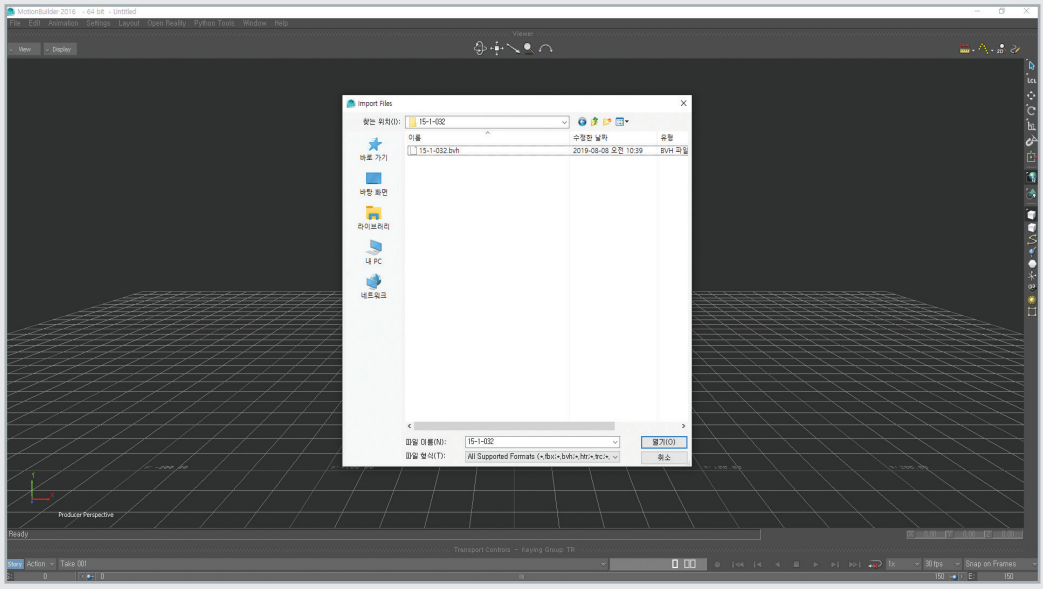


- 출력이 완료 된 BVH 파일은 240FPS 에서 30FPS 로 변환 과정이 필요
  (Motion Builder 를 활용해 변환)

2) Motion Builder (BVH파일 FPS 변환)

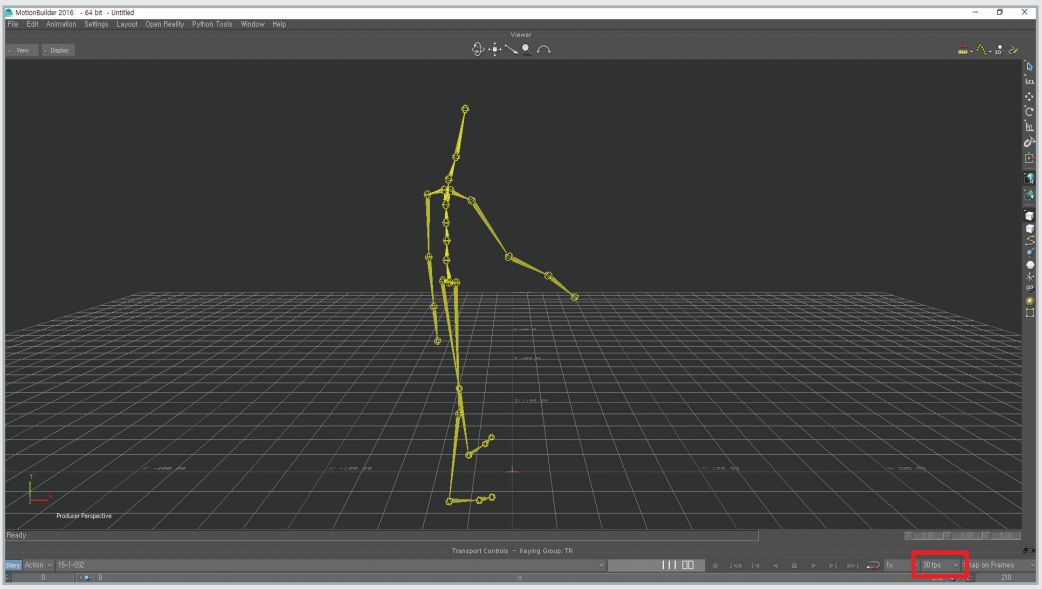
- Motion Builder 는 모션캡처 전용 에디팅 툴로써 촬영된 데이터를 편집 및 변환 가능
- 자이로 모션캡처로 촬영된 데이터를 30FPS 로 변환 하여 영상 데이터와 동일한 조건 형성

[그림 20] Motion Builder 작업1



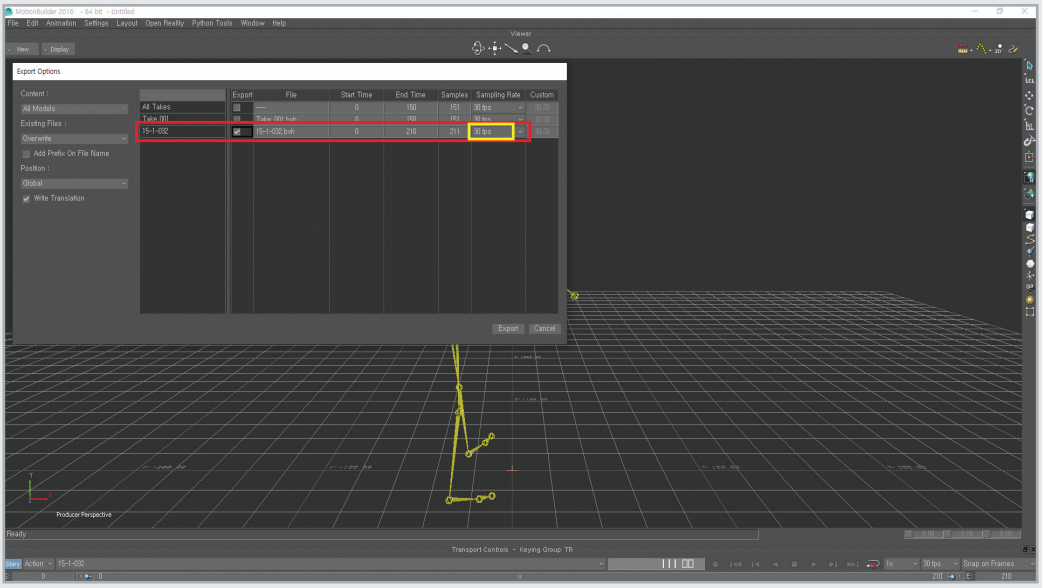
- 출력이 완료 된 BVH 파일을 Import

[그림 21] Motion Builder 작업2



- Import 가 완료 된 파일을 30FPS 로 변환

[그림 22] Motion Builder 작업3

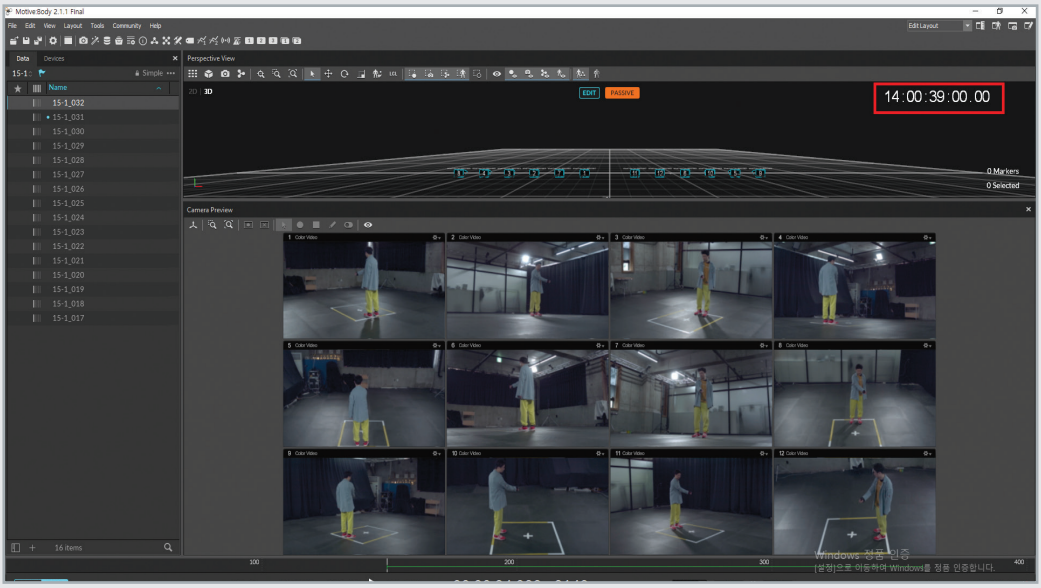


- 30FPS 로 변환 확인 후 Export

3) Optitrack System (영상)

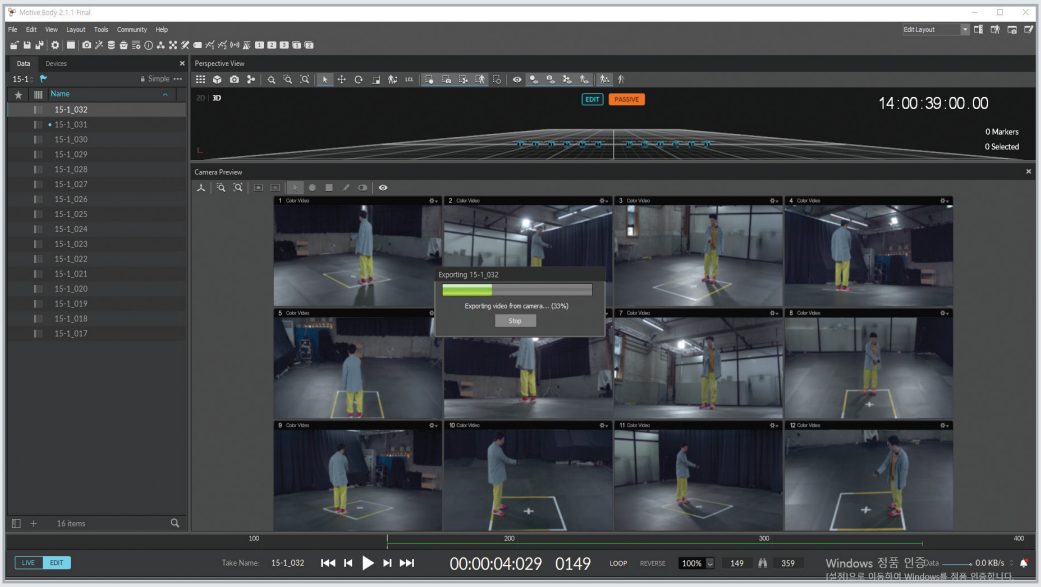
- Optitrack 에서는 E-sync 라는 별도의 장치에 외부 기기(Audio LTC)를 사용
- TimeCode 를 입력 하여 모션캡처 데이터와 Sync를 맞추는 작업이 필요

[그림 23] Optitrack System 작업1



- Time Code 를 활용 필요한 영역을 지정

[그림 24] Optitrack System 작업2

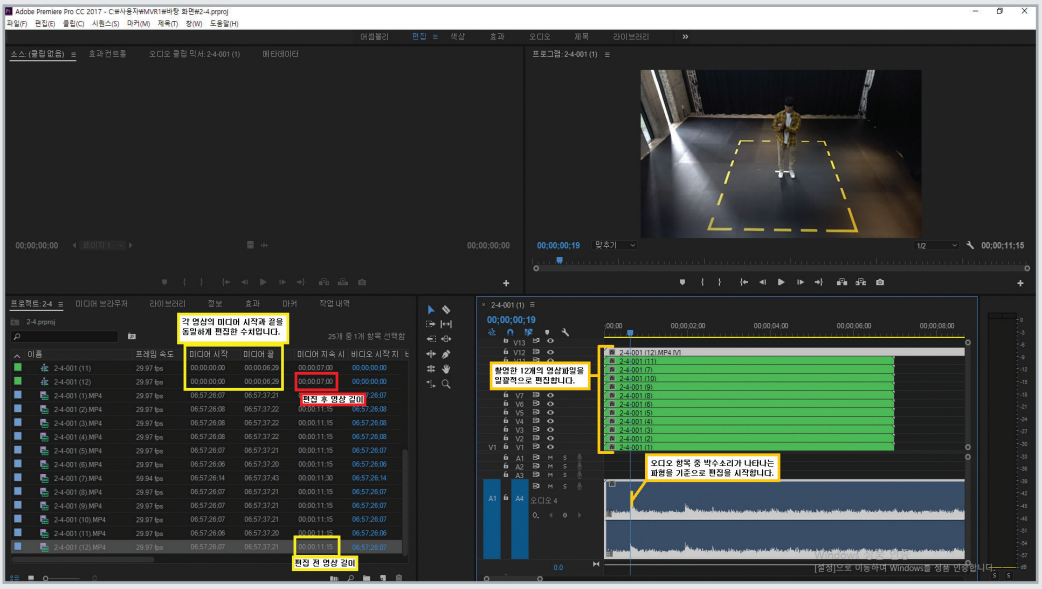


- 지정된 영역을 확인 후 Export

- Premiere (영상)

· 카메라를 이용한 영상에 경우 Audio 파형과 Sync 를 맞추는 작업이 필요
· 영상을 동일한 시점으로 편집

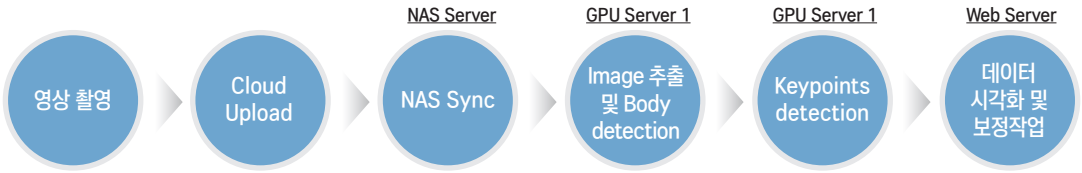
[그림 25] Premiere 영상편집



- Audio 신호를 확인 하여 필요한 영역을 지정 및 Export

- 일반 카메라의 경우 비디오카메라를 동기화 하고 타임코드를 기록하는 기능이 없음, 또한 다수의 카메라를 통해 촬영되는 만큼 동시 촬영을 위한 트리거 기능이 필요함. 이를 만족시키는 카메라는 방송급의 카메라뿐이나 이 경우 카메라의 가격이 수천만원을 호가해 사용이 어려움. 이에 고속 카메라의 기능 중 카메라 동기화, 타임코드 기록, 트리거 기능의 필요에 의해 고속 카메라를 사용
- 가공 결과물은 모션캡처 데이터와 촬영된 비디오 클립으로 구성
- 총 20만 비디오 클립 중 10%에 해당하는 2만 클립에 대하여 모션캡처 데이터를 포함할 예정임.

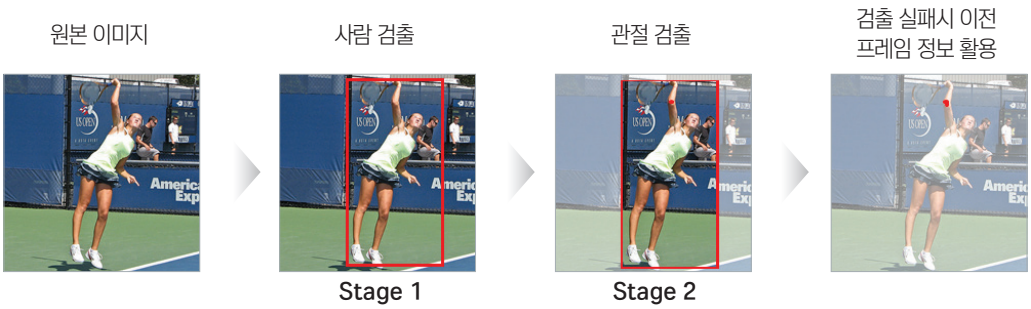
5.2.2.2 2D 데이터 전처리



- 촬영된 영상은 Cloud를 통하여 업로드되며 Cloud Sync 기능을 통하여 NAS에 동기화 됨
- GPU Server 1을 통하여 동영상으로부터 프레임당 3프레임의 이미지 추출 및 body detection 수행(이미지는 FTP를 통하여 GPU Server 2로 이관, body 정보는 DB 적재)
- GPU Server 2를 통하여 추출된 body 영역에 대하여 keypoint detection 수행(DB 적재)
- Web Server를 통하여 추출된 body 및 keypoint 데이터는 시각화되어 검수 단계에서 보정작업 대상이 됨
- 전처리에 의해 선별된 프레임에 대하여 이미지 추출
- 선별된 이미지에 대한 컨볼루션 기반의 Top-Down 모델을 활용하여 body bounding box 탐색

- 사람 검출을 통해 이미지를 crop하여 탐색 범위를 좁히고 정확도 향상
- Mask R CNN를 활용 사람을 검출 단계와 관절을 인식하는 단계로 나누어 2단계로 구성
- 연속형 이미지 정보를 활용하여 검출이 되지 않았을 경우 이전 프레임의 관절정보로 대체하여 검출 성능 향상

[그림 26] 관절 검출 과정



- 추출된 관절 데이터 JSON포맷 형식, 구조 요소

[표 8] 2D, 3D 관절정보 JSON 포맷 형식

2D Json 포맷정보	3D Json 포맷정보
{ "images":[{ "video_no":102, "img_no":17991, "img_path":"/studio_B1/20190701/3-1/3-1-08-Camera 9 (805316 625)_299.jpg", "width":1920, "height":1080, "action_category":"앉기" }, ...], "categories":[{ "supercategory":"person", "id":1, "name":"person", "keypoints":["right_ankle_x", "right_ankle_y", "right_ankle_view", "right_knee_x","right_knee_y", "right_knee_view", "right_hip_x", "right_hip_y", "right_hip_view", "left_hip_x", "left_hip_y", "left_hip_view", "left_knee_x", "left_knee_y", "left_knee_view", "left_ankle_x", "left_ankle_y", "left_ankle_view", "pelvis_x", "pelvis_y", "pelvis_view", "thorax_x", "thorax_y", "thorax_view", "neck_x", "neck_y", "neck_view", "head_x", "head_y", "head_view", "right_wrist_x", "right_wrist_y", "right_wrist_view", "right_elbow_x", "right_elbow_y", "right_elbow_view", "right_shoulder_x", "right_shoulder_y", "right_shoulder_view", "left_shoulder_x", "left_shoulder_y", "left_shoulder_view", "left_elbow_x", "left_elbow_y", "left_elbow_view", "left_wrist_x", "left_wrist_y", "left_wrist_view"]	{ "images":[{ "video_no":102, "img_no":17991, "img_path":"/studio_B1/20190701/3-1/3-1-08-Camera 9 (805316625)_ 299.jpg", "width":1920, "height":1080, "action_category":"앉기" }, ...], "categories":[{ "supercategory":"person", "id":1, "name":"person", "keypoints":["right_ankle_x", "right_ankle_y", "right_ankle_z", "right_ankle_roll", "right_ankle_pitch", "right_ankle_yaw", "right_knee_x", "right_knee_y", "right_knee_z", "right_knee_roll", "right_knee_pitch", "right_knee_yaw", "right_hip_x", "right_hip_y", "right_hip_z", "right_hip_roll", "right_hip_pitch", "right_hip_yaw", "left_hip_x", "left_hip_y", "left_hip_z", "left_hip_roll", "left_hip_pitch", "left_hip_yaw", "left_knee_x", "left_knee_y", "left_knee_z", "left_knee_roll", "left_knee_pitch", "left_knee_yaw", "left_ankle_x", "left_ankle_y", "left_ankle_z", "left_ankle_roll", "left_ankle_pitch", "left_ankle_yaw", "pelvis_x",

```
    }
  ],
  "annotations":[
    {
      "img_no":17991, "person_no":1,
      "bbox":[
        842, 445, 1154, 936
      ],
      "keypoints":[
        1077, 784, 2, 909, 529, 2, 866, 419, 2, 875,
        520, 2, 1067, 793, 2, 918, 558, 2, 909, 525,
        2, 827, 400, 2, 767, 460, 2, 1077, 529,
        2, 796, 393, 2, 777, 383, 2, 767, 393, 2,
        995, 424, 2, 822, 448, 2, 767, 383, 2],
      "num_keypoints":16
    },
    ...
  ]
}
```

```
        "pelvis_y", "pelvis_z", "pelvis_roll", "pelvis_pitch",
"pelvis_yaw",
        "thorax_x", "thorax_y", "thorax_z", "thorax_roll", "thorax_
pitch",
        "thorax_yaw", "neck_x", "neck_y", "neck_z", "neck_roll",
        "neck_pitch", "neck_yaw", "head_x", "head_y",
        "head_z", "head_roll", "head_pitch", "head_yaw",
        "right_wrist_x", "right_wrist_y", "right_wrist_z", "right_
wrist_roll",
        "right_wrist_pitch", "right_wrist_yaw", "right_elbow_x",
        "right_elbow_y", "right_elbow_z", "right_elbow_roll",
        "right_elbow_pitch", "right_elbow_yaw", "right_
shoulder_x",
        "right_shoulder_y", "right_shoulder_z", "right_shoulder_
roll",
        "right_shoulder_pitch", "right_shoulder_yaw",
        "left_shoulder_x", "left_shoulder_y", "left_shoulder_z",
        "left_shoulder_roll", "left_shoulder_pitch", "left_
shoulder_yaw",
        "left_elbow_x", "left_elbow_y", "left_elbow_z", "left_
elbow_roll",
        "left_elbow_pitch", "left_elbow_yaw", "left_wrist_x",
        "left_wrist_y", "left_wrist_z", "left_wrist_roll", "left_wrist_
pitch",
        "left_wrist_yaw"
      ]
    }
  ],
  "annotations":[
    {
      "img_no":17991,
      "person_no":1,
      "keypoints":[
        1077, 784, 2, 909, 529, 2, 866, 419, 2, 875, 520,
        2, 1067, 793, 2, 918, 558, 2, 909, 525, 2, 827, 400,
        2, 767, 460, 2, 1077, 529,2, 796, 393, 2, 777, 383, 2,
        767, 393, 2, 995, 424, 2, 822, 448, 2, 767, 383, 2,
        1077, 784, 2, 909, 529, 2, 866, 419, 2, 875, 520, 2,
        1067, 793, 2, 918, 558, 2, 909, 525, 2, 827, 400, 2,
        767, 460, 2, 1077, 529, 2, 796, 393, 2, 777, 383, 2, 767,
        393, 2, 995, 424, 2, 822, 448, 2, 767, 383, 2
      ],
      "num_keypoints":16
    },
    ...
  ]
}
```

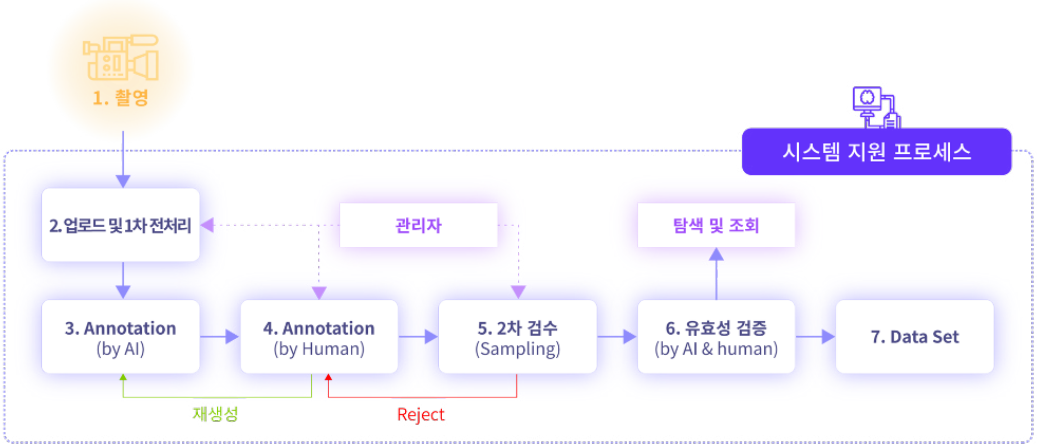

[표 9] 2D, 3D 관절정보 JSON포맷 요소

항목	내용
video_no	영상 번호
img_no	이미지 번호
img_path	이미지 파일 경로
action_category	동작 종류
keypoints	16개 관절 정보
person_no	이미지 내 annotation 대상 사람 번호
bbox	annotation 대상의 body box
num_keypoints	이미지 내 보여지는 관절의 수

5.2.3 시스템 기반 가공 프로세스

- 영상 및 이미지 추출
- Top-down 방식으로 Body Box 추출하여 1차 적재 수행
- keypoint-detection을 통하여 joint 추출
- 이미지, 2D(3D) joint 정보 DB 적재로 데이터 1셋을 구성함

[그림 27] 시스템 기반 가공 프로세스

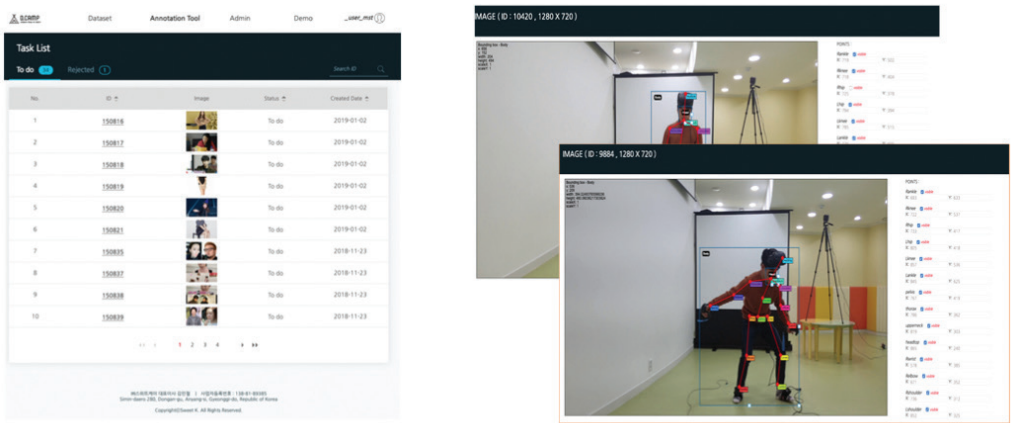


- 자동 생성된 데이터 셋은 Web view를 통하여 사용자가 검수를 수행함.

5.2.4 2차원 관절 데이터 검수

- keypoint detection에 의해 추출된 2차원 관절 정보에 대하여 2차 사람에 의한 이미지 위의 어노테이션 데이터 확인 및 수정(모든 데이터 대상)

[그림 28] 어노테이션 작업 검수 예시



작업 목록　　작업 및 검수

- 검수 및 수정은 학습데이터 생성 제품 DCAMP를 통하여 web 상에서 joint의 위치를 UI를 통하여 확인 및 마우스 드래그로 수정

5.3 데이터 검증

- 20만 클립의 2D annotation 데이터 검증에는 10% 정도 샘플링 검수하는 방법과 모델을 사용하는 방법이 가능하나, 10% 샘플 검수 자체도 2만 클립에 해당하는 많은 수량에 해당 함으로 모델에 의해 학습 후 테스트 데이터에 대하여 keyppoint를 추출하여 만들어진 학습 데이터와 비교를 통하여 검증하는 방식을 적용함.
- 추가 제안으로 만들어진 3D 데이터 검증은 오픈 포즈 데이터를 학습하여 추정된 좌표와 비교를 통하여 검증을 수행.

[표 10] 데이터 검증 절차

구축단계	세부 절차	설명	필수여부
5.3 데이터 검증	3.1 2D 데이터 검증	ground truth 데이터와 비교하여 성능 검증 방안	선택
	3.2 3D 데이터 검증	동일 모델로 오픈 포즈 데이터와의 비교 검증 방안	선택

5.3.1 2D 데이터 검증

- 2D 데이터 검증은 COCO에서 제공하는 툴을 사용하여 평가.
- 학습 모델은 많이 사용되고 있는 Simple Baselines for Human Pose Estimation and Tracking (ECCV 2018) of MSRA[2] 를 사용.
- 전체 50%의 데이터를 train 데이터로 하여 학습.
- 나머지 50%의 데이터에 대하여 10%씩 5개의 테스트 데이터 생성.
- 각 5개의 테스트 데이터에 대하여 COCO에서 제공하는 툴로 평가 수행
- 5개의 데이터 셋에서 관절별로 평가 결과 확인을 통하여 검증

5.3.2 3D 데이터 검증

- 먼저 비교작업에 사용할 validation set을 별도로 구분하여 저장
- 본 과제에서 개발된 딥러닝 학습 네트워크에 오픈 포즈 데이터를 입력데이터로 사용하여, 사람 동작 모델을 학습시키고, 이를 기반으로 구축한 추론시스템을 통해 3D 사람 동작 추론 결과로 추출되는 human body의 16개 keypoint의 3차원 좌표를 저장
- 본 과제에서 개발된 딥러닝 학습 네트워크에 본 과제에서 생성 및 구축한 사람 동작 데이터를 입력데이터로 사용하여, 사람 동작 모델을 학습시키고, 이를 기반으로 구축한 추론시스템을 통해 3D 사람 동작 추론 결과로 추출되는 human body의 16개 keypoint의 3차원 좌표를 저장
- 위의 첫 번째와 두 번째 결과물에서 각 관절의 keypoint별 3차원 좌표를 비교하여 정확도를 측정한다. 비교 방법에는 유클리디언 거리 비교 방식을 사용
- 활용하는 오픈 포즈 데이터로는 MPII[3], COCO, human 3.6m 등의 공개된 오픈 데이터 중 하나를 선정하여 활용

2) [1] Xiao, Bin, Haiping Wu, and Yichen Wei. "Simple Baselines for Human Pose Estimation and Tracking". ECCV 2018
3) http://human-pose.mpi-inf.mpg.de/#evaluation

- **Gound truth 정의** : 본 과제인 "사람 동작 영상 인공지능 학습용 데이터 구축 "사업에서는 여러 방향 및 각도에서의 사람동작에 대한 2차원의 동영상을 촬영하여 동기화된 시점 별로 각 이미지 프레임을 추출하여, 이를 기반으로 딥러닝 기반의 학습 네트워크를 통해 학습 모델을 추출함. 학습된 모델을 기반으로 2차원 이미지를 입력으로 받아 3차원 좌표의 추정값을 추론하는 것이므로, 데이터 수집 당시의 모든 사람 동작 데이터는 2차원 동영상 데이터 뿐만 아니라 3차원 좌표값도 또한 수집한다.
- **Gound truth 형태** : 사람 동작 영상을 촬영 시에 모든 actor는 모션 캡처 장비를 장착하게 되며, 동영상의 프레임과 정확하게 동기화되어 특정 시점에 위치한 key point의 3차원 좌표를 생성한다.

	frame 1	frame 2	...	frame 30
key point 1	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)
key point 2	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)
key point ...	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)
key point 17	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)	(x,y,z, p,r.y)

* x,y,z : x축, y축, z축의 position value.
* p,r,y : pitch, roll, yaw의 rotation value.

[그림 29] 모션캡처를 장비한 사용자 예시

5.4 데이터 활용

[표 11] 데이터 활용 절차

구축단계	세부 절차	설명	필수여부
5.4 데이터 활용	4.1 데이터 공개	AI hub를 통해 구축 데이터셋 공개	필수
	4.2 데이터 가공 시스템 공개	DCAMP 기능 중 pose어노테이션 툴 일부 수정하여 대외서비스 제공	필수
	4.3 모델 프로토 타입	확장된 mask r-cnn을 적용하여 3D 인간자세추정 프로토타입 개발	필수

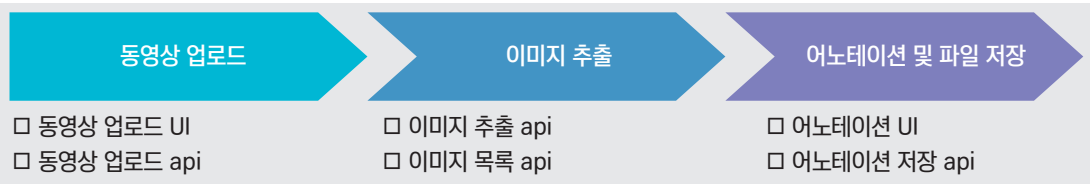
5.4.1 데이터 공개

- 사람 동작 영상 데이터 셋은 aihub 또는 github 등의 사이트를 통하여 다운로드 제공가능하며, 한국정보화진흥원의 정책에 따름
- 학습 데이터의 범용적인 활용 측면을 고려하여 카메라의 각도를 상/중/하로 나누었으며 이는 각각 CCTV, 홈 피트니스, 스마트시티 등에서 활용 할 수 있는 크게 3가지 범용적인 측면을 고려한 데이터 셋으로 공개되는 데이터 셋 또한 상단 셋, 중단 셋, 하단 셋으로 나누어서 공개

5.4.2 데이터 가공 시스템 공개

- 주관기관 스위트케이가 보유하고 있는 학습데이터 생성시스템 D.CAMP의 기능 중 pose 어노테이션 툴을 일부를 수정하여 동영상 업로드, 데이터 가공, 데이터 셋 다운로드 가능한 대외 서비스 제공

[그림 30] 대외 제공 소스 및 서비스



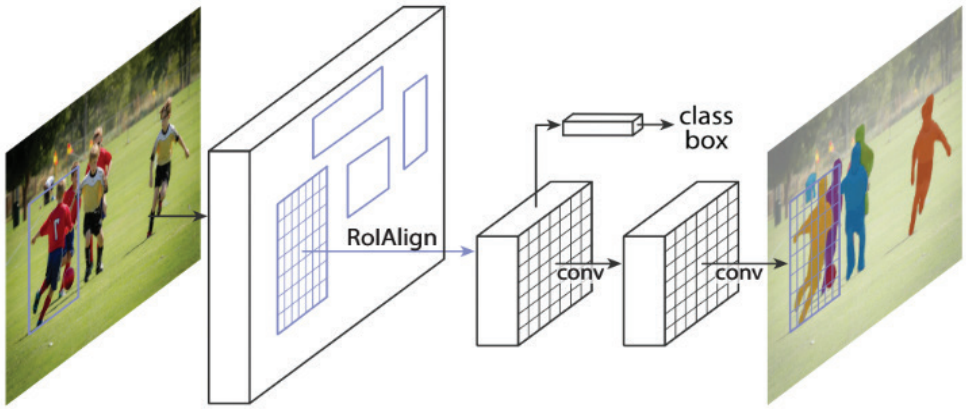
- 동영상 선택(동영상 선택(업로드) UI, 동영상 업로드 api)
- 이미지 추출(이미지 추출 api, 이미지 목록 api)
- 어노테이션 조회 및 데이터 셋 내려받기(어노테이션 UI, 어노테이션 저장 api)

5.4.3 모델 프로토 타입

- **5.4.3.1 모델 프로토타입 개발 개요**
- 프로토타입 모델은 각 사람의 3D 인간 자세 추정이 가능하도록 확장한 mask r-cnn (ICCV 2017)으로 정함

- Mask r-cnn은 개체기반의 컴퓨터비젼 task에서 높고 일관성 있는 성능을 보여주는 시스템으로서 기존의 mask r-cnn은 convolutional neural network (CNN) 기반의 시스템으로, 그림 13과 같이 여러 사람이 여러 크기, 위치에 존재하는 입력 이미지를 받아서 각 사람의 2D 인간 자세를 추정하는 시스템

[그림 31] Mask R-CNN의 개요도



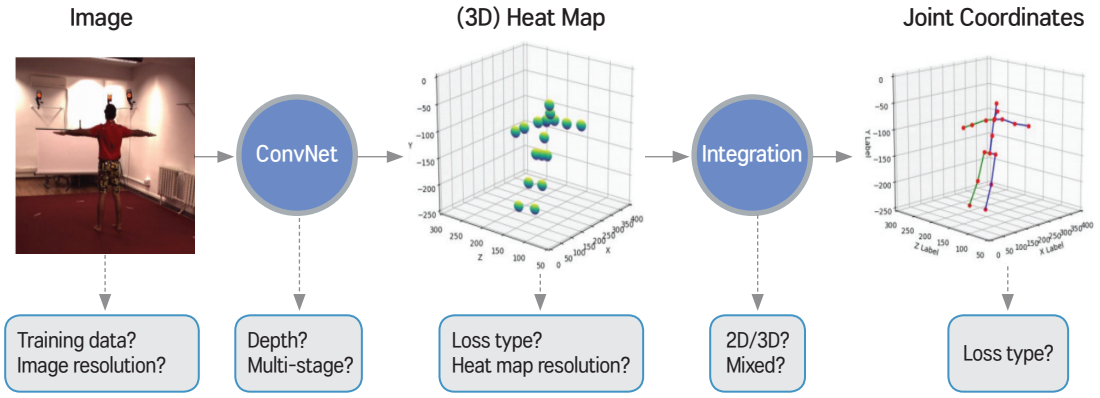
- 입력 영상에서 backbone network (residual network, CVPR 2016)이 유용한 feature를 추출하고, feature pyramid network (FPN, CVPR 2017)가 multi-scale feature map을 생성
- 생성된 multi-scale feature map으로부터 region proposal network (RPN)이 사람의 후보 영역을 추정하고, 해당 영역의 feature를 RoIAlign layer가 추출 함

[그림 32] Mask R-CNN의 2D 인간 자세 추정 결과

- 추출된 각 사람의 feature들은 2D 인간 자세 추정을 위한 branch로 넘겨지며, 이 branch에서 각 관절 당 2D heatmap을 추정하여 최종적인 인간 자세 추정을 수행
- 각 heatmap은 관절의 위치에서 큰 response를 갖도록 cross-entropy loss를 통해 supervise 됨

- **5.4.3.2 모델 프로토타입 개발 세부 내용**
- 기존 mask r-cnn을 3D 인간 자세 추정이 가능하도록 확장함
- 이를 위해 mask r-cnn이 2D heatmap이 아닌, 3D heatmap을 추정하도록 함
- 이 때 입력 이미지에 z축 (깊이) 정보가 없기 때문에 CNN model에서 channel dimension을 z축을 포함한 축으로 간주하여 3D heatmap을 추정하도록 함
- 또한, 정확도를 향상시키기 위해 그림 15와 같이 Integral human pose regression (ECCV 2018)에서 제안된 soft-argmax 방식의 추정을 수행

[그림 33] Integral human pose regression의 개요도



- Soft-argmax는 기존의 argmax가 갖는 한계점인 미분불가능을 해결하기 위한 연산으로, heatmap에서 관절의 좌표를 미분 가능하게 추출할 수 있게 해줌
- 이러한 soft-argmax 연산을 바탕으로 Integral Human Pose Regression은 단일 사람에 대한 3D 인간 자세 검출에서 높은 성능을 이루었음
- 이러한 soft-argmax를 확장된 mask r-cnn에 적용하기 위해 mask r-cnn에서 추정한 3D heatmap에서 3D 좌표를 미분 가능한 방식으로 추출
- 모델을 학습시킬 때는 추정된 좌표가 정답 좌표가 되도록 L1 loss를 통해 supervise를 하여 모델을 학습시킴

부속서 A

시나리오 및 데이터 예시

A.1 액션에 따른 상세 시나리오

※ 일부 시나리오는 촬영 환경에 따라 조금의 변경사항이 발생할 수 있으나 전체 수량에는 변화가 없습니다

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
1	1-1	기초	걷기	- 피사체가 왼손에 작은 책을 쥐고 오른쪽 팔꿈치를 15~30º 정도로 가볍게 구부리며 7Km/h 이하로 걷는다.	실내	1	2	2	1	2	1	12	96
1	1-2	기초	걷기	- 피사체가 오른쪽 어깨에 가방을 걸고 오른손으로 가방 끈을 쥐고 왼팔을 크게 흔들며 빠른걸음으로 걷는다.	실내	1	2	2	1	2	1	12	96
1	1-3	기초	걷기	- 운동화/구두/슬리퍼를 신고 빠르게 걷는다	실내	1	2	2	1	2	1	12	96
1	1-4	기초	걷기	- 운동화/구두/슬리퍼를 신고 소리 안나게 조용히 걷는다	실내	1	2	2	1	2	1	12	96
1	1-5	기초	걷기	- 피사체가 책/핸드폰을 보며 앞으로 천천히 걷는다	실내	1	2	2	1	2	1	12	96
1	1-6	기초	걷기	- 정면/오른쪽/왼쪽을 주시하며 걷는다	실내	1	2	2	1	2	1	12	96
1	1-7	기초	걷기	- 2명이 같이 정면을 보며 걷기	실내	2	2	2	1	2	1	12	96
1	1-8	기초	걷기	- 2명이 같이 핸드폰을 보며 걷기	실내	2	2	2	1	2	1	12	96
1	1-9	기초	걷기	- 2명이 어깨동무하고 걷기	실외	2	2	2	1	2	1	12	96
1	1-10	기초	걷기	- 2명중 한명은 핸드폰을 보며 걷기	실외	2	2	2	1	2	1	12	96
1	1-11	기초	걷기	- 한명이 다른 사람의 허리에 손을 두르고 걷기	실내	2	2	2	1	2	1	12	96
1	1-12	기초	걷기	- 2명이 같이 정면을 걷기 - 둘 중 한명은 가방을 들고	실내	2	2	2	1	2	1	12	96
1	1-13	기초	걷기	- 2명중 한명은 핸드폰을 보며 걷기 - 둘 중 한명은 가방을 들고	실내	2	2	2	1	2	1	12	96
1	1-14	기초	걷기	- 2명이 같이 정면을 걷기 - 둘 중 한명은 쇼핑백을 들고	실내	2	2	2	1	2	1	12	96
1	1-15	기초	걷기	- 2명중 한명은 핸드폰을 보며 걷기 - 둘 중 한명은 쇼핑백을 들고	실내	2	2	2	1	2	1	12	96
1	1-16	기초	걷기	- 2명이 같이 정면을 걷기 - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	1	12	96
1	1-17	기초	걷기	- 2명중 한명은 핸드폰을 보며 걷기 - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	1	12	96
1	1-18	기초	걷기	- 4명이 같은 방향으로 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-19	기초	걷기	- 4명이 같은 방향으로 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-20	기초	걷기	- 4명이 각자 다른 방향으로부터 교차하며 걷기(동,서,남,북) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-21	기초	걷기	- 4명이 각자 다른 방향으로부터 교차하며 걷기(동,서,남,북) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-22	기초	걷기	- 3명이 같은 방향으로 걷고 나머지 한명이 다른 방향으로 걷기(동, 남) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
1	1-23	기초	걷기	- 3명이 같은 방향으로 걷고 나머지 한명이 다른 방향으로 걷기(동, 서) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-24	기초	걷기	- 3명이 같은 방향으로 걷고 나머지 한명이 다른 방향으로 걷기(동, 남) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-25	기초	걷기	- 3명이 같은 방향으로 걷고 나머지 한명이 다른 방향으로 걷기(동, 서) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-26	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 남) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-27	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 서) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-28	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 남) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-29	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 서) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-30	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 서, 동, 서) - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-31	기초	걷기	- 2명이 같은 방향으로 걷고 나머지 - 2명이 다른 방향으로 걷기(동, 서, 동, 서) - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-32	기초	걷기	- 3명이 주위를 원 방향으로 돌고 나머지 한명이 원 안에서 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-33	기초	걷기	- 3명이 주위를 원 방향으로 돌고 나머지 한명이 원 안에서 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-34	기초	걷기	- 1명이 주위를 원 방향으로 돌고 나머지 - 3명이 원 안에서 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-35	기초	걷기	- 1명이 주위를 원 방향으로 돌고 나머지 - 3명이 원 안에서 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-36	기초	걷기	- 2명이 주위를 같은 원 방향으로 돌고 나머지 - 2명이 원 안에서 같은 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-37	기초	걷기	- 2명이 주위를 같은 원 방향으로 돌고 나머지 - 2명이 원 안에서 같은 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-38	기초	걷기	- 2명이 주위를 다른 원 방향으로 돌고 나머지 - 2명이 원 안에서 같은 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-39	기초	걷기	- 2명이 주위를 다른 원 방향으로 돌고 나머지 - 2명이 원 안에서 같은 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-40	기초	걷기	- 2명이 주위를 같은 원 방향으로 돌고 나머지 - 2명이 원 안에서 다른 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-41	기초	걷기	- 2명이 주위를 같은 원 방향으로 돌고 나머지 - 2명이 원 안에서 다른 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
1	1-42	기초	걷기	- 2명이 주위를 다른 원 방향으로 돌고 나머지 - 2명이 원 안에서 다른 직선방향으로 왕복 걷기 - 전부 정면을 보며 걷기	실내	4	2	2	1	2	1	12	96
1	1-43	기초	걷기	- 2명이 주위를 다른 원 방향으로 돌고 나머지 - 2명이 원 안에서 다른 직선방향으로 왕복 걷기 - 정면, 폰 보기, 백팩, 쇼핑백	실내	4	2	2	1	2	1	12	96
2	2-1	기초	달리기	- 피사체가 팔꿈치 가볍게 흔들며 조깅을 한다.	실외	1	2	2	1	2	1	12	96

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
2	2-2	기초	달리기	- 피사체가 팔꿈치 크게 흔들며 빠르게 달린다.	실외	1	2	2	1	2	1	12	96
2	2-3	기초	달리기	- 피사체가 백팩/가방을 매고 달린다.	실내	1	2	2	1	2	1	12	96
2	2-4	기초	달리기	- 피사체가 옆구리에 책/가방/쇼핑백 을 안고 달린다	실내	1	2	2	1	2	1	12	96
2	2-5	기초	달리기	- 100미터 달리기 준비 자세를 취한 후 출발	실내	1	2	2	1	2	1	12	96
2	2-6	기초	달리기	- 하얀선이 그어져 있고 선 띠를 달려서 통과한다(100미터 finish line)	실내	1	2	2	1	2	1	12	96
2	2-7	기초	달리기	- 뒤를 돌아보며 달려간다	실내	1	2	2	1	2	1	12	96
2	2-8	기초	달리기	- 2명이 같이 팔꿈치 가볍게 흔들며 조깅을 한다.	실내	2	2	2	1	2	1	12	96
2	2-9	기초	달리기	- 2명이 같이 팔꿈치 크게 흔들며 빠르게 달린다.	실내	2	2	2	1	2	1	12	96
2	2-10	기초	달리기	- 2명중 1명이 옆구리에 가방을 안고 달린다.	실외	2	2	2	1	2	1	12	96
2	2-11	기초	달리기	- 2명 중 1명이 뒤를 돌아보며 달려간다	실외	2	2	2	1	2	1	12	96
2	2-12	기초	달리기	- 2명이 같이 팔꿈치 가볍게 흔들며 조깅을 한다. - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	2	12	192
2	2-13	기초	달리기	- 2명이 같이 팔꿈치 크게 흔들며 빠르게 달린다. - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	2	12	192
2	2-14	기초	달리기	- 2명중 1명이 옆구리에 가방을 안고 달린다. - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	1	12	96
2	2-15	기초	달리기	- 2명 중 1명이 뒤를 돌아보며 달려간다 - 둘 중 한명은 백팩을 메고	실내	2	2	2	1	2	1	12	96
2	2-16	기초	달리기	- 2명 중 1명은 서서 신문을 보며 나머지 한명이 가볍게 조깅을 한다.	실내	2	2	2	1	2	1	12	96
2	2-17	기초	달리기	- 2명 중 1명은 서서 신문을 보며 나머지 한명이 빠르게 조깅을 한다.	실내	2	2	2	1	2	1	12	96
2	2-18	기초	달리기	- 2명 중 1명은 서서 신문을 보며 나머지 한명이 옆구리에 가방을 안고 달린다.	실내	2	2	2	1	2	1	12	96
2	2-19	기초	달리기	- 2명 중 1명은 서서 신문을 보며 나머지 한명이 뒤를 돌아보며 달려간다	실내	2	2	2	1	2	1	12	96
3	3-1	기초	앉기	- 피사체가 무릎을 굽혀 제 자리에 쪼그려 앉는다.	실외	1	2	2	1	2	1	12	96
3	3-2	기초	앉기	- 피사체가 제 자리에서 양반다리를 하며 앉는다.	실외	1	2	2	1	2	1	12	96
3	3-3	기초	앉기	- 피사체가 의자에 앉는다.	실내	1	2	2	3	2	1	12	288
3	3-4	기초	앉기	- 피사체가 의자에 다리를 꼬고 앉는다.	실내	1	2	2	3	2	1	12	288
3	3-5	기초	앉기	- 피사체가 두 무릎을 바닥에 대고 꿇어 앉는다.	실내	1	2	2	1	2	1	12	96
3	3-6	기초	앉기	- 피사체가 무릎을 꿇어 앉고 옆으로 비스듬히 앉는다.	실내	1	2	2	1	2	1	12	96
3	3-7	기초	앉기	- 피사체가 의자에 앉아 양 팔꿈치를 무릎에 대고 앉는다.	실내	1	2	2	3	2	1	12	288
3	3-8	기초	앉기	- 피사체가 의자에 앉아 다리를 꼰 상태로 펴고 앉는다.	실내	1	2	2	3	2	1	12	288
3	3-9	기초	앉기	- 피사체가 책상에 걸터 앉는다.	실내	1	2	2	2	2	1	12	192
3	3-10	기초	앉기	- 피사체가 엉덩이를 바닥에 대고 양 발바닥을 서로 붙여 앉는다.	실내	1	2	2	1	2	1	12	96

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
3	3-11	기초	앉기	- 왼쪽 무릎 세워 앉기	실내	1	2	2	1	2	1	12	96
3	3-12	기초	앉기	- 오른쪽 무릎 세워 앉기	실내	1	2	2	1	2	1	12	96
3	3-13	기초	앉기	- 피사체가 한 쪽 무릎을 바닥에 대고 나머지 무릎을 구부려 바닥에 앉는다.	실내	1	2	2	1	2	1	12	96
4	4-1	기초	인사 (목례+절)	- 피사체가 시선을 아래로 향하고 고개를 15º 정도 가볍게 숙인다.	실내	1	4	2	2	2	2	12	768
4	4-2	기초	인사 (목례+절)	- 의자에 앉은 - 피사체가 목례로 인사를 한다.	실내	1	4	2	2	2	2	12	768
4	4-3	기초	인사 목례+절)	- 90도로 인사	실내	1	4	2	2	2	2	12	768
4	4-4	기초	인사 (목례+절)	- 한국 전통 절 - 평절(남/여)	실내	1	4	2	2	2	2	12	768
4	4-5	기초	인사 (목례+절)	- 한국 전통 절 - 큰절(남/여)	실내	1	4	2	2	2	2	12	768
18	18-1	기초	인사 (손흔들기)	- 피사체가 오른손/왼손/양손을 손바닥이 보이게 머리 위로 올리고 흔들며 인사한다.	실외	1	4	2	2	2	2	12	768
18	18-2	기초	인사 (손흔들기)	- 피사체가 오른손/왼손/양손을 손바닥이 보이게 얼굴 옆에서 흔들며 인사한다.	실외	1	4	2	2	2	2	12	768
18	18-3	기초	인사 (손흔들기)	- 피사체가 오른손/왼손/양손을 손바닥이 보이게 가슴 앞에서 흔들며 인사한다.	실내	1	4	2	2	2	2	12	768
18	18-4	기초	인사 (손흔들기)	- 피사체가 오른손/왼손/양손을 손바닥이 보이게 머리 위로 올리고 흔들지 않고 인사한다.	실내	1	4	2	2	2	2	12	768
18	18-5	기초	인사 (손흔들기)	- 피사체가 왼손은 차렷자세로, 오른손은 눈썹으로 들어 경례한다.	실내	1	4	2	2	2	2	12	768
19	19-1	기초	인사 (악수)	- 2명이 서로 가볍게 인사한다.	실외	2	3	2	1	2	2	12	288
19	19-2	기초	인사 (악수)	- 2명이 서로 90도로 인사한다.	실외	2	3	2	1	2	2	12	288
19	19-3	기초	인사 (악수)	- 2명이 서로 악수한다.	실내	2	3	2	1	2	2	12	288
19	19-4	기초	인사 (악수)	- 2명이 서로 경례한다.	실내	2	3	2	1	2	2	12	288
19	19-5	기초	인사 (악수)	- 2명중 1명은 90도로 인사하고 다른 한명은 인사를 받아준다.	실내	2	3	2	1	2	2	12	288
5	5-1	기초	포옹	- 피사체가 양 팔로 상대방 어깨 위를 팔로 감싸 품에 껴안는다.	실외	2	4	2	1	2	2	12	384
5	5-2	기초	포옹	- 피사체가 양 팔로 상대방 허리를 팔로 감싸 품에 껴안는다.	실외	2	4	2	1	2	2	12	384
5	5-3	기초	포옹	- 피사체가 한 팔로 상대방 어깨 위를 팔로 감싸 품에 껴안는다.	실내	2	4	2	1	2	2	12	384
5	5-4	기초	포옹	- 피사체가 한 팔로 상대방 허리를 팔로 감싸 품에 껴안는다.	실내	2	4	2	1	2	2	12	384
5	5-5	기초	포옹	- 왼쪽 사람이 뒤에서 안아주며 포옹한다.	실내	2	2	2	1	2	2	12	192
5	5-6	기초	포옹	- 손을 목 뒤로 두른며 포옹을 한다.	실내	2	2	2	1	2	2	12	192
5	5-7	기초	포옹	- 온 몸을 밀착하며 포옹한다.	실내	2	2	2	1	2	2	12	192
6	6-1	기초	쓰러짐	- 피사체가 빠르게 바닥에 눕는 상태가 된다.	실내	1	5	2	1	2	2	12	480
6	6-2	기초	쓰러짐	- 피사체가 좌측/우측/앞으로 바닥에 누우며 쓰러진다.	실내	1	5	2	1	2	2	12	480
6	6-3	기초	쓰러짐	- 피사체가 뒤로 바닥에 누우며 쓰러진다.	실내	1	5	2	1	2	2	12	480

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
6	6-4	기초	쓰러짐	- 피사체가 걸어가다 넘어지며 쓰러진다.	실내	1	5	2	1	2	2	12	480
6	6-5	기초	쓰러짐	- 피사체가 뛰어가다 넘어지며 쓰러진다.	실내	1	5	2	1	2	2	12	480
6	6-6	기초	쓰러짐	- 피사체가 비틀거리다 넘어지며 쓰러진다.	실내	1	5	2	1	2	2	12	480
6	6-7	기초	쓰러짐	- 피사체가 쓰러진 후 일어나서 다시 쓰러진다.	실내	1	5	2	1	2	2	12	480
7	7-1	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 안보이는 경우 - 왼손이 위	실내	1	4	2	1	2	1	12	192
7	7-2	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 안보이는 경우 - 왼손이 위	실내	1	4	2	1	2	1	12	192
7	7-3	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 보이는 경우 - 왼손이 위	실내	1	4	2	1	2	1	12	192
7	7-4	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 보이는 경우 - 왼손이 위	실내	1	4	2	1	2	1	12	192
7	7-5	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 보이는 경우 - 왼손이 위 - 장갑착용	실내	1	4	2	1	2	1	12	192
7	7-6	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 보이는 경우 - 왼손이 위 - 장갑착용	실내	1	4	2	1	2	1	12	192
7	7-7	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 안보이는 경우 - 오른손이 위	실내	1	4	2	1	2	1	12	192
7	7-8	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 안보이는 경우 - 오른손이 위	실내	1	4	2	1	2	1	12	192
7	7-9	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 보이는 경우 - 오른손이 위	실내	1	4	2	1	2	1	12	192
7	7-10	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 보이는 경우 - 오른손이 위	실내	1	4	2	1	2	1	12	192
7	7-11	기초	팔짱끼기	- 앉아서 팔짱을 꼈을 때, 손등이 보이는 경우 - 오른손이 위 - 장갑착용	실내	1	4	2	1	2	1	12	192
7	7-12	기초	팔짱끼기	- 서서 팔짱을 꼈을 때, 손등이 보이는 경우 - 오른손이 위 - 장갑착용	실내	1	4	2	1	2	1	12	192
8	8-1	기초	식사동작	- 피사체가 식기를 이용하여 음식을 먹는다. - 한식	실내	1	4	2	2	2	1	12	384
8	8-2	기초	식사동작	- 피사체가 식기를 이용하여 음식을 먹는다. - 양식	실내	1	4	2	2	2	1	12	384
8	8-3	기초	식사동작	- 피사체가 손을 이용하여 음식을 먹는다.	실내	1	4	2	2	2	1	12	384
8	8-4	기초	식사동작	- 허리를 곧게 펴고 앉아 밥을 먹는다.	실내	1	4	2	2	2	1	12	384
8	8-5	기초	식사동작	- 오른 손을 무릎위에 놓고 밥을 먹는다.	실내	1	4	2	2	2	1	12	384
8	8-6	기초	식사동작	- 햄버거를 먹는다	실내	1	4	2	2	2	1	12	384
9	9-1	기초	점프	- 양발로 점프 하면서 양팔을 하늘로 올리는 동작	실내	1	4	2	1	2	2	12	384
9	9-2	기초	점프	- 양발로 점프 하면서 오른팔을 하늘로 올리는 동작	실내	1	4	2	1	2	2	12	384
9	9-3	기초	점프	- 양발로 점프 하면서 왼팔을 하늘로 올리는 동작	실내	1	4	2	1	2	2	12	384
9	9-4	기초	점프	- 제자리에서 오른발로 앞으로 점프	실내	1	4	2	1	2	2	12	384
9	9-5	기초	점프	- 제자리에서 왼발로 앞으로 점프	실내	1	4	2	1	2	2	12	384
9	9-6	기초	점프	- 제자리에서 양발로 앞으로 점프 (제자리 멀리 뛰기)	실내	1	4	2	1	2	2	12	384

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
9	9-7	기초	점프	- 달리다가 오른발/왼발/양발로 앞으로 점프	실내	1	4	2	1	2	2	12	384
10	10-1	기초	계단 오르기	- 발바닥 전체를 계단면에 닿으며 한/두/세 계단씩 걸어 올라간다.	실내	1	4	2	1	2	2	12	384
10	10-2	기초	계단 오르기	- 발바닥 앞부분을 계단면에 닿으며 한/두/세 계단씩 걸어 올라간다.	실내	1	4	2	1	2	2	12	384
10	10-3	기초	계단 오르기	- 발바닥 전체를 계단면에 닿으며 세 계단씩 뛰어 올라간다.	실내	1	4	2	1	2	2	12	384
10	10-4	기초	계단 오르기	- 발바닥 앞부분을 계단면에 닿으며 세 계단씩 뛰어 올라간다.	실내	1	4	2	1	2	2	12	384
10	10-5	기초	계단 오르기	- 발바닥 전체를 계단면에 닿으며 세 계단씩 뛰어 내려간다.	실내	1	4	2	1	2	2	12	384
10	10-6	기초	계단 오르기	- 발바닥 앞부분을 계단면에 닿으며 세 계단씩 뛰어 내려간다.	실내	1	4	2	1	2	2	12	384
10	10-7	기초	계단 오르기	- 발바닥 전체를 계단면에 닿으며 한/두/세 계단씩 걸어 내려간다.	실내	1	4	2	1	2	2	12	384
10	10-8	기초	계단 오르기	- 발바닥 앞부분을 계단면에 닿으며 한/두/세 계단씩 걸어 내려간다.	실내	1	4	2	1	2	2	12	384
11	11-1	기초	푸쉬업	- 바닥에 손을 짚고 기본적인 푸쉬업을 실시한다.	실내	1	6	2	1	2	2	12	576
11	11-2	기초	푸쉬업	- 의자 등의 물체에 손을 짚어 상체를 띄우고 푸쉬업을 실시한다.	실내	1	6	2	1	2	2	12	576
11	11-3	기초	푸쉬업	- 푸쉬업 기본 동작에서 무릎을 바닥에 닿게하여 푸쉬업을 실시한다.(무릎 푸쉬업)	실외	1	6	2	1	2	2	12	576
11	11-4	기초	푸쉬업	- 기본 푸쉬업동작에서 왼쪽다리를 왼쪽 팔꿈치쪽으로, 오른쪽 다리는 오른쪽 팔꿈치 쪽으로 굽혀서 닿게하며 동작을 실시한다.(스파이더맨 푸쉬업)	실외	1	6	2	1	2	2	12	576
12	12-1	기초	방향 지시	- 손가락으로 방향지시 - 앉아서	실내	1	4	2	1	2	2	12	384
12	12-2	기초	방향 지시	- 손바닥이 보이게 방향 지시 - 앉아서	실내	1	4	2	1	2	2	12	384
12	12-3	기초	방향 지시	- 양손으로 공손하게 방향 지시 - 앉아서	실내	1	4	2	1	2	2	12	384
12	12-4	기초	방향 지시	- 손가락으로 방향지시	실내	1	4	2	1	2	2	12	384
12	12-5	기초	방향 지시	- 손바닥이 보이게 방향 지시	실내	1	4	2	1	2	2	12	384
12	12-6	기초	방향 지시	- 양손으로 공손하게 방향 지시	실내	1	4	2	1	2	2	12	384
13	13-1	기초	박수치기	- 자리에 앉았다가 일어서면서 박수를 친다.	실내	1	4	2	1	2	2	12	384
13	13-2	기초	박수치기	- 의자에 앉아 가슴높이에서 가볍게 박수를 친다.	실내	1	4	2	1	2	2	12	384
13	13-3	기초	박수치기	- 팔을 들어 머리 위에서 손뼉을 마주친다.	실내	1	4	2	1	2	2	12	384
13	13-4	기초	박수치기	- 몸을 흔들며 박수를 친다. (5:24~6:20)	실내	1	4	2	1	2	2	12	384
13	13-5	기초	박수치기	- 대~한민국 응원 박수를 친다	실내	1	4	2	1	2	2	12	384
14	14-1	기초	윗몸 일으키기	- 혼자 바닥에 등을 대고 누워 무릎을 구부리고 발이 바닥에 닿게 위치한다. 몸을 일으키며 오른쪽 팔꿈치가 왼쪽 무릎에 닿도록 상체를 일으킨다. 다시원 위치한다.	실내	1	4	2	1	2	2	12	384

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
14	14-2	기초	윗몸 일으키기	- 한사람은 바닥에 눕고 다른 한 사람은 다리를 잡아 주고 윗몸일으키기를 완벽하게 한다(팔굼치가 무릎에 닿으면 안됨)	실내	1	4	2	1	2	2	12	384
14	14-3	기초	윗몸 일으키기	- 바닥에 누워 왼쪽 팔꿈치가 오른쪽 무릎에 닿게 상체를 일으키고 다리를 움직인다.(트위스트 크런치)	실내	1	4	2	1	2	2	12	384
14	14-4	기초	윗몸 일으키기	- 윗몸일으키기 자세로 몸을 일으켜 상체를 좌우로 비튼다.	실내	1	4	2	1	2	2	12	384
14	14-5	기초	윗몸 일으키기	- 한사람은 바닥에 눕고 다른 한 사람은 다리를 잡아 주고 윗몸일으키기를 완벽하게 한다(팔굼치가 무릎에 닿아야 함)	실내	2	4	2	1	2	2	12	384
14	14-6	기초	윗몸일으키기	- 한사람은 바닥에 눕고 다른 한 사람은 다리를 잡아 주고 윗몸일으키기를 완벽하게 한다(팔굼치가 무릎에 닿으면 안됨)	실내	2	4	2	1	2	2	12	384
15	15-1	기초	포복	- (군인) 낮은포복으로 전진한다	실내	1	4	2	1	2	9	12	1,728
15	15-2	기초	포복	- (군인) 높은포복으로 전진한다	실내	1	4	2	1	2	9	12	1,728
15	15-3	기초	포복	- (군인) 응용포복으로 전진한다	실내	1	4	2	1	2	9	12	1,728
15	15-4	기초	문열기	- 도어락 키패드를 눌러 문을 연다.	실내	1	-	2	1	2	2	12	-
15	15-5	기초	문열기	- 방안의 창문을 연다. - 미닫이 / 여닫이	실내	1	-	2	1	2	2	12	-
15	15-6	기초	문열기	- 버튼을 눌러 문을 연다	실내	1	-	2	1	2	2	12	-
15	15-7	기초	문열기	- 냉장고 문을 연다.	실내	1	-	2	1	2	2	12	-
16	16-1	기초	태권도	- 태극 1장	실내	1	1	1	1	2	12	12	288
16	16-2	기초	태권도	- 태극 2장	실내	1	1	1	1	2	12	12	288
16	16-3	기초	태권도	- 태극 3장	실외	1	1	1	1	2	12	12	288
16	16-4	기초	태권도	- 태극 4장	실내	1	1	1	1	2	12	12	288
16	16-5	기초	태권도	- 태극 5장	실내	1	1	1	1	2	12	12	288
16	16-6	기초	태권도	- 태극 6장	실내	1	1	1	1	2	12	12	288
16	16-7	기초	태권도	- 태극 7장	실내	1	1	1	1	2	12	12	288
16	16-8	기초	태권도	- 태극 8장	실내	1	1	1	1	2	12	12	288
16	16-9	기초	싸움	- 앞차기로상대방을가격하는동작을한다.	실내	2	-	2	1	2	2	12	-
16	16-10	기초	싸움	- 몽둥이로상대방을가격하는동작을한다.	실내	2	-	2	1	2	2	12	-
16	16-11	기초	싸움	- 칼로상대방을위협하는동작을한다.	실내	2	-	2	1	2	2	12	-
16	16-12	기초	싸움	- 상대방을잡아서넘어뜨리는동작을한다.	실내	2	-	2	1	2	2	12	-
16	16-13	기초	싸움	- 상대방을뒤에서양손으로감아잡는동작을한다.	실내	2	-	2	1	2	2	12	-
17	17-1	기초	국군도수체조	- 제자리에 서서 다른 사람과 마주보며 이야기를 나눈다.	실내	1	1	2	1	1	30	12	720
17	17-2	기초	국군도수체조	- 제자리에 서서 다른 사람과 마주보며 이야기를 나눈다.	실외	4	1	2	1	1	30	12	720
17	17-3	기초	대화하기	- 2명 모두 의자에 앉아서 대화한다.	실내	2	-	2	1	2	2	12	-

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
17	17-4	기초	대화하기	- 2명 모두 책상을 사이에 두고 앉아서 대화한다.	실내	2	-	2	1	2	2	12	-
17	17-5	기초	대화하기	- 2명 중 한명은 서있고 한명은 의자에 앉아 대화한다.	실내	2	-	2	1	2	2	12	-
17	17-6	기초	대화하기	- 2명 중 한명은 책상 반대에 서있고 나머지 한명은 책상 반대편에 의자에 앉아 대화한다.	실내	2	-	2	1	2	2	12	-
20	20-1	의미론적	요가	- KT 서비스 예정인 요가 프로그램 자세를 따라한다	실내	1	4	2	1	1	50	12	4,800
20	20-2	의미론적	요가	- 황아영의 요가자세 1~ 8편을 따라하며 자세를 취한다	실내	1	4	2	1	1	50	12	4,800
21	21-1	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(레드벨벳 -빨간맛) (0:00~0:21)	실내	1	2	2	1	2	20	12	1,920
21	21-2	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(트와이스 -likey) (1:29~1:40)	실내	1	2	2	1	2	20	12	1,920
21	21-3	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(idol)(0:20~3:58)	실내	1	2	2	1	2	20	12	1,920
21	21-4	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(solo)	실내	1	2	2	1	2	20	12	1,920
21	21-5	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(벌써 12시)	실내	1	2	2	1	2	20	12	1,920
21	21-6	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(레드벨벳-빨간맛) (0:00~0:21)	실외	4	1	2	1	2	20	12	960
21	21-7	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(트와이스-likey) (1:29~1:40)	실외	4	1	2	1	2	20	12	960
21	21-8	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(idol)(0:20~3:58)	실외	4	1	2	1	2	20	12	960
21	21-9	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(solo)	실외	4	1	2	1	2	20	12	960
21	21-10	의미론적	댄스 (아이돌) 여러명	- 특정한 안무의 춤을 춘다.(벌써 12시)	실외	4	1	2	1	2	20	12	960
22	22-1	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(오나나나춤)(0:21)	실내	1	2	2	1	2	20	12	1,920
22	22-2	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(baby shark)(0:45)	실내	1	2	2	1	2	20	12	1,920
22	22-3	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(외계인춤)(0:58)	실내	1	2	2	1	2	20	12	1,920
22	22-4	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(셔플댄스)(3:20~)	실내	1	2	2	1	2	20	12	1,920
22	22-5	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(댑댄스)(9:20~)	실내	1	2	2	1	2	20	12	1,920
22	22-6	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(오나나나춤)(0:21)	실내	4	1	2	1	2	20	12	960
22	22-7	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(baby shark)(0:45)	실내	4	1	2	1	2	20	12	960
22	22-8	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(외계인춤)(0:58)	실내	4	1	2	1	2	20	12	960

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
22	22-9	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(셔플댄스)(3:20~)	실내	4	1	2	1	2	20	12	960
22	22-10	의미론적	댄스 (클럽 외) 여러명	- 특정한 안무의 춤을 춘다.(댑댄스)(9:20~)	실내	4	1	2	1	2	20	12	960
23	23-1	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (룸바)	실내	1	2	2	1	2	20	12	1,920
23	23-2	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (자이브)	실내	1	2	2	1	2	20	12	1,920
23	23-3	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (차차차)	실내	1	2	2	1	2	20	12	1,920
23	23-4	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (왈츠)	실내	1	2	2	1	2	20	12	1,920
23	23-5	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (탱고)	실내	1	2	2	1	2	20	12	1,920
23	23-6	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (룸바)	실내	4	1	2	1	2	20	12	960
23	23-7	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (자이브)	실내	4	1	2	1	2	20	12	960
23	23-8	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (차차차)	실내	4	1	2	1	2	20	12	960
23	23-9	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (왈츠)	실내	4	1	2	1	2	20	12	960
23	23-10	의미론적	댄스 (스포츠) 여러명	- 댄스스포츠를 춘다. (탱고)	실내	4	1	2	1	2	20	12	960
24	24-1	의미론적	악기연주 (기타)	- 기타를 연주한다. (3 -4가지 곡 선택)	실내	1	4	2	1	2	20	12	3,840
25	25-1	의미론적	악기연주 (전자 피아노)	- 전자피아노를 연주한다. (3 -4가지 곡 선택)	실내	1	4	2	1	2	20	12	3,840
26	26-1	의미론적	악기연주 (드럼)	- 드럼을 연주한다. (3 -4가지 곡 선택)	실내	1	4	2	1	2	20	12	3,840
27	27-1	의미론적	훌라후프	- 훌라후프를 10회 한다.	실내	1	10	2	1	2	2	12	960
27	27-2	의미론적	훌라후프	- 훌라후프를 20회 한다.	실내	1	10	2	1	2	2	12	960
27	27-3	의미론적	훌라후프	- 훌라후프를 30회 한다.	실내	1	10	2	1	2	2	12	960
28	28-1	의미론적	웨이트트레이닝	- 피사체가 앉았다 일어나는 자세를 반복한다(스쿼트)	실내	1	4	2	1	2	3	12	576
28	28-2	의미론적	웨이트트레이닝	- 피사체가 다리를 앞뒤로 벌리고 앉았다 일어나는 자세를 반복한다(런지)	실내	1	4	2	1	2	3	12	576
28	28-3	의미론적	웨이트트레이닝	- 피사체가 업드린채로 팔을 굽혔다 펴는 자세를 반복한다(푸쉬업)	실내	1	4	2	1	2	3	12	576
28	28-4	의미론적	웨이트트레이닝	- 피사체가 팔꿈치로 상체를 받친채로 엎드린 자세를 유지한다(플랭크)	실내	1	4	2	1	2	3	12	576
28	28-5	의미론적	웨이트트레이닝	- 피사체가 누운채로 상체를 일으키는 동작을 반복한다(크런치)	실내	1	4	2	1	2	3	12	576
28	28-6	의미론적	웨이트트레이닝	- 피사체가 누운채로 하체를 상하로 일으키는 동작을 반복한다(레그레이즈)	실내	1	4	2	1	2	3	12	576
28	28-7	의미론적	웨이트트레이닝	- 피사체가 누운채로 하체를 좌우로 움직이는 동작을 반복한다(와이퍼)	실내	1	4	2	1	2	3	12	576

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
28	28-8	의미론적	웨이트트레이닝	- 피사체가 서있는 자세에서 측면기준으로 위아래로 올렸다내렸다하는 동작을 반복한다(사이드레터럴 레이즈)	실내	1	4	2	1	2	3	12	576
28	28-9	의미론적	웨이트트레이닝	- 피사체가 서있는 자세에서 정면기준으로 위아래로 올렸다내렸다하는 동작을 반복한다(프론트 레이즈)	실내	1	4	2	1	2	3	12	576
28	28-10	의미론적	웨이트트레이닝	- 피사체가 뒤에 팔을 받친 자세로 팔을 굽혔다 펴는 동작을 반복한다(벤치 딥스)	실내	1	4	2	1	2	3	12	576
28	28-11	의미론적	웨이트트레이닝	- 피사체가 두손으로 물체를 든 채로 상하로 움직이는 동작을 반복한다(케틀벨 스윙)	실내	1	4	2	1	2	3	12	576
28	28-12	의미론적	웨이트트레이닝	- 피사체가 업드린채로 팔을 굽혔다 펴는 자세를 반복한다(푸쉬업)	실내	1	4	2	1	2	3	12	576
28	28-13	의미론적	웨이트트레이닝	- 피사체가 앉은채로 한손으로 물체를 잡아 머리 뒤로 두고 다른 한손으로 팔꿈치를 받쳐준 후 상하로 움직이는 동작을 반복한다(원 암 트라이셉스 익스텐션)	실내	1	4	2	1	2	3	12	576
28	28-14	의미론적	웨이트트레이닝	- 피사체가 의자에 앉은 자세로 물체를 양손에 각각 쥔 후 상하로 움직이는 동작을 반복한다(덤벨 숄더 프레스)	실외	1	4	2	1	2	3	12	576
28	28-15	의미론적	웨이트트레이닝	- 피사체가 서있는 자세에서 양손에 물체를 든 채로 허리를 숙인 후 앞뒤로 팔을 움직이는 동작을 반복한다(덤벨 킥백)	실외	1	4	2	1	2	3	12	576
29	29-1	의미론적	홈트레이닝	- 인기 홈트레이닝 top 1/5 따라하기	실내	1	3	2	1	2	3	30	1,080
29	29-2	의미론적	홈트레이닝	- 인기 홈트레이닝 top 2/5 따라하기	실내	1	3	2	1	2	3	30	1,080
29	29-3	의미론적	홈트레이닝	- 인기 홈트레이닝 top 3/5 따라하기	실내	1	3	2	1	2	3	30	1,080
29	29-4	의미론적	홈트레이닝	- 인기 홈트레이닝 top 4/5 따라하기	실내	1	3	2	1	2	3	30	1,080
29	29-5	의미론적	홈트레이닝	- 인기 홈트레이닝 top 5/5 따라하기	실내	1	3	2	1	2	3	30	1,080
30	30-1	의미론적	탁구(서브)	- 드라이브 서브를 한다	실내	1	4	2	1	2	3	12	576
30	30-2	의미론적	탁구(서브)	- 커트 서브를 한다.	실내	1	4	2	1	2	3	12	576
30	30-3	의미론적	탁구(서브)	- 너클 서브를 한다.	실내	1	4	2	1	2	3	12	576
31	31-1	의미론적	탁구(리시브)	- 날아오는 탁구공을 포핸드 스윙으로 받아친다.	실내	1	4	2	1	2	3	12	576
31	31-2	의미론적	탁구(리시브)	- 날아오는 탁구공을 백핸드 스윙으로 받아친다.(0:15~0:20)	실내	1	4	2	1	2	3	12	576
31	31-3	의미론적	탁구(리시브)	- 피사체가 날아오는 탁구공을 커트 스윙으로 받아친다.(01:30~1:40)	실내	1	4	2	1	2	3	12	576
32	32-1	의미론적	야구(투구)	- 오버핸드 방식으로 공을 던진다.(우완/좌완)	실내	1	4	2	2	2	3	12	1,152
32	32-2	의미론적	야구(투구)	- 언더핸드 방식으로 공을 던진다.(0:58~1:00)(우완/좌완)	실내	1	4	2	2	2	3	12	1,152
33	33-1	의미론적	야구(배팅)	- 피사체가 양손으로 야구 배트를 잡고 휘두른다(우완/좌완)	실내	1	4	2	2	2	3	12	1,152
33	33-2	의미론적	야구(배팅)	- 피사체가 배트를 짧게 잡고 번트 자세를 취한다(우완/좌완)	실내	1	4	2	2	2	3	12	1,152
34	34-1	의미론적	축구(드리블)	- 피사체가 가슴을 이용하여 공을 받는다	실내	1	4	2	1	2	3	12	576

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
34	34-2	의미론적	축구(드리블)	- 왼발/오른발 패스	실내	1	4	2	1	2	3	12	576
34	34-3	의미론적	축구(드리블)	- 2명이서 서로 패스한다.	실내	2	3	2	1	2	2	12	288
34	34-4	의미론적	축구(드리블)	- 2명 중 1명이 드리블을 하고 나머지 한명이 수비를 한다. 수비를 제치고 지나감	실내	2	3	2	1	2	2	12	288
34	34-5	의미론적	축구(드리블)	- 2명 중 1명이 드리블을 하고 나머지 한명이 수비를 한다. 수비에게 공을 뺏김	실내	2	3	2	1	2	2	12	288
34	34-6	의미론적	축구(드리블)	- 2명이 같은 방향으로 동시에 드리블 한다	실내	2	3	2	1	2	2	12	288
34	34-7	의미론적	축구(드리블)	- 2명이 서로 다른 방향으로동시에 드리블을 한다.	실내	2	3	2	1	2	2	12	288
34	34-8	의미론적	축구(드리블)	- 3명이 각각 공을 가지고 서로 다른 방향으로 드리블 하기	실내	3	3	2	1	2	3	12	432
34	34-9	의미론적	축구(드리블)	- 한명이 공을 가지고 드리블 하고 나머지 - 2명이 수비 하기	실내	3	3	2	1	2	3	12	432
34	34-10	의미론적	축구(드리블)	- 3명이 공하나를 서로에게 패스하기	실내	3	3	2	1	2	3	12	432
34	34-11	의미론적	축구(드리블)	- 4명이 각각 공을 가지고 서로 다른 방향으로 드리블 하기	실내	4	3	2	1	2	3	12	432
34	34-12	의미론적	축구(드리블)	- 한명이 공을 가지고 드리블 하다가 다른 한명에게 패스 하고 나머지 - 2명이 수비하기	실내	4	3	2	1	2	3	12	432
34	34-13	의미론적	축구(드리블)	- 한명이 공을 가지고 드리블 하고 나머지 - 3명이 수비 하기	실외	4	3	2	1	2	3	12	432
34	34-14	의미론적	축구(드리블)	- 4명이 공하나를 서로에게 패스하기	실외	4	3	2	1	2	3	12	432
35	35-1	의미론적	축구(슛팅)	- 정상 슛팅 - 오른발/왼발	실외	1	4	2	1	2	3	12	576
35	35-2	의미론적	축구(슛팅)	- 밀어차기 - 오른발/왼발	실외	1	4	2	1	2	3	12	576
35	35-3	의미론적	축구(슛팅)	- 찍어차기 - 오른발/왼발	실외	1	4	2	1	2	3	12	576
36	36-1	의미론적	골프(스윙)	- 피사체가 골프채를 이용해 골프 스윙을 하여 골프공을 친다	실내	1	5	2	1	2	6	12	1,440
36	36-2	의미론적	골프(스윙)	- 피사체가 자세를 잡으며 스윙전에 골프채를 골프공에 두세번 갖다 대며 준비자세를 취한다(어드레스)	실내	1	5	2	1	2	6	12	1,440
37	37-1	의미론적	골프(퍼팅)	- 피사체가 골프채를 이용해 퍼팅을 하여 골프공을 친다	실내	1	5	2	1	2	5	12	1,200
37	37-2	의미론적	골프(퍼팅)	- 피사체가 골프채를 세워잡은채로 앉는다	실내	1	5	2	1	2	5	12	1,200
38	38-1	의미론적	배드민턴(서브)	- 피사체가 셔틀콕을 위쪽으로 던지고 라켓으로 친다.	실내	1	5	2	1	2	5	12	1,200
38	38-2	의미론적	배드민턴(서브)	- 피사체가 라켓으로 셔틀콕을 아래서부터 위로 올려친다	실내	1	5	2	1	2	5	12	1,200
38	38-3	의미론적	배드민턴(서브)	- 피사체가 라켓을 뒤집어 셔틀콕을 아래서부터 위로 올려친다	실내	1	5	2	1	2	5	12	1,200
39	39-1	의미론적	배드민턴(리시브)	- 피사체가 날아오는 셔틀콕을 라켓으로 받아 친다.	실내	1	5	2	1	2	5	12	1,200
39	39-2	의미론적	배드민턴(리시브)	- 피사체가 날아오는 셔틀콕을 라켓을 뒤집어 받아 친다.	실내	1	5	2	1	2	5	12	1,200
39	39-3	의미론적	배드민턴(리시브)	- 피사체가 바닥에 떨어져있는 셔틀콕을 집어든다	실내	1	5	2	1	2	5	12	1,200
40	40-1	의미론적	줄넘기	- 모둠발뛰기는 양발 모아뛰기라고도 불리는 동작으로 1회선 1도약 / 1회선 2도약 합니다.	실내	1	4	2	1	2	3	12	576

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
40	40-2	의미론적	줄넘기	- 발바꿔뛰기 : 양손으로 줄을 잡은 후, 시선은 정면으로 하고 조깅 하듯 가볍게 줄을 돌리면서 한 발씩 넘어야 해요.	실내	1	4	2	1	2	3	12	576
40	40-3	의미론적	줄넘기	- 2단 3단 뛰기	실내	1	4	2	1	2	3	12	576
41	41-1	의미론적	음료 마시기	- 서서 - 피사체가 왼손으로 음료를 들고 마신다	실내	1	4	2	1	2	1	12	192
41	41-2	의미론적	음료마시기	- 서서 - 피사체가 빨대를 꽂은채로 왼손으로 음료수를 마신다	실내	1	4	2	1	2	1	12	192
41	41-3	의미론적	음료 마시기	- 서서 - 피사체가 두손을 이용하여 음료수를 마신다	실내	1	4	2	1	2	1	12	192
41	41-4	의미론적	음료 마시기	- 서서 - 피사체가 오른손으로 음료를 들고 마신다	실내	1	4	2	1	2	1	12	192
41	41-5	의미론적	음료 마시기	- 서서 - 피사체가 빨대를 꽂은채로 오른손으로 음료수를 마신다	실내	1	4	2	1	2	1	12	192
41	41-6	의미론적	음료 마시기	- 앉아서 - 피사체가 두손을 이용하여 음료수를 마신다	실내	1	4	2	1	2	1	12	192
41	41-7	의미론적	음료 마시기	- 앉아서 - 피사체가 왼손으로 음료를 들고 마신다	실내	1	4	2	1	2	1	12	192
41	41-8	의미론적	음료 마시기	- 앉아서 - 피사체가 오른손으로 음료를 들고 마신다	실내	1	4	2	1	2	1	12	192
41	41-9	의미론적	음료 마시기	- 앉아서 - 피사체가 빨대를 꽂은채로 왼손으로 음료수를 마신다	실외	1	4	2	1	2	1	12	192
41	41-10	의미론적	음료 마시기	- 앉아서 - 피사체가 빨대를 꽂은채로 오른손으로 음료수를 마신다	실외	1	4	2	1	2	1	12	192
42	42-1	의미론적	핸드폰 조작	- 피사체가 핸드폰을 응시하며 두손으로 조작한다	실내	1	4	2	1	2	1	12	192
42	42-2	의미론적	핸드폰 조작	- 피사체가 핸드폰으로 게임을 한다	실내	1	4	2	1	2	1	12	192
42	42-3	의미론적	핸드폰 조작	- 피사체가 핸드폰을 응시하며 왼손으로 조작한다	실내	1	4	2	1	2	1	12	192
42	42-4	의미론적	핸드폰 조작	- 피사체가 핸드폰을 응시하며 오른손으로 조작한다	실내	1	4	2	1	2	1	12	192
42	42-5	의미론적	핸드폰 조작	- 피사체가 셀카봉을 이용하여 자신의 모습을 촬영한다	실내	1	4	2	1	2	1	12	192
42	42-6	의미론적	핸드폰 조작	- 피사체가 셀프카메라를 세로로 들고 자신의 모습을 촬영한다	실내	1	4	2	1	2	1	12	192
42	42-7	의미론적	핸드폰 조작	- 피사체가 셀프카메라를 가로로 들고 자신의 모습을 촬영한다	실내	1	4	2	1	2	1	12	192
42	42-8	의미론적	핸드폰 조작	- 피사체가 휴대폰을 세로로 들고 다른 - 피사체를 촬영한다	실외	1	4	2	1	2	1	12	192
42	42-9	의미론적	핸드폰 조작	- 피사체가 휴대폰을 가로로 들고 다른 - 피사체를 촬영한다	실외	1	4	2	1	2	1	12	192
43	43-1	의미론적	통화	- 피사체가 핸드폰을 귀에대고 통화한다.	실내	1	4	2	2	2	2	12	768
43	43-2	의미론적	통화	- 피사체가 핸드폰을 이어폰을 낀 채로 통화한다.	실내	1	4	2	2	2	2	12	768
43	43-3	의미론적	통화	- 피사체가 핸드폰을 스피커폰으로 통화한다	실내	1	4	2	2	2	2	12	768
43	43-4	의미론적	통화	- 피사체가 핸드폰을 블루투스이어폰을 낀 채로 통화한다.	실외	1	4	2	2	2	2	12	768
44	44-1	의미론적	물건 운반	- 피사체가 물건을 들고 이동한다	실내	1	4	2	2	2	2	12	768
44	44-2	의미론적	물건 운반	- 수레/카트에 짐을 실어서 밀면서 운반한다.	실내	1	4	2	2	2	2	12	768

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
44	44-3	의미론적	물건 운반	- 피사체가 물건을 손으로 밀어 운반한다	실내	1	4	2	2	2	2	12	768
44	44-4	의미론적	물건 운반	- 서로 같은 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-큰 크기 상자)	실내	2	3	2	1	2	2	12	288
44	44-5	의미론적	물건 운반	- 서로 같은 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-중간 크기 상자)	실내	2	3	2	1	2	2	12	288
44	44-6	의미론적	물건 운반	- 서로 같은 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-작은 크기 상자)	실내	2	3	2	1	2	2	12	288
44	44-7	의미론적	물건 운반	- 서로 다른 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-큰 크기 상자)	실내	2	3	2	1	2	2	12	288
44	44-8	의미론적	물건 운반	- 서로 다른 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-중간 크기 상자)	실외	2	3	2	1	2	2	12	288
44	44-9	의미론적	물건 운반	- 서로 다른 방향으로 - 2명이 각각 물건을 나른다. (큰 크기 상자-작은 크기 상자)	실외	2	3	2	1	2	2	12	288
45	45-1	의미론적	담배 피우기	- 피사체가 서있는 상태로 왼손으로 담배를 피운다	실내	1	5	2	1	2	5	12	1,200
45	45-2	의미론적	담배 피우기	- 피사체가 서있는 상태로 오른손으로 담배를 피운다	실내	1	5	2	1	2	5	12	1,200
45	45-3	의미론적	담배 피우기	- 피사체가 앉아있는 상태로 왼손으로 담배를 피운다	실내	1	5	2	1	2	5	12	1,200
45	45-4	의미론적	담배 피우기	- 피사체가 앉아있는 상태로 오른손으로 담배를 피운다	실내	1	5	2	1	2	6	12	1,440
45	45-5	의미론적	담배 피우기	- 피사체가 이동하는 상태로 왼손으로 담배를 피운다	실외	1	5	2	1	2	8	12	1,920
45	45-6	의미론적	담배 피우기	- 피사체가 이동하는 상태로 오른손으로 담배를 피운다	실외	1	5	2	1	2	8	12	1,920
46	46-1	의미론적	사진 찍기	- 피사체가 서있는 상태로 세로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-2	의미론적	사진 찍기	- 피사체가 서있는 상태로 가로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-3	의미론적	사진 찍기	- 피사체가 서있는 상태로 세로로 사진기를 카메라삼발이에 놓고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-4	의미론적	사진 찍기	- 피사체가 서있는 상태로 가로로 사진기를 카메라삼발이에 놓고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-5	의미론적	사진 찍기	- 피사체가 앉아있는 상태로 세로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-6	의미론적	사진 찍기	- 피사체가 앉아있는 상태로 가로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-7	의미론적	사진 찍기	- 피사체가 누워있는 상태로 세로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
46	46-8	의미론적	사진 찍기	- 피사체가 누워있는 상태로 가로로 사진기를 들고 사진을 찍는다	실내	1	4	2	2	2	1	12	384
47	47-1	의미론적	농구	- 1점 자유투를 던진다(왼손/오른손)	실내	1	4	2	1	2	2	12	384
47	47-2	의미론적	농구	- 2점 구역에서 슛을 쏜다(왼손/오른손)	실내	1	4	2	1	2	2	12	384
47	47-3	의미론적	농구	- 3점 라인에서 점프하며 슛을 쏜다(왼손/오른손)	실내	1	4	2	1	2	2	12	384
47	47-4	의미론적	농구	- 달려가며 레이업 슛을 쏜다(왼손/오른손)	실내	1	4	2	1	2	2	12	384
47	47-5	의미론적	농구	- 덩크 슛을 한다.(왼손/오른손)	실내	1	4	2	1	2	2	12	384

번호	시나리오 번호	동작 분류	항목	시나리오	실내/실외	액터 수	촬영조	의상 수	소품 수	조도	반복 촬영 수	카메라 수	영상 클립 수
47	47-6	의미론적	농구	- 단거리 패스를 한다	실내	1	4	2	1	2	2	12	384
47	47-7	의미론적	농구	- 중거리 패스를 한다	실내	1	4	2	1	2	2	12	384
48	48-1	의미론적	역도 (물건 들어 올리기)	- 피사체가 양 발이 바닥에 닿은 채로 무릎을 굽혀 물건을 집는다 → 다시 무릎을 펴 물건을 들어올린다	실내	1	4	2	1	2	3	12	576
48	48-2	의미론적	역도 (물건 들어 올리기)	- 피사체가 물건의 한쪽만 들어올린다	실내	1	4	2	1	2	3	12	576
48	48-3	의미론적	역도 (물건 들어 올리기)	- 피사체가 물건을 어깨에서 머리위로 들어올린다	실내	1	4	2	1	2	3	12	576
49	49-1	의미론적	펜싱 (찌르기)	- 다리를 벌리고 무릎을 굽혀 자세를 낮추고 팔을 뻗어 상대방을 찌른다	실내	1	4	2	1	2	3	12	576
49	49-2	의미론적	펜싱 (찌르기)	- 서서 상대의 몸통을 찌르거나 팔을 벤다	실내	1	4	2	1	2	3	12	576
49	49-3	의미론적	펜싱 (찌르기)	- 뛰어서 당대의 몸통을 찌른다	실내	1	4	2	1	2	3	12	576
50	50-1	의미론적	집안일	- 청소 - 쓸기 (큰 빗자루, 중간 빗자루, 작은 빗자루)	실내	1	4	2	2	1	3	12	576
50	50-2	의미론적	집안일	- 청소 - 큰 청소기 사용	실내	1	4	2	2	1	3	12	576
50	50-3	의미론적	집안일	- 청소 - 핸드 청소기 사용 (상 중 하 높이 청소)	실내	1	4	2	2	1	3	12	576
50	50-4	의미론적	집안일	- 청소 - 걸레로 닦기(상 중 하 높이 청소)	실내	1	4	2	2	1	3	12	576
50	50-5	의미론적	집안일	- 빨래 - 세탁기에서 빼기	실내	1	4	2	2	1	3	12	576
50	50-6	의미론적	집안일	- 빨래 - 세탁기에 넣기	실내	1	4	2	2	1	3	12	576
50	50-7	의미론적	집안일	- 빨래 - 널기	실내	1	4	2	2	1	3	12	576
50	50-8	의미론적	집안일	- 빨래 - 먼지 털기	실내	1	4	2	2	1	3	12	576
50	50-9	의미론적	집안일	- 빨래 - 개어넣기	실내	1	4	2	2	1	3	12	576
50	50-10	의미론적	집안일	- 설거지 - 식기세척기에 넣기	실내	1	4	2	2	1	3	12	576
50	50-11	의미론적	집안일	- 설거지 - 식기세척기에서 빼기	실내	1	4	2	2	1	3	12	576
50	50-12	의미론적	집안일	- 설거지 - 손으로 하기	실내	1	4	2	2	1	3	12	576
50	50-13	의미론적	집안일	- 설거지 - 식기 정리하기	실내	1	4	2	2	1	3	12	576
50	50-14	의미론적	집안일	- 요리 - 재료 칼로 썰기	실내	1	4	2	2	1	3	12	576
50	50-15	의미론적	집안일	- 요리 - 재료 볶기	실내	1	4	2	2	1	3	12	576
50	50-16	의미론적	집안일	- 요리 - 음식 무침	실내	1	4	2	2	1	3	12	576
합계													200,088

07

한국인 안면 이미지 AI 데이터 구축 가이드라인

한국과학기술연구원
Korea Institute of Science and Technology

	성명	소속	직위
가이드라인 제안	조정현	KIST	선임 연구원
가이드라인 작성	조정현	KIST	선임 연구원
	김학섭	KIST	선임 연구원
전문가 검토 위원	이상근	한양대학교	조교수
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 한국과학기술연구원 / 참여: 휴먼아이씨티, 에쓰큐아이소프트

서 문

1. 가이드라인의 목적

이 가이드라인은 다양한 안면 응용 분야의 인공지능 연구 및 개발을 위한 한국인 안면 이미지 데이터 구축 방법을 제공하기 위하여 작성되었다. 인공지능 기술 개발을 위해서는 학습용 안면 이미지 데이터가 필수적으로 필요하며, 효과적인 데이터 활용을 위해서는 데이터의 획득, 가공, 공유, 활용에 대한 세부 지침이 필요하다고 할 수 있다. 이에 본 가이드라인은 한국인 안면 이미지 데이터셋의 폭넓은 활용 및 데이터의 증식을 목적으로 작성되었다.

2. 주요 내용 요약

딥러닝 기반의 안면 응용 기술의 발전을 위해서는 대규모의 정제되고, 색인된 안면 이미지 데이터셋이 반드시 필요한 실정이며, 기술 고도화를 위해서는 다양한 환경이 고려된 데이터셋이 필요하다. 이에 본 가이드라인은 아래와 같은 환경을 정의하고, 이에 대한 안면 이미지 데이터셋 구축 세부 지침을 제공한다.

- 다양한 포즈, 각도, 조도, 해상도에 따른 안면 이미지 데이터셋
- 균일하게 분포된 성별/연령별 안면 이미지 데이터셋
- 열악한 환경 (저조도, 액세서리, 저해상도 등)이 고려된 데이터셋
- 국내 상황에 맞는 알고리즘 개발을 위한 한국인 안면 이미지 데이터셋
- 직관적인 색인 및 정제되고, 유효성이 검증된 데이터셋

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

한국인 안면 이미지 데이터 구축 지침

Guideline for Constructing a Large-scale Dataset of Korean Face Images

07

1. 적용 범위

해당 가이드라인은 인공지능 학습용 한국인 안면 이미지 데이터셋 구축 지침을 제시하도록 한다. 먼저 본 안면 이미지 데이터셋 구축을 위해 획득되는 원천 데이터의 구조 및 정제 방법에 대한 세부 지침을 제공하며, 구체적인 저장 방법을 제시한다. 뿐만아니라, 데이터셋 구축을 위한 촬영 장비의 설계 및 구축에 대한 기준을 제공하여 추후 데이터 증식을 위한 데이터 추가 구축 시, 기 구축된 데이터셋과의 호환성을 유지할 수 있도록 한다. 또한 인공지능 학습을 위해 본 데이터셋이 활용될 수 있도록 활용 지침을 제공하도록 하여, 인공지능 기반 안면 응용 기술 분야에 본 데이터셋이 폭넓게 활용될 수 있도록 한다.

2. 인용 표준

- 해당 사항 없음

3. 용어 정의

3.1 포즈

– 한국인 안면 이미지 촬영 시의 카메라의 각도를 의미하며, 본 데이터 구축 가이드라인에서는 촬영 대상자의 코를 기준으로 수직각도 -15°~30°에서 수평각도 -90°~90°까지의 다양한 각도의 포즈가 존재한다.

3.2 API

– API (Application Programming Interface)는 응용 프로그램 프로그래밍 인터페이스로 응용 프로그램에서 사용할 수 있도록, 운영 체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있게 만든 인터페이스를 뜻한다.

3.3 NAS

— NAS (Network Attached Storage)는 네트워크 결합 저장장치로, 네트워크로 연결되는 외장 하드디스크로, 컴퓨터에 직접 연결하지 않고 네트워크를 통해 데이터를 주고 받는 저장장치를 의미한다.

3.4 Facial Landmark Position

— 촬영된 안면이미지를 활용한 가공데이터 중의 하나로, 얼굴 내의 특징점 및 바운딩 박스를 의미한다.

3.5 Cropping

— 이미지 내의 일부 영역을 잘라내는 것을 의미한다.

4. 약어

- API Application Programming Interface
- NAS Network Attached Storage

5. 인공지능 학습용 한국인 안면 이미지 데이터 구축

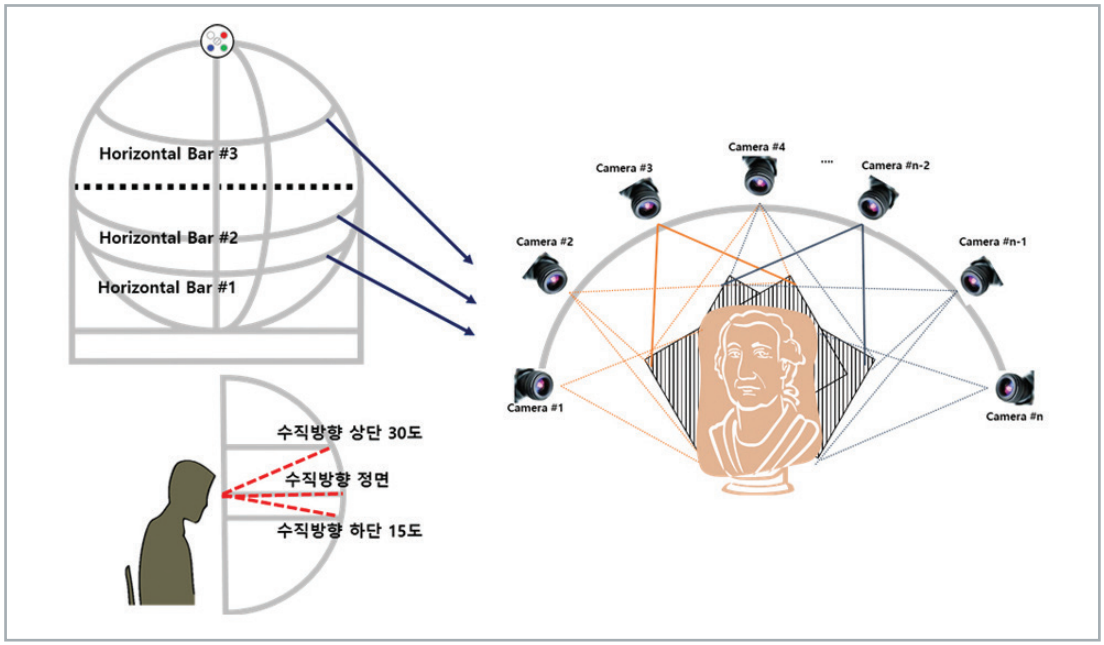
● 5.1 한국인 안면 이미지 데이터 수집

○ 5.1.1 원천데이터의 종류

- 한국인 안면 이미지 데이터 구축을 위한 원천데이터는 성별, 연령대별(20대~60대)로 한국인을 대상으로 촬영한 고화질 얼굴 영상 데이터라고 할 수 있다. 한국인 안면 이미지 원천데이터는 얼굴 포즈, 조명, 액세서리, 표정, 해상도별 조건을 부여하였다.

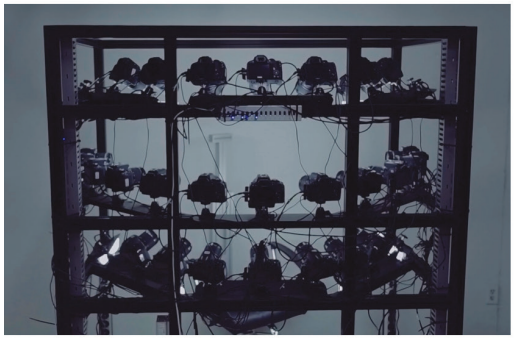
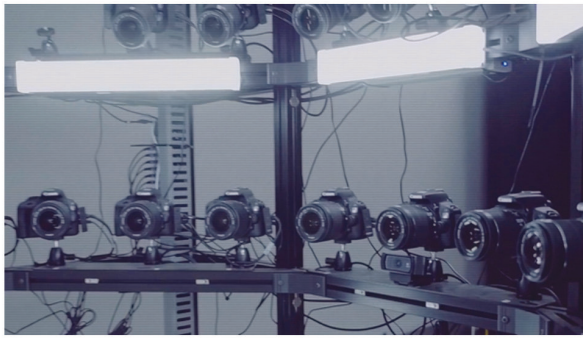
○ 5.1.2 촬영 장비 제작 및 운용

[그림 5.1-1] 한국인 안면 이미지 데이터셋 구축을 위한 촬영 장비 설계



- 상기 그림과 같이 영상 촬영 장비는 촬영 대상자를 중심으로 다수의 카메라를 포함하며, 카메라의 설치 위치는 촬영 대상자와 카메라 간의 각도 (수직 방향 상단 30도, 0도, 하단 15도 / 수평 방향 좌우 90도 범위)를 각기 달리함으로써, 단 한 번의 촬영으로 아래 그림과 같이 여러 포즈 변화를 포함하는 영상을 획득하도록 하였다.
- 포즈별 촬영을 위해서 카메라는 수직방향 0도 기준에서 수평방향(±90, ±75, ±60, ±45 ,±30, ±15, 0), 수직 상단 30도 기준에서 수평방향(±45, ±15, 0), 수직 하단 15도 기준에 수평방향(±30) 총 20대의 카메라를 설치하였다.
- 촬영 장비에는 다양한 조명 환경(0 Lux ~ 1000 Lux)을 제공하기 위해 조도를 제어할 수 있는 조명 제어 시스템을 포함하였다.

[그림 5.1-2] 한국인 안면 이미지 데이터셋 구축 촬영 장비

* 촬영 카메라 : Cannon EOS 100D
* 카메라 세팅 및 저장 방식
- Tv : 40, Av : 5.6, ISO : 800, Meteringmode : Evaluate
- Image Quality : Small1 Fine Jpeg
- 해상도 : 고해상도 (864x576), 중해상도 (346x230), 저해상도 (173x115)

○ 5.1.3 안면 이미지 데이터셋 구축

- 아래의 표와 같이 얼굴 포즈 20종, 조명 30종, 액세서리 6종, 표정 3종, 해상도 3종을 포함하여, 촬영 대상자 1명당 32,400장의 안면 이미지 데이터셋을 구축하도록 하였다.

[표 5.1-1] 안면 이미지 데이터셋 구성요소 및 상세 설명

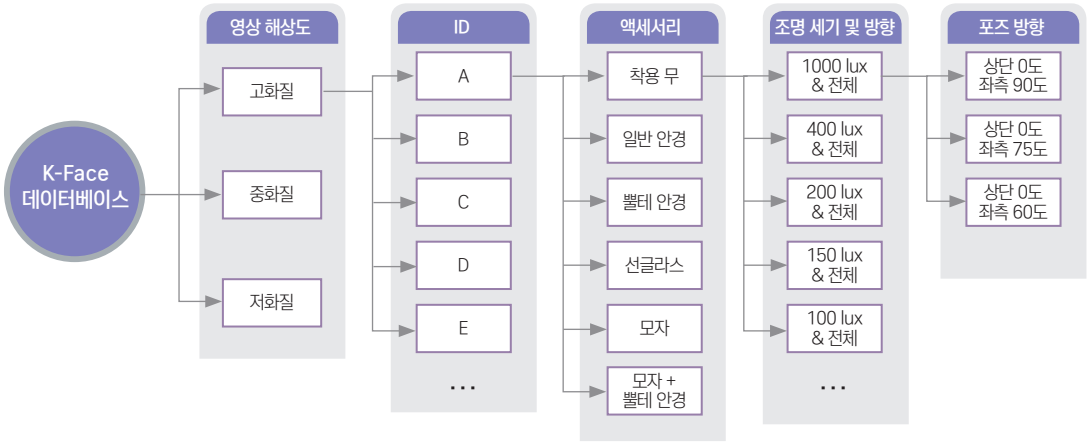
구성 요소		상세 설명
얼굴 촬영 시스템	반구 형태의 프레임	좌우 90도,상단 30도 / 하단 15도 커버
	최소 20대의 DSLR 및 webcam	Canon 700D급 렌즈 이상
	Auto illuminator	각도별, 조도별(Lux단위)조절 가능
얼굴 데이터	포즈	좌우 90도, 상단 30도/하단 15도 이내
	조명	날씨, 시간대, 실내외, 그림자 및 빛 반사 고려
	가림	일반 안경, 뿔테안경, 선글라스,모자
	표정	일반, 웃음, 찡그림
	해상도	고화질, 중화질, 저화질

[그림 5.1-3] 한국인 안면 이미지 데이터셋 구축 샘플 영상

○ 5.1.4 한국인 안면 이미지 데이터셋 구조 및 형태

- 효과적인 데이터 활용 및 데이터 증식을 위해 한국정보화진흥원과 한국과학기술연구원에서 구축한 K-Face와의 호환성을 유지하도록, 다양한 촬영환경과 연령분포를 갖도록 한국인 안면 이미지 데이터셋을 구축하도록 하였다.

[그림 5.1-4] 한국인 안면 이미지 데이터셋 구조



- 수직각도 -15°~ 30°에서 수평각도 -90°~ 90°까지 다양한 각도로 안면 이미지를 촬영하며, 1인당 얼굴 포즈 20종, 조명 30종, 액세서리 6종, 표정 3종, 해상도 3종을 포함하도록 하였다.

[그림 5.1-5] 한국인 안면 이미지 데이터셋 포즈 구조

방향		
	+	-
수직	상단	하단
수평	우측	좌측

Camera	방향	
	수직	수평
C1	0°	+90°
C2	0°	+75°
C3	0°	+60°
C4	0°	+45°
C5	0°	+30°
C6	0°	+15°
C7	0°	+0°
C8	0°	-15°
C9	0°	-30°
C10	0°	-45°
C11	0°	-60°
C12	0°	-75°
C13	0°	-90°
C14	+30°	+45°
C15	+30°	+15°
C16	+30°	0°
C17	+30°	-15°
C18	+30°	-45°
C19	-15°	+30°
C20	-15°	-30°

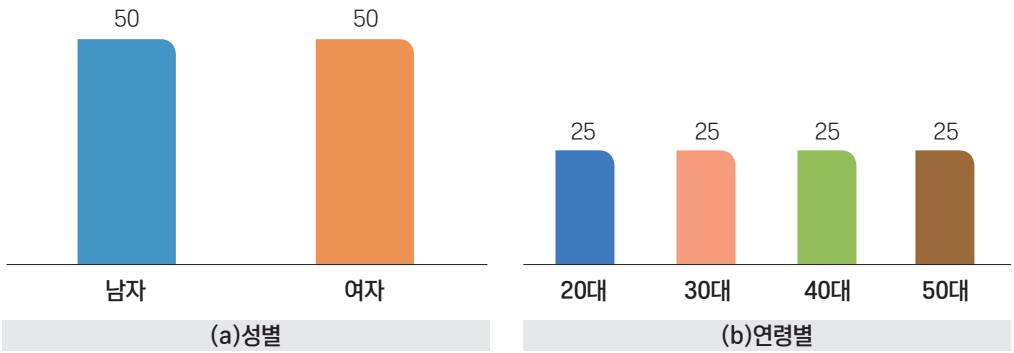
Light	Lux	방향	
		수직	수평
L1	1000	전체	
L2	400	전체	
L3	200	전체	
L4	150	전체	
L5	100	전체	
L6	40	전체	
L7	0	전체	
L8	400	+30°	전체
L9	200	+30°	전체
L10	100	+30°	전체
L11	40	+30°	전체
L12	400	-15°	전체
L13	200	-15°	전체
L14	100	-15°	전체
L15	40	-15°	전체
L16	400	전체	+90°
L17	200	전체	+90°
L18	100	전체	+90°
L19	400	전체	+45°
L20	200	전체	+45°
L21	100	전체	+45°
L22	400	전체	+0°
L23	200	전체	+0°
L24	100	전체	+0°
L25	400	전체	-45°
L26	200	전체	-45°
L27	100	전체	-45°
L28	400	전체	-90°
L29	200	전체	-90°
L30	100	전체	-90°

Accessory	
S001	보통
S002	일반 안경
S003	뿔테 안경
S004	선글라스
S005	모자
S006	모자+뿔테안경

Expression	
E01	무표정
E02	활짝웃음
E03	찡그림

- 연령대는 20대 ~ 50대로 다양한 연령대에 대한 이미지 촬영을 진행하며, 최종적으로 아래의 그림과 같이 남녀 및 연령분포가 균일한 한국인 안면 이미지 데이터셋을 구축하였다.

[그림 5.1-6] 한국인 안면이미지 데이터셋 분포



- 한국인 안면 이미지 데이터셋은 랜덤 샘플링을 통해 학습용:평가용 데이터셋을 8:2로 구분하여, 인공지능 학습에 활용이 용이하게 하였다.

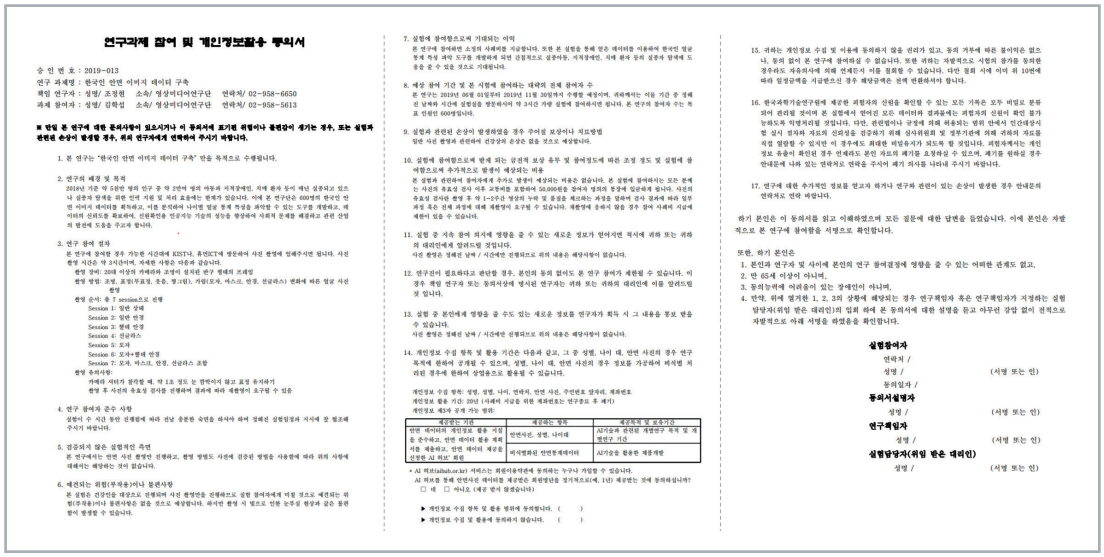
5.1.5 촬영 대상자 모집

- 촬영 대상자는 성별 및 연령대의 균형을 맞춰서 섭외하였으며, 인력채용 사이트(알바몬 등), 엑스트라 협회, 지자체 협조등을 통한 다양한 경로를 통해서 모집하였다. (개인정보 활용 동의 내용 포함)

[그림 5.1-7] 구인 사이트 활용 예시

- 모집된 촬영자는 반드시 개인정보 활용 동의서를 작성하도록 하며, 법적 의무사항을 준수할 수 있도록 개인정보 활용 동의서에 대한 IRB 심사 및 전문가 자문을 구하였다.

[그림 5.1-8] 개인정보 활용 동의서 예시



연구과제 참여 및 개인정보활용 동의서

승 인 번 호 : 2019-013
연구 과제명 : 한국인 안면 이미지 데이터 구축
책임 연구자 : 성명/ 조정현 소속/ 영상미디어연구단 연락처/ 02-958-6650
과제 참여자 : 성명/ 김학섭 소속/ 영상미디어연구단 연락처/ 02-958-5613

※ 만일 본 연구에 대한 문제사항이 있으시거나 이 동의서에 표기된 위험이나 불편감이 생기는 경우, 또는 실험과 관련된 손상이 발생할 경우, 위의 연구자에게 연락하여 주시기 바랍니다.

1. 본 연구는 "한국인 안면 이미지 데이터 구축" 만을 목적으로 수행됩니다.

2. 연구의 배경 및 목적
2018년 기준 약 5천만 명의 연구 중 약 2만여 명의 사람과 지적장애인, 치매 환자 등이 매년 실종되고 있으나 실종자 발생을 과한 인력 기원 및 치의 보조에는 한계가 있습니다. 이에 본 연구단은 400명의 한국인 안면 이미지 데이터를 확득하고, 이를 분석하여 나타낼 얼굴 특징을 파악할 수 있는 도구를 개발하고, 데이터의 신뢰도를 확보하여, 신원확인을 연상시킬 기술의 성능을 향상하여 사회적 문제를 해결하고 관련 산업의 발전에 도움을 주고자 합니다.

3. 연구 참여 절차
본 연구에 참여할 경우 가능한 시간대에 KISTI, 휴먼ICT에 방문하여 사진 촬영에 임해주시면 됩니다. 사진 촬영 시간은 약 3시간이며, 자세한 사항은 다음과 같습니다.
촬영 장비: 20대 이상의 카메라와 조명이 설치된 연구 형태의 조립실
촬영 방법: 조명, 표정(무표정, 웃음, 찡그림), 가림(모자, 마스크, 안경, 선글라스) 변화에 따른 일상 사진 촬영
촬영 순서: 총 7 session으로 진행
Session 1: 일반 안경
Session 2: 일반 안경
Session 3: 블랙 안경
Session 4: 선글라스
Session 5: 모자
Session 6: 모자+블랙 안경
Session 7: 모자, 마스크, 안경, 선글라스 조합
촬영 유의사항:
카메라 시야가 충만할 때, 약 1.2, 정도 는 잔뜩이거 많고 표정 유지하기
촬영 후 사진의 유효성 문제를 발견하여 재촬영이 요구될 수 있음

4. 연구 참여자 준수 사항
실험자 수 시간 동안 진행함에 따라 선상 유발한 욕만을 하거나 하여 정해진 실험일정과 시지에 꼭 협조해 주시기 바랍니다.

5. 검증되지 않은 실험적인 측면
본 연구에서는 안면 사진 촬영만 진행하고, 촬영 방법도 사진에 검증된 방법을 사용함에 따라 위의 사항에 대해서는 해당하는 것이 없습니다.

6. 예견되는 위험(부작용)이나 불편사항
본 실험은 건강인을 대상으로 진행되며 사진 촬영만을 진행하므로 실험 참여자에게 미칠 것으로 예견되는 위험(부작용)이나 불편사항은 없을 것으로 예상됩니다. 하지만 촬영 시 빛으로 인한 눈부심 현상과 같은 불편함이 발생할 수 있습니다.

7. 실험에 참여함으로써 기대되는 이익
본 연구에 참여하면 소정의 사례비를 지급합니다. 또한 본 실험을 통해 얻은 데이터를 이용하여 한국인 얼굴 특징 학습 위한 도구를 개발하게 되면 신뢰도로 공증하여, 지적장애인, 치매 환자 등의 실종자 발생에 도움을 줄 수 있을 것으로 기대됩니다.

8. 예상 참여 기간 및 본 시험에 참여하는 대략의 전체 참여자 수
본 연구는 2019년 04월 01일부터 2019년 11월 30일까지 수행될 예정이며, 위와 예시는 이들 기간 중 정해진 날짜와 시간에 실험실을 방문하시어 약 3시간 가량 실험에 참여하시면 됩니다. 본 연구의 참여자 수는 적요 단원인 600명입니다.

9. 실험자 관련된 손상이 발생하였을 경우 주어질 보상이나 치료방법
일반 사진 촬영과 관련하여 건강상의 손상은 없을 것으로 예상됩니다.

10. 실험에 참여함으로써 받게 되는 금전적 보상 유무 및 참여정도에 따른 조정 정도 및 실험에 참여함으로써 추가적으로 발생하게 예상되는 비용
본 실험과 관련하여 참여자에게 추가로 발생이 예상되는 비용은 없습니다. 본 실험에 참여하시는 모든 분에는 사진의 유효성 검사 시간 고려하여 보장하여 50,000원을 참여자 명의의 통장에 입금해 드립니다. 사진의 유효성 검사는 촬영 후 약 1~2주간 정상적 누락 및 훼손을 확인하는 과정을 입증하여 검사 결과에 따라 일부 혹은 전체 과정에 대해 재촬영이 요구될 수 있습니다. 재촬영에 응하지 않을 경우 참여 사례비 지급에 제한이 있을 수 있습니다.

11. 실험 중 지속 참여 의지에 영향을 줄 수 있는 새로운 정보가 연이어진 목시에 귀하 또는 귀하의 대리인에게 알려드릴 것입니다.
사진 촬영은 정해진 날짜 / 시간에만 진행되므로 위의 내용은 해당사항이 없습니다.

12. 연구진이 필요하다고 판단할 경우, 본인의 동의 없이도 본 연구 참여가 제한될 수 있습니다. 이 경우 해당 연구자 또는 동의서상에 명시된 연구자는 귀하 또는 귀하의 대리인에 이를 알려드릴 것 입니다.

13. 실험 중 본인에게 영향을 줄 수도 있는 새로운 정보를 연구자가 획득 시 그 내용을 통보 받을 수 있습니다.
사진 촬영은 정해진 날짜 / 시간에만 진행되므로 위의 내용은 해당사항이 없습니다.

14. 개인정보 수집 항목 및 활용 기간은 다음과 같고, 그 중 성별, 나이 대, 안면 사진의 경우 연구 목적에 필하여 공개될 수 있으며, 성별, 나이 대, 안면 사진의 경우 정보를 가공하여 비식별 처리된 경우에 한하여 상업용으로 활용될 수 있습니다.

개인정보 수집 항목: 성명, 성별, 나이, 연락처, 안면 사진, 주민번호 앞자리, 계좌번호
개인정보 활용 기간: 20년 (사례비 지급을 위한 계좌번호는 연구종료 후 폐기)
개인정보 제3자 공개 가능 범위:

개인정보 기관	제공받는 범위	제공목적 및 보유기간
안면 데이터의 개인정보 활용 기관을 준수하고, 안면 데이터 활용 계획 서를 제출하고, 안면 데이터 제공을 신청한 AI 학교" 회원	안면사진, 성별, 나이대	AI기술과 관련된 개발연구 목적 및 개발연구 기간
	비식별화된 안면정제데이터	AI기술을 활용한 제품개발

* AI 학교(aihub.or.kr) 사이트는 회원이용약관에 동의하는 누구나 가입할 수 있습니다.
AI 학교를 통해 안면사진 데이터를 제공받은 회원명단을 정기적으로(예, 1년) 제공받는 것에 동의하십니까?
☐ 네 ☐ 아니오 (제공 받지 않겠습니다)

▶ 개인정보 수집 항목 및 활용 방법에 동의합니다. ( )
▶ 개인정보 수집 및 활용에 동의하지 않습니다. ( )

15. 귀하는 개인정보 수집 및 사용에 동의하지 않을 권리가 있고, 동의 거부에 따른 불이익은 없으나, 동의 없이 본 연구에 참여하실 수 없습니다. 또한 귀하는 자발적으로 시험의 참가를 동의한 경우라도 사유의 이에 연계되어 이들 철회할 수 있습니다. 다만 철회 시에 이미 위 10번에 따라 일정금액을 지급받으신 경우 해당금액은 전액 반환하셔야 합니다.

16. 한국과학기술연구원에 제공한 피험자의 신원을 확인할 수 있는 모든 기록은 모두 비밀로 분류되어 관리될 것이며 본 실험에서 얻어진 모든 데이터의 결과물에는 피험자의 신원이 확인 불가능하도록 익명처리되될 것입니다. 다만, 관련법령나는 규정에 의해 허용되는 범위 안에서 인간대상시험 심사 위원회 자료의 신뢰성을 검증하기 위해 심사위원회 및 정부기관에 의해 귀하의 자료를 직접 열람할 수 있지만 이 경우에도 해당한 비밀유지가 되도록 할 것입니다. 피험자에서는 개인 정보 유출이 확인된 경우 연세대도 본인 자료의 폐기를 요청하실 수 있으며, 폐기를 원하실 경우 안내문에 나와 있는 연락처로 연락을 주시어 폐기 의사를 나타내 주시기 바랍니다.

17. 연구에 대한 추가적인 정보를 얻고자 하거나 연구자 관련이 있는 손상이 발생한 경우 안내문의 연락처로 연락 바랍니다.

하기 본인은 이 동의서를 읽고 이해하였으며 모든 질문에 대한 답변을 들었습니다. 이에 본인은 자발적으로 본 연구에 참여함을 서명으로 확인합니다.

또한, 하기 본인은
1. 본인과 연구자 및 사이에 본인의 연구 참여결정에 영향을 줄 수 있는 어떠한 관계도 없고,
2. 만 65세 이상이 아니며,
3. 동의능력에 어려움이 있는 장애인이 아니며,
4. 만약, 위에 열거한 1, 2, 3의 상황에 해당되는 경우 연구책임자 혹은 연구책임자가 지정하는 실험 담당자(위임 받은 대리인)의 입회 하에 본 동의서에 대한 설명을 듣고 아무런 강압 없이 전적으로 자발적으로 아래 서명을 하였음을 확인합니다.

실험참여자
연락처 /
성명 / (서명 또는 인)
동의일자 /
동의서설명자
성명 / (서명 또는 인)
연구책임자
성명 / (서명 또는 인)
실험담당자(위임 받은 대리인)
성명 / (서명 또는 인)

5.2 한국인 안면 이미지 데이터 정제

5.2.1 원천데이터의 확보 및 정제

- 다수의 성별, 연령별 촬영 대상자를 모집하여, 사람 별로 32,400 장의 영상을 촬영하도록 하며, 집단 지성을 활용하여 촬영 대상자 및 영상 정보 분류를 진행하였다.

- 2단계에 수작업 검수를 통한 사용불가 사진 제거 및 미촬영 사진에 대한 즉각 추가 촬영을 진행하였으며, 데이터베이스 구축 요원의 반복적 작업으로 인한 업무 효율성 저하를 제거하기 위해서 촬영 일정에 따른 휴식 시간을 지정하여 데이터 품질 및 업무 효율성을 보장하도록 하였다.

○ 5.2.2 원천데이터의 자료형태 및 규모

- 촬영 대상자 개인별로 20개의 포즈 조건, 30개의 조명 조건, 6개의 액세서리 조건, 3개의 표정 조건 및 3개의 해상도 조건을 고려하여 총 인당 32,400장 (20 × 30 × 6 × 3 × 3)을 구축하도록 하였다.
- 구축하고자 하는 얼굴 영상 데이터베이스는 얼굴 인식 성능의 고도화뿐만 아니라, 얼굴 대상자 모델링 등 다양한 얼굴 관련 분야에서의 활용성을 높이는 것을 목적으로 하며, 이에 따라 데이터베이스 구조는 ID, 얼굴 영역 정보, 포즈, 조명 위치 및 세기, 액세서리 유무, 영상 해상도가 포함되도록 하였다.
- 촬영 대상자로부터 획득한 32,400장의 영상을 촬영 조건에 따라 자동으로 데이터베이스 구조에 맞추어 구성된 디렉토리에 저장함으로써, 작업 시간 및 영상 정리 작업에 대한 효율성 증대를 이루도록 하였다.

● 5.3 한국인 안면 이미지 데이터 가공

○ 5.3.1 안면 이미지의 Facial Landmark Position 생성

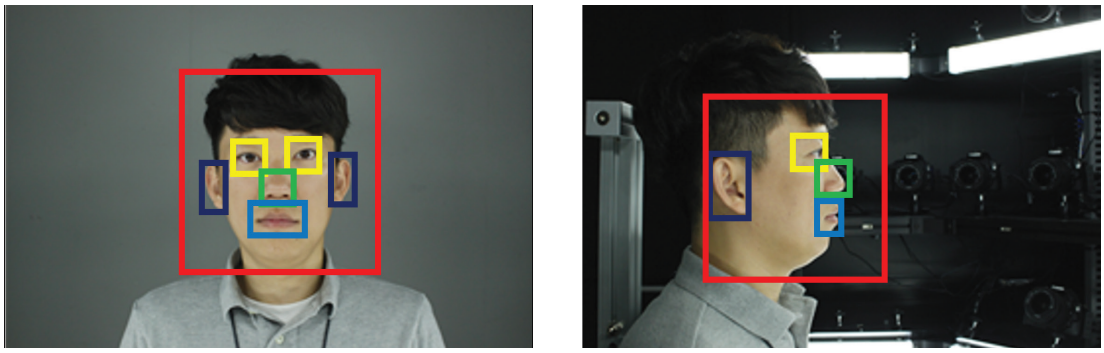
- 안면 이미지의 효율적 활용을 위한 Facial Landmark Position을 아래와 같이 추출하였다.

[그림 5.3-1] 안면 이미지의 특징점 예시

* 특징점 (총 7개)

0 : 코 끝
1 : 오른쪽 눈 중심
2 : 왼쪽 눈 중심
3 : 오른쪽 입꼬리
4 : 왼쪽 입꼬리
5 : 오른쪽 귀 중심
6 : 왼쪽 귀 중심

[그림 5.3-2] 안면 이미지의 바운딩 박스 예시



* 바운딩박스 (총 5종류)
  얼굴, 눈, 코, 입, 귀 영역에 대한 바운딩 박스

○ 5.3.2 고해상도 안면이미지를 이용한 중/저해상도 이미지 자동 생성 도구 개발

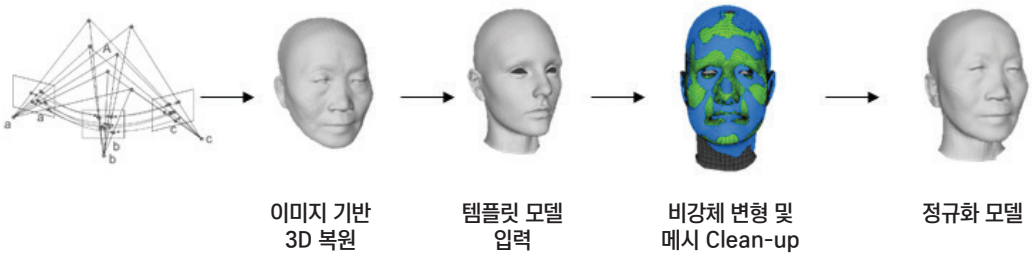
- 2차 검수가 완료된 고해상도 이미지를 대상으로 중해상도와 저해상도 이미지를 각각 자동으로 생성하도록 하였다.
- 고해상도 이미지를 보다 낮은 해상도로 변환 및 생성하는 과정에서 해상도 변화에 대한 비율을 이용하여 Facial Landmark Position 정보를 생성하도록 하였다.
- 단, 해상도를 낮춰 새롭게 생성된 이미지의 가로, 세로 비율은 유지하며, 해상도가 변경된 이미지와 Facial Landmark Position 정보는 완료 이미지 보관용 NAS에 저장하도록 하였다.

  * 고해상도 (864x576), 중해상도 (346x230), 저해상도 (173x115)

○ 5.3.3 안면 이미지의 3D 모델 데이터 생성

- 구축된 안면 이미지의 폭넓은 활용을 위하여, 구축된 안면 이미지에 대한 3D 모델 데이터 샘플 생성을 고려하도록 하였다.

[그림 5.3-3] 정규화된 3D 모델 데이터 생성 흐름



- 구축된 안면 이미지 데이터는 하나의 피사체에 대한 다양한 각도의 이미지를 포함하고 있으므로, 이미지 기반 3D 복원 및 이를 템플릿 모델에 적용, 보정하는 방식으로 정규화된 3D 모델 데이터를 생성하고자 하였다.

- 3D 모델 데이터는 활용 목적에 적합한 안면 이미지 데이터에 대해서 선택적으로 구축하였으며, 이는 다양한 안면 응용 기술에 활용될 수 있을 것으로 예상된다.

[그림 5.3-4] 정규화된 3D 모델 데이터 예시

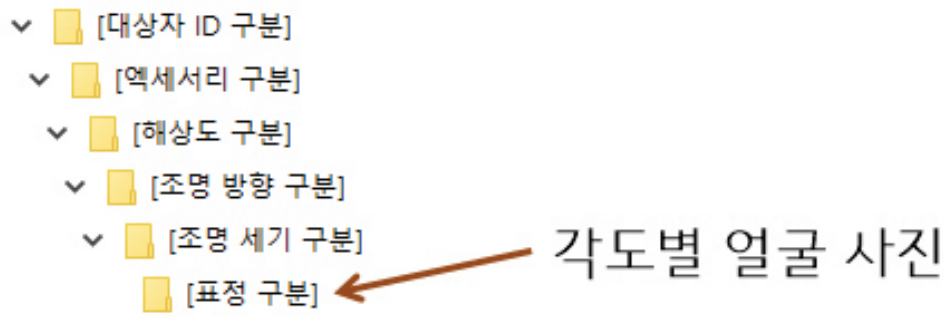
안면 이미지의 3D 모델 예시 정규화된 3차원 메시 데이터

5.4 한국인 안면 이미지 데이터 검수

5.4.1 한국인 안면 이미지 데이터베이스 구축 작업 규칙

- 촬영된 안면 이미지는 촬영 당시 1차 확인을 거치고, 검수 팀에서 2차 확인을 실시하도록 하였으며, 촬영 조건이 일치하지 않는 경우에는 재촬영을 실시하도록 하였다.
- 얼굴 영상 데이터베이스의 자료 형태는 아래와 같이 디렉터리로 구분하여 파일을 저장하도록 하였다.

[그림 5.4-1] 한국인 안면 이미지 데이터셋 디렉토리 구조

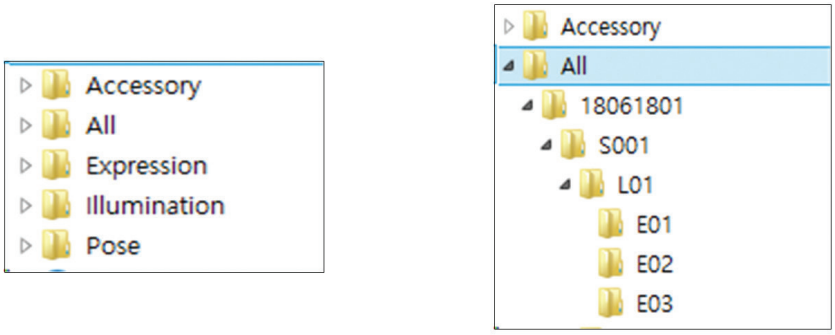


① 촬영 장비를 이용하여 각 조건별로 모든 각도의 얼굴을 촬영
② 최종적으로 저장되는 파일은 JPG 형식으로 저장
③ 표정으로 구분된 디렉터리에 각도별 이미지 저장

- 각 Case 별로 디렉터리로 구분하기 때문에 별도의 Labeling을 하지 않았다.

* 구분: 대상자 ID, 액세서리, 해상도, 조명 방향, 조명 세기, 표정, 얼굴 각도 등
* 구분별 디렉토리명은 (그림 5.1-5)의 각 요소별 코드를 활용하였다.

[그림 5.4-2] (a) 분류 예시, (b) 상세구분 예시

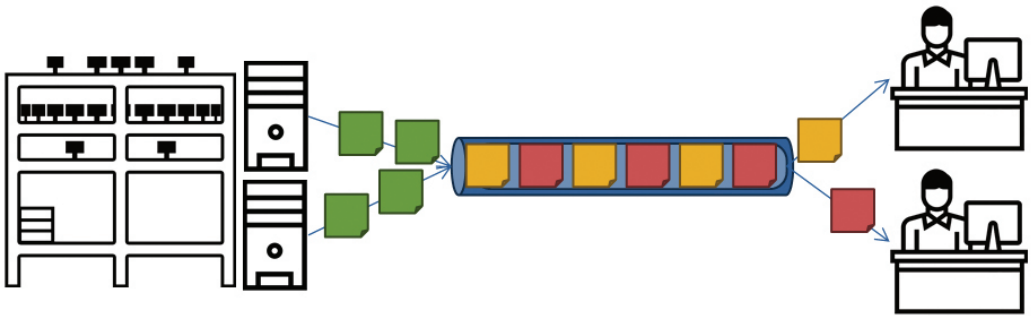


(a) (b)

○ 5.4.2 데이터베이스 구축 효율화를 위한 관리도구 활용

- 촬영된 안면 이미지의 누락, 번짐 및 조건오류를 실시간으로 점검하기 위한 도구를 개발하여 활용하였다.

[그림 5.4-3] 데이터셋 검수 이미지 자동 할당 예시

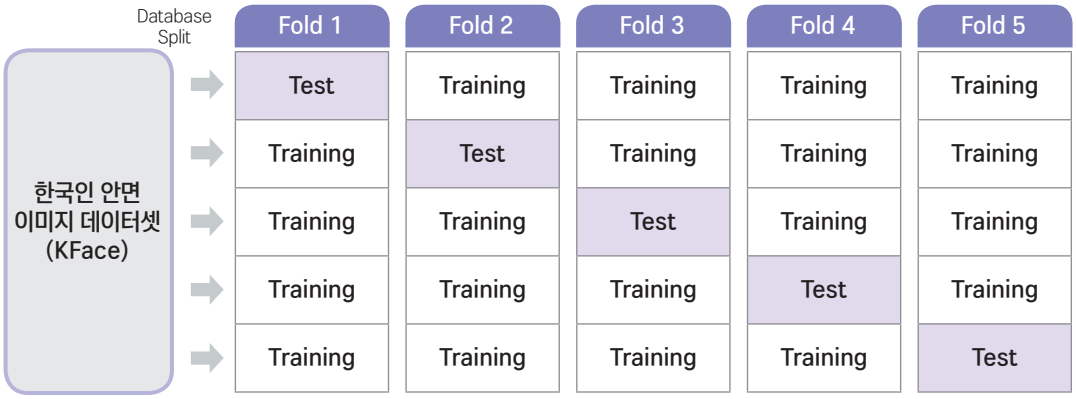


- 실시간으로 모든 촬영된 이미지에 대한 전수 검수를 실시하며, 자동으로 할당되는 서로 다른 이미지를 검수함으로써 중복 없이 빠른 검수가 가능하게 하고, 검수가 완료된 이미지는 데이터베이스 저장용 NAS에 저장하도록 하였다.

○ 5.4.3 한국인 안면 이미지 데이터셋 유효성 검증

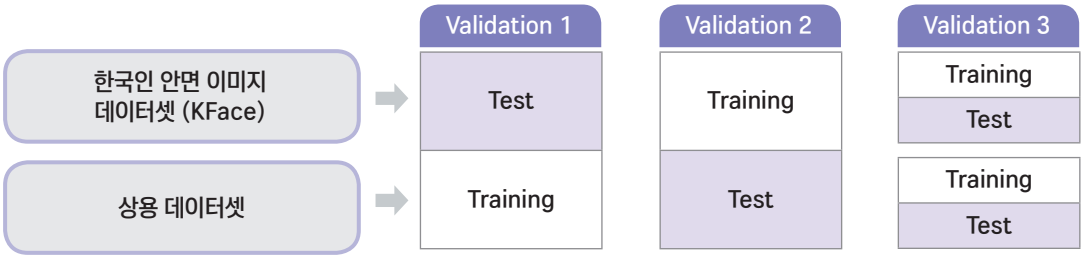
- 구축되는 한국인 안면 이미지 데이터셋의 유효성을 검증하기 위하여, 교차 유효성 검증 (Cross Validation) 방식을 활용할 수 있다.
- 기개발된 얼굴 인식 AI 알고리즘을 통해 본 데이터셋을 검증하도록 할 수 있으며, 본 데이터셋을 이용한 학습 모델의 성능 평가 또는 기학습된 모델(Pre-trained model)로 본 데이터셋을 평가하고, 이와 같은 교차 검증 성능을 통해 유효성을 최종 검증할 수 있다.

[그림 5.4-4] 데이터셋 유효성 검증 예시: N-Fold Cross Validation (N=5)



- 상기 그림과 같이 구축된 데이터베이스를 랜덤하게 섞은 후, 겹치지 않는 n개의 데이터셋으로 나눈 이후, 하나의 데이터셋 (테스트용)을 제외한 나머지 데이터셋을 활용하여 기개발된 AI 알고리즘을 학습한 후, 학습된 모델에 대한 테스트를 진행할 수 있다.
- 이러한 테스트를 n번 반복한 (Iteration) 후, 각 Fold에 대한 성능 비교를 통해 본 데이터셋의 유효성을 검증할 수 있다.

[그림 5.4-5] 데이터셋 유효성 검증 예시: Cross Validation



- 또 다른 유효성 검증으로는 상기 그림과 같이 기구축된 다양한 상용 데이터셋과의 학습 성능 비교를 통한 데이터셋 유효성 검증이 있을 수 있다.
- 구축된 한국인 안면 이미지 데이터셋과 상용 데이터셋을 교차하여 기개발된 AI 알고리즘을 학습시킨 이후, 성능 비교를 통해 본 데이터셋의 유효성을 검증할 수 있다.
- 유효성 검증을 위해서는 다양한 공개용 학습 모델을 활용할 수 있으며, 대표적인 안면 응용 학습 모델로는 얼굴 인식을 위한 LightCNN, 가상 얼굴을 생성하는 PGGAN 등이 있다.
- 추후 본 사업에서 구축되는 한국인 안면이미지 데이터를 활용하여 해당 모델을 학습시킬 수 있을 것으로 예상되며, 아래의 두 예시 이외에도 다양한 안면 응용 기술에 대한 공개 학습 모델이 존재하며, 목적에 따라 활용할 수 있을 것으로 전망된다.

1) LightCNN

- 논문 : W. Xiang, et al. “A light cnn for deep face representation with noisy labels.” IEEE Transactions on Information Forensics and Security vol. 13, no. 11, pp. 2884-2896, 2018

- Github : https://github.com/AlfredXiangWu/LightCNN
- LightCNN은 CNN기반의 얼굴인식 알고리즘으로 기존에 존재하는 LFW, YTF, YTC, MegaFace, IJB-A, IJB-B 데이터베이스를 활용한 학습 모델이다.
- Light CNN은 딥러닝 구조에 따라 총 3개의 모델이 존재하며, 각 모델은 아래 표와 같은 구조를 가진다.

[표 5.4-1] Light CNN-4 Model

Type	Filter Size /Stride	Output Size	#Params
Conv1	$9 \times 9/1$	$120 \times 120 \times 96$	7.7K
MFM1	-	$120 \times 120 \times 48$	-
Pool1	$2 \times 2/2$	$60 \times 60 \times 48$	-
Conv2	$5 \times 5/1$	$56 \times 56 \times 192$	230.4K
MFM2	-	$56 \times 56 \times 96$	-
Pool2	$2 \times 2/2$	$28 \times 28 \times 96$	-
Conv3	$5 \times 5/1$	$24 \times 24 \times 256$	614K
MFM3	-	$24 \times 24 \times 128$	-
Pool3	$2 \times 2/2$	$12 \times 12 \times 128$	-
Conv4	$4 \times 4/1$	$9 \times 9 \times 384$	786K
MFM4	-	$9 \times 9 \times 192$	-
Pool4	$2 \times 2/2$	$5 \times 5 \times 192$	-
fc1	-	512	2,457K
MFM_fc1	-	256	-
Total	-	-	4,095K

[표 5.4-2] Light CNN-9 Model

Type	Filter Size /Stride, Pad	Output Size	#Params
Conv1	$5 \times 5/1, 2$	$128 \times 128 \times 96$	2.4K
MFM1	-	$128 \times 128 \times 48$	-
Pool1	$2 \times 2/2$	$64 \times 64 \times 48$	-
Conv2a	$1 \times 1/1$	$64 \times 64 \times 96$	4.6K
MFM2a	-	$64 \times 64 \times 48$	-
Conv2	$3 \times 3/1, 1$	$64 \times 64 \times 192$	165K
MFM2	-	$64 \times 64 \times 96$	-
Pool2	$2 \times 2/2$	$32 \times 32 \times 96$	-
Conv3a	$1 \times 1/1$	$32 \times 32 \times 192$	18K
MFM3a	-	$32 \times 32 \times 96$	-
Conv3	$3 \times 3/1, 1$	$32 \times 32 \times 384$	331K
MFM3	-	$32 \times 32 \times 192$	-
Pool3	$2 \times 2/2$	$16 \times 16 \times 192$	-
Conv4a	$1 \times 1/1$	$16 \times 16 \times 384$	73K
MFM4a	-	$16 \times 16 \times 192$	-
Conv4	$3 \times 3/1, 1$	$16 \times 16 \times 256$	442K
MFM4	-	$16 \times 16 \times 128$	-
Conv5a	$1 \times 1/1$	$16 \times 16 \times 256$	32K
MFM5a	-	$16 \times 16 \times 128$	-
Conv5	$3 \times 3/1, 1$	$16 \times 16 \times 256$	294K
MFM5	-	$16 \times 16 \times 128$	-
Pool4	$2 \times 2/2$	$8 \times 8 \times 128$	-
fc1	-	512	4,194K
MFM_fc1	-	256	-
Total	-	-	5,556K

[표 5.4-3] Light CNN-29 Model

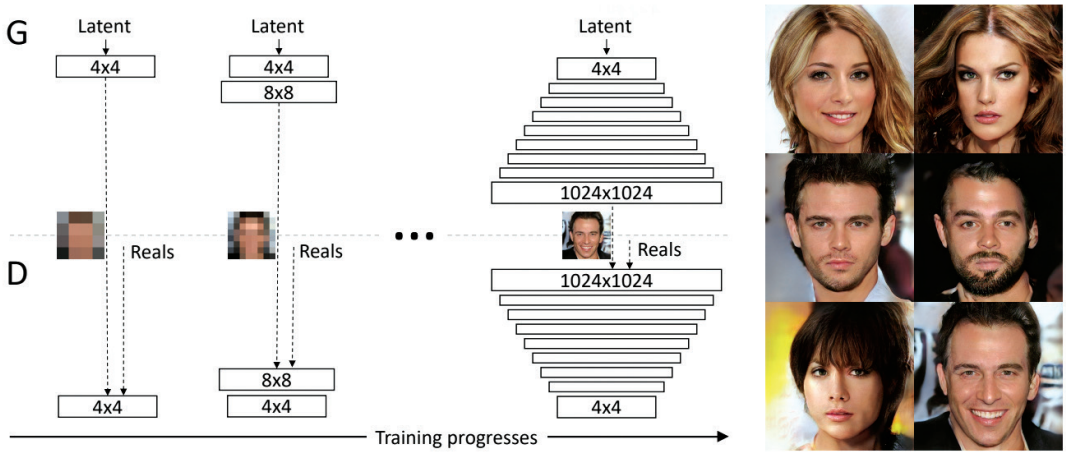
Type	Filter Size /Stride, Pad	Output Size	#Params
Conv1	$5 \times 5/1, 2$	$128 \times 128 \times 96$	2.4K
MFM1	-	$128 \times 128 \times 48$	-
Pool1	$2 \times 2/2$	$64 \times 64 \times 48$	-
Conv2_x	$\begin{bmatrix} 3 \times 3/1, 1 \\ 3 \times 3/1, 1 \end{bmatrix} \times 1$	$64 \times 64 \times 48$	82K
Conv2a	$1 \times 1/1$	$64 \times 64 \times 96$	4.6K
MFM2a	-	$64 \times 64 \times 48$	-
Conv2	$3 \times 3/1, 1$	$64 \times 64 \times 192$	165K
MFM2	-	$64 \times 64 \times 96$	-
Pool2	$2 \times 2/2$	$32 \times 32 \times 96$	-
Conv3_x	$\begin{bmatrix} 3 \times 3/1, 1 \\ 3 \times 3/1, 1 \end{bmatrix} \times 2$	$32 \times 32 \times 96$	662K
Conv3a	$1 \times 1/1$	$32 \times 32 \times 192$	18K
MFM3a	-	$32 \times 32 \times 96$	-
Conv3	$3 \times 3/1, 1$	$32 \times 32 \times 384$	331K
MFM3	-	$32 \times 32 \times 192$	-
Pool3	$2 \times 2/2$	$16 \times 16 \times 192$	-
Conv4_x	$\begin{bmatrix} 3 \times 3/1, 1 \\ 3 \times 3/1, 1 \end{bmatrix} \times 3$	$16 \times 16 \times 192$	3981K
Conv4a	$1 \times 1/1$	$16 \times 16 \times 384$	73K
MFM4a	-	$16 \times 16 \times 192$	-
Conv4	$3 \times 3/1, 1$	$16 \times 16 \times 256$	442K
MFM4	-	$16 \times 16 \times 128$	-
Conv5_x	$\begin{bmatrix} 3 \times 3/1, 1 \\ 3 \times 3/1, 1 \end{bmatrix} \times 4$	$16 \times 16 \times 128$	2356K
Conv5a	$1 \times 1/1$	$16 \times 16 \times 256$	32K
MFM5a	-	$16 \times 16 \times 128$	-
Conv5	$3 \times 3/1, 1$	$16 \times 16 \times 256$	294K
MFM5	-	$16 \times 16 \times 128$	-
Pool4	$2 \times 2/2$	$8 \times 8 \times 128$	-
fc1	-	512	4,194K
MFM_fc1	-	256	-
Total	-	-	12,637K

- Light CNN-29 모델의 얼굴 인식 성능이 가장 높으며, 해당 Github에서 공개 알고리즘 및 기 학습 모델(Pre-trained Model)을 다운로드 후 인식 알고리즘 개발에 활용할 수 있다.
- 본 사업에서 구축하는 한국인 안면이미지 또한 LightCNN 모델의 학습용으로 활용할 수 있으며, 이를 통해 해당 알고리즘의 성능 개선을 기대할 수 있다.

2) Progressive Growing of Generative Adversarial Network (PGGAN)

- 논문 : Karras, Tero, et al. "Progressive growing of gans for improved quality, stability, and variation." arXiv preprint arXiv:1710.10196, 2017.
- Github : https://github.com/tkarras/progressive_growing_of_gans
- PGGAN은 현재 이미지 생성 기술의 기존 GAN보다 업그레이드된 성능을 갖는 모델로서, 해상도를 높인 CelebA (178×218) 데이터인 CelebA-HQ (1024×1024) 데이터베이스를 활용한 모델이라고 할 수 있다.

[그림 5.4-6] PGGAN 구조



- PGGAN은 Generator와 Discrimiator의 구조를 점진적으로 저해상도(4×4)에서 고해상도(1024×1024)로 레이어를 쌓는 구조를 통해 기존보다 향상된 속도와 이미지 생성 성능을 지닌 알고리즘이라고 할 수 있다.
- 위의 Github에서 PGGAN의 공개 알고리즘 및 기학습 모델을 다운로드할 수 있으며, 이를 활용한 이미지 생성 모델의 성능 테스트가 가능하다.
- 본 사업에서 구축하는 한국인 안면이미지 또한 PGGAN 모델의 학습용으로 활용할 수 있으며, 이를 통해 한국인 얼굴 분포에 맞게 해당 알고리즘의 성능 개선이 가능할 것으로 예상된다.

5.5 한국인 안면 이미지 데이터 활용

5.5.1 한국인 안면 이미지 데이터셋 대외 공개 방안

- 안면 이미지 데이터셋은 개인 정보 유출에 민감한 데이터로서, 연구 목적 혹은 기업에서의 활용에 있어서, 정교한 설계가 선행되어야 한다.

[표 5.5-1] 데이터 공개 항목 및 공개 범위

공개 항목		공개 범위
데이터	연구용	원본 데이터, 학습모델
	상업용	학습모델
전처리 툴		실행 프로그램
얼굴 인식기		학습 및 테스트 코드
얼굴 변환기		학습 및 테스트 코드
서비스 모델		비즈니스 모델 및 성능 개선 리포트

- 이에 한국과학기술연구원 내에 데이터 공유 샌드박스 시스템을 구축하여, 개인정보유출을 방지하면서 동시에, 연구자 혹은 기업에서 비교적 자유로운 데이터 학습을 가능할 수 있도록 하였다.
- 또한, 개인정보보호를 위해 피실험자에게 개인정보동의서 상으로 연구 목적 활용에 대한 공유 동의를 받고 있으며, 상업적 목적으로는 데이터베이스를 직접 공유하지 않고, 학습 모델만을 제공할 수 있도록 하였다.
- 또한 촬영자의 ID는 개인정보 이외의 임의의 ID를 할당하여 저장하도록 하였으며, 개인을 특정화할 수 있는 어떠한 정보도 제공하지 않고, KIST내의 보안서버에 데이터베이스를 저장하여 외부 유출을 방지하도록 하였다.
- 이에 따르는 IRB 심사 및 법제도 자문을 수행하여 법적 문제가 발생하지 않도록 추진하였다.

[그림 5.5-1] 한국인 안면이미지 데이터셋 대외 공개 구조도

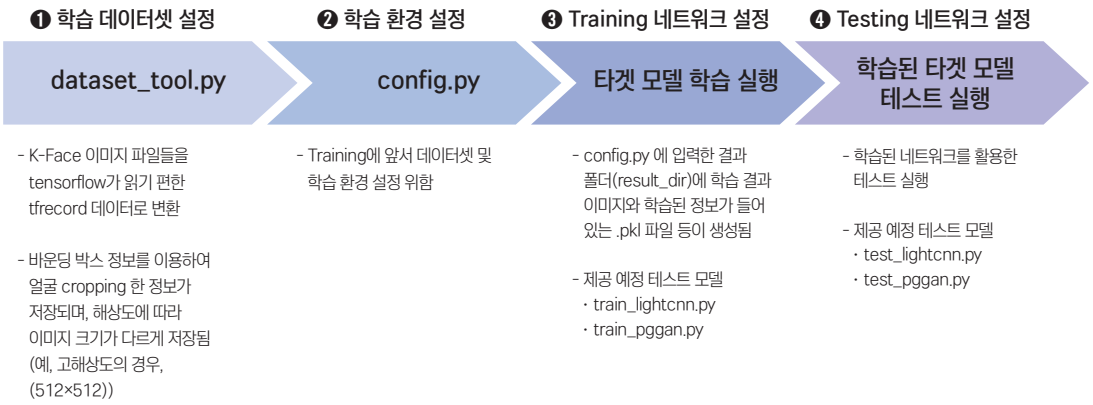
- 한국정보화진흥원에서 운영하고 있는 AIHub 사이트를 통해 안면 이미지 데이터셋을 요청하고 접근할 수 있도록 시스템을 구축하고, 이에 대한 시스템 매뉴얼을 공개도록 하였다.

- 연구 목적 활용의 경우, 기관 및 활용 목적 등에 대한 검토 후 데이터셋을 제공하도록 하며, 상업적 목적으로 활용하는 기업의 경우에는 학습하고자 하는 인공지능 네트워크를 제공 받아, 이를 샌드박스 시스템 내에서 학습을 진행한 이후 학습된 모델을 기업에 제공하도록 한다.
- 상업적 목적으로 기업이 데이터 샌드박스 시스템을 이용할 경우 정보 보안 문제없이 대용량의 데이터를 효율적으로 처리할 수 있도록 데이터 접근 및 학습 API를 개발하여 활용하도록 하였다.

○ 5.5.2 한국인 안면 이미지 데이터셋 활용 방안 및 API 구조

- 한국인 안면 이미지 데이터셋의 이미지는 촬영 각도에 따라 조명 혹은 의자 등 피실험자 이외의 영역이 포함되어 있는 경우가 있으므로, 목적에 따라 영상 내의 원하는 영역 (얼굴, 귀 등)을 Cropping하여 학습에 활용하는 것을 권장한다.
- K-Face Database 활용을 위한 API 구조는 다음과 같으며, 본 가이드라인에서는 텐서플로우(Tensorflow) 딥러닝 프레임워크를 활용하였다.
- 본 가이드라인에서 제시하는 API는 다양한 딥러닝 프레임워크 중 하나의 예시로, 다른 형태의 API가 존재할 수 있으며, 사용자가 본인의 프레임워크에 맞도록 새롭게 구현하는 것도 권장한다.

[그림 5.5-2] 코드 실행 순서 요약



① 학습 데이터셋 생성

- **코드명**: dataset_tool.py
- **목적**: 딥러닝 모델 학습에 사용될 이미지들을 추출하여, 텐서플로우 프레임워크에서 데이터를 불러올 수 있도록 함

[그림 5.5-3] 학습 데이터셋 생성 코드(dataset_tool.py) 도움말 화면

```
Anaconda Prompt (Anaconda3)

(base) D:₩KIST₩_code₩K-Face>python dataset_tool.py create_kface -h
usage: dataset_tool.py create_kface [-h] [--resolution RESOLUTION]
                                    [--session SESSION] [--light LIGHT]
                                    [--emotion EMOTION] [--camera CAMERA]
                                    [--shuffle SHUFFLE]
                                    tfrecord_dir kface_image_dir

Create dataset from a directory full of K-Face images.

positional arguments:
  tfrecord_dir          New tfrecord dataset directory to be created
  kface_image_dir       Directory containing the K-Face images

optional arguments:
  -h, --help            show this help message and exit
  --resolution RESOLUTION
                        Select resolution type. Input one of three types:
                        High, Middle, Low (default: High)
  --session SESSION     Select one session type [S001~S006] (default: S001)
  --light LIGHT         Select one light type [L1~L30] (default: L1)
  --emotion EMOTION     Select one emotion type [E01, E02, E03] (default: E01)
  --camera CAMERA       Select one camera(degree) type [C1~C20] (default: C7)
  --shuffle SHUFFLE     Randomize image order (default: 1)

Example: dataset_tool.py create_kface datasets/my/tfrecord/dir my/image/dir
```

- 생성하고자 하는 데이터의 해상도(resolution), 액세서리(session), 조명 방향 및 세기(light), 표정(emotion), 얼굴각도(camera), 데이터 무작위 섞음(shuffle) 값을 입력한다.
- 바운딩 박스 정보를 이용하여 딥러닝 학습 네트워크를 실행시킬 수 있도록 해상도에 맞게 데이터 Cropping 한 이후, 텐서플로우 프레임워크가 데이터를 읽도록 되어 있으며, 이때, 바운딩 박스 정보는 동일 폴더 내에 존재해야 한다.

② 학습 환경 설정

- **코드명**: config.py
- **목적**: 데이터셋 정보 및 학습 환경을 구체적으로 설정하기 위함
- **활용 방법**
  - data_dir: 위의 ①번 과정을 통해 획득된 데이터가 저장되어있는 디렉토리
  - result_dir: 학습 결과물이 저장될 디렉토리
  - 기본적으로 single-GPU로 설정되어 있으며, GPU 수를 비롯한 학습 환경은 코드 내의 주석 수정을 통해 설정 가능

③ Training 네트워크 설정

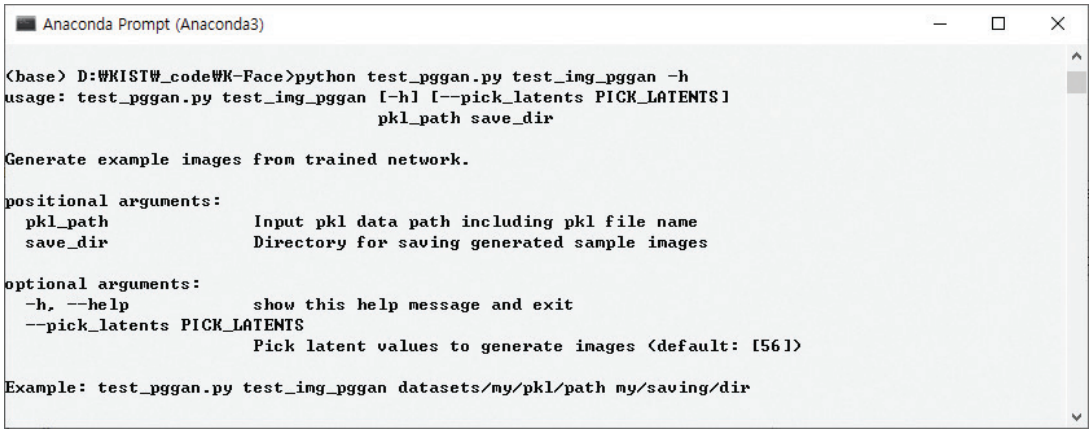
- 본 가이드라인에서는 학습모델 PGGAN을 예시로 설명하도록 한다.
- **코드명** : train_pggan.py
- **목적** : PGGAN 모델 학습 코드로, 이를 실행시켜 네트워크 학습을 진행함

- 진행 순서
  1. 위의 ②번의 'config.py'에서 학습 환경 설정
  2. 'python train_pggan.py'를 실행하여 모델 학습 진행
  3. 'config.result_dir' 내의 결과물 확인 가능
  4. GPU 사양에 따라 수행 시간은 다르며, 학습 완료 후 결과 분석 진행 가능

④ Testing 네트워크 설정

- 본 가이드라인에서는 학습모델 PGGAN을 예시로 설명하도록 한다.
- **코드명**: test_pggan.py
- **목적**: 학습된 PGGAN 네트워크를 이용하여 테스트 및 결과 분석 진행

[그림 5.5-4] 학습이 완료된 PGGAN 네트워크를 이용한 이미지 예시 생성 코드(test_pggan.py) 도움말 화면



08

위해물품 엑스레이 이미지 AI 데이터 구축 가이드라인

MPOLE system

	성명	소속	직위
가이드라인 제안	이상혁	엠폴시스템	연구소장
가이드라인 작성	이상혁	엠폴시스템	연구소장
전문가 검토 위원	이영한	KETI	팀장
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 엠폴시스템/ 참여: 인씨스, 국가사이버안전연합회

서 문

1. 가이드라인의 목적

이 가이드라인의 목적은 공항 및 항만 등 물품의 검사를 위한 X-Ray Scanner 장비의 영상을 기반으로 위해 또는 금지 물품을 자동으로 검출 하는 AI 자동검출 기술 분야에서 필요로 하는 물품의 이미지와 Annotation 정보의 집합체를 구축하기 위한 가이드라인을 정의하고 구축을 위한 절차와 세부 지침을 제공하기 위해 작성되었다.

2. 주요 내용 요약

이 가이드라인은 X-Ray Scanner 장비에 투과되는 다양한 목적 물체(위해물품 및 저장 매체)를 자동으로 탐지 하기 위해 필요한 AI(Artificial Intelligence) 학습용 이미지와 어노테이션 정보로 구성이 된다.

1. X-Ray Scan Image

무손실 압축 방식을 기반으로 한 데이터 셋 공개 서비스를 위한 저장, 조회 등과 같이 운영에 효율적인 파일 포맷을 이용하여 데이터를 생성한다.
해당 Scan Image는 무손실 압축 형태의 Png 형태로 생성한다. 왜곡이나 노이즈 없는 이미지 생성을 목표로 하고 있으며 생성 단계에 비정상 스캔 이미지의 경우
제외 하여 진행 한다.

2. Annotation Information

X-Ray Scan Image 내에 다양한 목적 물체(위해물품 & 저장 물품)를 다양한 방법으로 생성하며 이에 대한 AI 학습을 위한 Annotation 정보를 제작한다.
기본 CoCoDataset포맷을 기준으로 메타정보를 생성하며, 이미지내의 물체를 Bounding Box / Segmentation 으로 Annotation하여 포함되어 제공한다.

3. 인용 문서와의 비교

3.1 인용 문서와의 관련성

- 해당 사항 없음

3.2 인용 표준 또는 문서와 본 가이드라인의 비교표

- 해당 사항 없음

위해물품 엑스레이 이미지 데이터 구축

Building X-Ray Image AI Dataset for Hazardous Materials

08

1. 적용 범위

이 표준은 위해물품 엑스레이 이미지 데이터 구축을 위한 절차와 어노테이션 지침을 제시하여 국토교통부(국토교통부고시 제2016-1091호: 항공기 내 반입금지 위해물품) 기준을 통해 선정된 물품과 물리보안 대기업 및 연구소 반입 금지 물품(저장 매체)을 대상으로 데이터-셋을 구축하고 이에 대한 Annotation 정보를 생성하고 AI 기반의 자동 물품 검지 학습을 위한 데이터-셋 구축 및 물품을 대상으로 한다.

이미지 데이터는 X-Ray Scan을 통해 표출되는 Pseudo-6 Color 대상의 제품으로 선정하며 국내 공항 항만에 주로 설치되는 장비 제조사의 제품으로 구성한다.

(제조사 : Rapiscan, Smith-Dectection, Astro-Physics)

2. 인용 표준

- 국토 교통부 고시 제 2016-1091호 항공기 내 반입 금지 위해 물품
- 「항공안전 및 보안에 관한 법률」을 「항공보안법」으로, 「훈령·예규 등의 발령 및 관리에 관한 규정」(대통령훈령 제248호)을 「훈령·예규 등의 발령 및 관리에 관한 규정」
- 「항공보안법」제21조제1항에 따라 항공기 내에 반입해서는 아니되는 물품 규정
- 반입 금질 물품 리스트 「부록 1 참조」

3. 용어 정의

- 해당 사항 없음

4. 약어

- 해당 사항 없음

5. 데이터 수집

● 5.1 물품 종류 및 데이터-셋 수량 산정

데이터 셋 구축을 위한 물품 및 데이터 셋 수량 산정을 위해 국토부 고시를 참조하여 아래와 같이 정의한다.

[그림 1] 위해물품 엑스레이 이미지 데이터-셋 내역

위해물품			90%	저장매체			10%
품명	샘플수	샘플당 이미지수	총 데이터 량	품명	샘플수	샘플당 이미지수	총 데이터 량
탄약	15	800	12,000	핸드폰	15	500	7,500
폭죽	25	800	20,000	USB	20	500	10,000
라이터	25	800	20,000	노트북	20	500	10,000
에어로졸	10	700	7,000	Tablet	10	500	5,000
휴대용가스	10	700	7,000	SSD	10	500	5,000
알코올	5	500	2,500	HDD	5	500	2,500
신나	5	500	2,500	외장하드	5	500	2,500
라이터기름	5	500	2,500				
성냥	10	500	5,000				
고체연료	10	500	5,000				
건전지	5	500	2,500				
휴대용 배터리	30	500	15,000				
손톱깎이	20	500	10,000				
칼	35	700	24,500				
톱	35	700	24,500				
가위	35	700	24,500				
송곳	20	700	14,000				
드라이버	25	600	15,000				
스패너	25	600	15,000				
액체(100ml이상)	25	500	12,500				
수갑	5	500	2,500				
도끼	20	800	16,000				
망치	20	800	16,000				
야구배트	10	500	5,000				
총기	20	800	16,000				
금속 파이프	25	500	12,500				
총기부품	25	800	20,000				
끌	25	700	17,500				
플라이어	20	500	10,000				
표창	5	500	2,500				
전자담배(액상)	10	500	5,000				
전자담배(결련)	10	500	5,000				
32개 품목	570		369,000	7개 품목	85		42,500

주로 항공기, 선박의 운행에 방해되는 물품 및 안전에 심각한 위협을 초래 하는 물품을 기준으로 구성하였으며 해당 데이터-셋의 X-Ray Scan을 통한 이미지 획득을 100% 기준으로 사용한다.
해당 이미지별 Unbalanced Problem의 제거를 위해 영상의 복잡도 및 특장점 등 가중치에 따라 최소 500~최대 800개 사이에서 데이터-셋의 수를 조정한다.
학습용 데이터와 검증용 데이터 그리고 수행기관 보유 데이터를 포함하여 아래와 같은 데이터-셋을 공개한다.

[표 1] 위해물품 엑스레이 이미지 데이터-셋 목적별 데이터 수

	학습용	검증용	당사 데이터 제공	합계
위해물품	369,000	36,900	-	405,900
저장매체	42,500	4,250	30,000	76,750
	411,500	41,150	30,000	482,650

● 5.2 이미지 수집

엑스레이 영상의 입력을 통해 영상을 취득(수집)하는 단계로 목적 물체의 선정 이후 데이터셋 생성 및 분포를 정의하여 X-Ray Scan을 통해 이미지를 생성 하는 단계.

○ 5.2.1 이미지 포맷 및 생성 정보

- File Format : PNG(무손실 압축 방식)
- Resolution : 1920 × 1080(Full HD 또는 장비 제조사별 최대 해상도)
- 입력신호 : HDMI
- 생성 방식 : HDMI Grabber를 이용한 Image Capture 방식
- Software Tool(Image Capture SW)를 이용한 자동 생성

○ 5.2.2 X-Ray Scanner 하드웨어 정보

제조사	모델명	표출
Rapiscan	RAP620XR	Psuedo 6 Color(HDMI)
Smith Detection	HI-SCAN60402is	Psuedo 6 Color(HDMI)
Astro Physics	XIS6040	Psuedo 6 Color(HDMI)

- RAP620XR

Steel Penetration	30 mm typical; 28 mm standard
Wire Resolution	40 AWG typical; 38 AWG standard
Generator Cooling	Sealed oil bath with forced air
Anode Voltage	Operating at 140kV
Tube Current	0.7 mA
Orientation	Vertically Upward

- HI-SCAN60402is

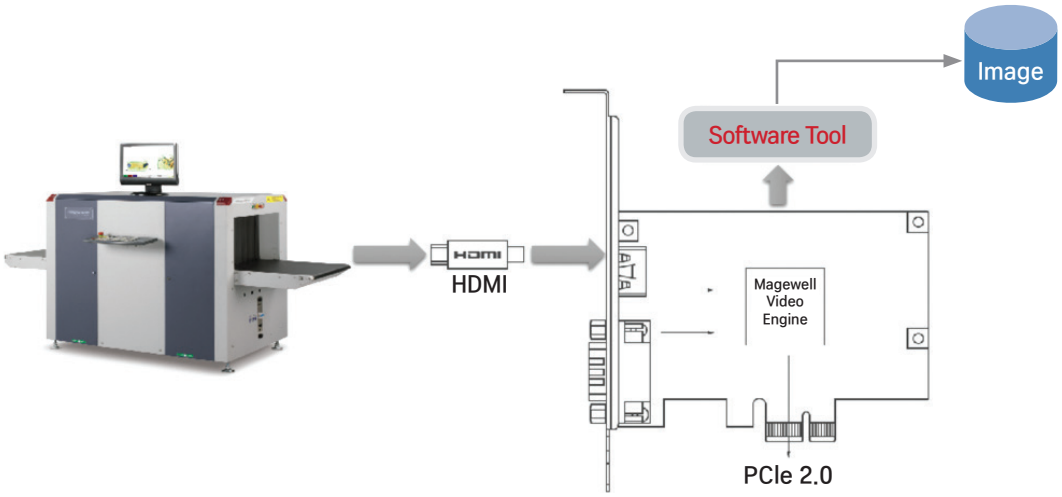
X-ray converter	dual-view detector system with Large-scale integrated, monolithic amplifier
Grey levels stored	4096
Image presentation	B/W, color, HI-MAT
Digital video memory	1280 x 1024 / 24 bit
Image evaluation functions	VARI-MAT, O^2, OS, HIGH, LOW, NEG stepless, electronic Zoom: enlargement: up to 16 times
Monitor	Flat Panel LCD Monitor

- XIS6040

Wire Resolution	40 AWG typical / 38 AWG standard
Steel Penetration	39 mm typical / 37 mm standard
Spatial Resolution	1.0 mm horizontal / 1.0 mm vertical
Film Safety	Guaranteed up to ISO 1600 (33DIN)
Organic/Inorganic Material Discrimination	The system is capable of discriminating between organicand inorganic materials as measured by the ASTM F792-08 Test Piece - Test #5 & #7.
Organic Differentiation	The system is capable of classifying organic material masked by 0.48 mm ofsteel as measured by the ASTM F792-08 Test Piece - Test #9.

3개의 장비를 기반으로 데이터셋 물량은 1:1:1로 정의하여 진행.
시장점유율에서 차이를 보이고 있으나 3개사 제품 모두 데이터-셋 활용도에서 동등한 데이터-셋으로 기술 발전을 시켜야 하는 점을 고려하여 해당 데이터-셋의 물량을 동일한 비율로 계획 수립하여 진행.

○ 5.2.3 HDMI Grabber Board 하드웨어 정보



1) Image Capture Work Station

Item	Specification
CPU	Intel Xeon E5-2640 v2 2.00Ghz 20MB 1600 8C 1st CPU
Memory	8GB DDR4 또는 그 이상
VGA	NDIVIA GTX 1080 Ti 11GB
Capture Card	HDMI Capture BD
OS	Windows 10 or 7 64Bit
Application	X-Ray AI Image Inspection Application

Image Capture를 위한 소프트웨어 툴은 물품을 엑스레이에 넣고 Scan이 완료 되면 자동으로 PNG 이미지를 생성하여 정제인원에서 전송한다.

○ 5.2.4 물품 데이터 Scan 방식 정의

생성방식		데이터 생성 방식	
단일	기본 (15%)	해당 품목의 샘플 1개에 해당 물품 이외에 어떠한 물품을 넣지 않고 생성하는 방식	Single_Default
	비품목 (15%)	해당 품목의 샘플 1개에 다른 위해물품 없이 일반 품목을 동시에 넣고 생성하는 방식	Single_Other
복합	품목 (40%)	해당 품목의 샘플 복수 수량과 다른 위해물품과 동시에 넣고 생성 하는 방식	Multiple_Categories
	비품목 (30%)	해당 품목의 샘플 복수 수량과 다른 위해물품 없이 일반 품목을 동시에 넣고 생성하는 방식	Multiple_Other

위 표와 같이 4가지 방식으로 이미지를 생성 한다. 단일 기본의 경우 해당 샘플 하나에 대한 멀티뷰 및 싱글뷰를 적용하여 다양한 각도 및 진입 방향을 고려하여 목적 이미지 총 데이터의 15% 정도를 생성한다. 단일 비품목의 경우 해당 샘플과 위해 물품이 아닌 일반 비품목(책, 옷, 기타 물품) 들과 동시에 다양한 각도 및 긴입 방향을 고려하여 목적 이미지 총 데이터의 15% 정도를 생성한다. 복합 품목의 경우 해당 샘플과 동일 품목의 다른 샘플과의 동시 이미지 생성하는 방식으로 총 데이터의 40% 정도를 생성한다. 복합 비품목의 경우 복합 품목에 비품목(책, 옷, 기타 물품)을 조합하여 총 데이터의 30% 정도를 생성한다. 위 데이터 생성은 가방, 케리어, 검색용 바스킷 등을 이용하여 데이터를 생성하게 된다.

위 정의된 생성 방식은 물리적인 폴더 구조에 정의 되며 명명 방안은 아래와 같이 정의됨
Annotation 정보에 폴더를 통해 기준 물품의 정의가 되어 있음.

예) “path”: “/datasets/Astrophysics/USB/Single_Default/S_8544-42-2090_001_0200.png”,
장비제조사/기준물품/생성방식/물품HSCODE_샘플번호_이미지번호

• 비품목 리스트

여행용 세면도구	신발	화장품	섹션케이스
의류	악세사리	가방	손거울
케이블류	문구류	카메라	치실
우산	액자	휴대용 선풍기	물티슈
안경	열쇠고리	셀카봉	머리빗
벨트	약품	파우치	가공류

○ 5.2.5 엑스레이 물품 Color별 속성 정의

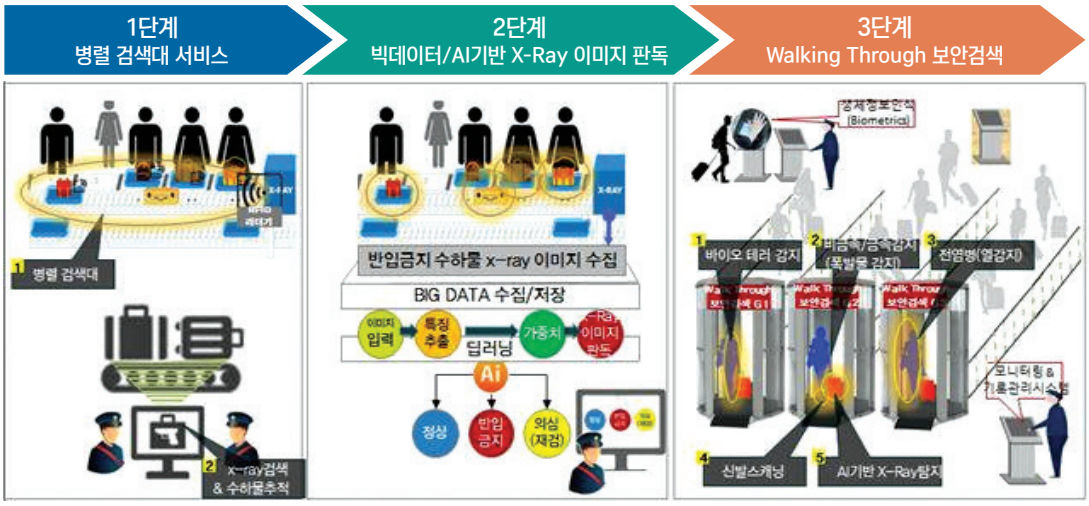
Color imaging3기법은 Astrophysics. Inc. 등에서 개발되었는데, 6가지의 색상을 통해화물의 물질적 속성 정보 판독원에게 제공한다. 3가지 혹은 6가지로써 화물이 가질 수 있는 매우 다양한 속성 정보 중 극히 일부만을 제한적으로 알려주므로, 평균적인 물질 정보만을 제공.

Z-Number	Material Type	3 Color	6 Color	Examples	Possible Threats
0-8	Organic	Orange	Brown	Wood, Oil	C-4, TNT, Semtex
8-10	Low Inorganic	Orange	Orange	Paper	Cocaine, Heroin
10-12	High Inorganic	Green	Yellow	Glass	Propellants
12-17	Light Metals	Green	Green	Aluminum, Silicon	Gunpowder, Trigger Devices
17-29	Heavy Metals	Blue	Blue	Iron, Steel	Guns, Bullets, Knives
29+	Dense Metals	Blue	Violet	Gold, Silver	High Value Contraband
-	Impenetrable	Black	Black	Lead	Shielding for Above Threats

* 제조사별 및 장비별 특징은 약간의 차이가 발생할 수 있으나 실제 Capture-board로 입력되는 이미지는 RGB기반의 24Bit 데이터를 활용하고 이를 png 포맷으로 압축을 하기 때문에 해당 파일의 헤더를 이용하여 컬러 값 확인이 가능
* 엑스레이 물품 컬러는 원소기호(물체 투과율)와 이를 표시하는 제조사의 특성으로 구성되어 차후 재질별 Color를 통해 분류 가능

○ 5.2.6 촬영 조건 및 중첩 이미지 처리

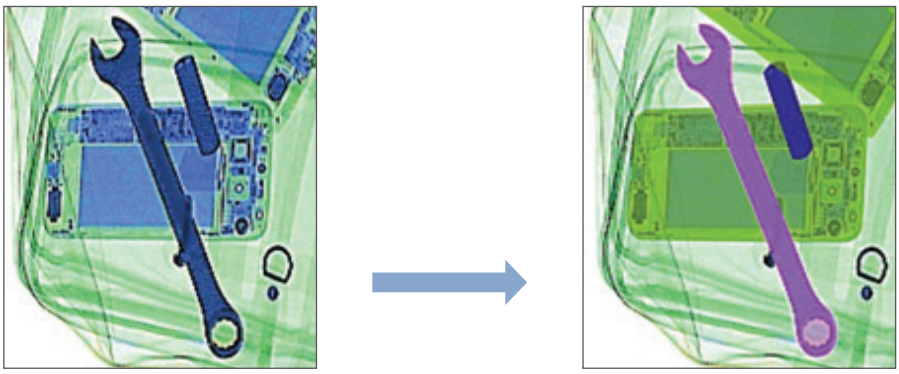
1) 촬영 조건



AI 분석을 통해 나온 결과는 정상, 반입금지, 의심(재검)의 3가지 상태로 표출 되게 된다. 특히 의심(재검) 상황에서는 X-Ray Scan 방향을 역방향으로 처리하거나 다시 재진입 하는 업무를 실행 하게 되는데 이 상황에서 세밀한 검색을 위한 보안 업무가 속하게 됩니다. 예를 들어 X-Ray가 투과 되지 않는 물체를 제거하여 재검 하거나 다른 각도로 데이터를 처리하거나 의심 물체의 투과를 확인하기 위해 재-검증 하는 업무 방식 적용

- 다양한 재질의 Baggage 및 내부 물품 적용
  경력직 보안 검색 요원의 업무 Know-How중 다양한 환경에서의 업무 경험이 필요.
  PVC, Fabric, Leather, 등 Baggage의 재질에 따른 투과 결과가 다른 부분이 존재하며 이는 자동 학습에 고려되어야 함.
  또 의류 이외에 기내 반입 및 항공기 반입의 품목에 대한 경험도 해당 데이터셋에 동일하게 적용되어야 함.

2) 중첩 이미지 처리



각 물품에 대한 특징 등이 지나치게 가려 식별이 어려울 정도가 아닌 이상(투과율로 인한 식별이 불가능한 수준 이 아닌 이상) Annotation을 진행

3) 이미지 각도 및 진입 방향

실제 이미지 각도와 진입 방향의 구분은 학습을 위한 프로퍼티 조정을 통해 진행이 가능하기 때문에 실제 중요도는 중첩에 비해 중요성이 매우 떨어진다.
실제 Rotate 파라메터를 이용하여 동일 물체에 대한 다양한 학습 옵션을 제공하기도 하고 Annotation 내부의 데이터를 참조하여 학습하기에 진입 방향이나 각도는 우선 순위에서 떨어지며 대부분의 물품이 Flat한 장비는 기울기에 대한 특징이 없는 것이 보통이다.
이에 Flat하지 않은 물품에 대한 데이터-셋 생성 시 고려하여 진행

[기본 데이터 생성 고려 사항]

- 중요 특징에 대한 식별이 가능(필수 요소 : 인간의 판단이 불확실한 경우 제외)
- 색상이나 형태가 기본 형태에 비해 매우 훼손 되는 경우 이는 Annotation에서 제외
- 물체가 겹쳐 있어도 특징에 대한 구분이 가능하며 윤곽이 뚜렷한 경우 Annotation
- 기본 특징에 공통점을 찾을 수 없을 정도로 각도가 있을 경우 Annotation 제외

○ 5.2.7 생성 이미지 파일 구조

1) Image 정보

이미지 특성 비교

File Format	BMP	JPG	PNG
압축여부	무손실	손실	무손실
사이즈 (Full HD기준)	4MB	900KB	1.5MB
표현 색상수	24Bit	24Bit	24Bit
생성속도	느림	빠름	빠름
적합성	2nd	3rd	1st

PNG 파일 포맷은 X-Ray Scanner 업체의 Raw데이터를 제외하고 AI 데이터 구축을 하는데 가장 최적의 이미지 데이터로 활용 가능성이 높습니다.
이러한 부분은 구축 및 활용 차원에서도 데이터 처리 속도 대비 용량에서도 좋은 File Format으로 선정될 수 있으며 파일 변환을 원하는 고객에게 JPG Converting Tool를 제공함으로 써 다양한 데이터 배포에 탄력적인 File Format이라 할 수 있습니다. PNG File Format으로 데이터를 구축 할 경우 제안 데이터의 이미지 부분의 용량은 다음과 같다.

$$1.5\ MB(\text{이미지1장평균사이즈}) \times 482{,}650(\text{데이터총수}) \fallingdotseq 725\ GB$$

2) 학습 메타 파일 포맷 정보

- 조정위원회의 요구 사항에 대한 PolyLine Annotation을 추가로 요청하여 데이터셋 구축시 동시에 구성이 가능하게 제공

- Json 기준으로 되어있으며, 한개의 json파일 안에 이미지 정보, annotation 정보 등 모든 정보 보유
- 큰 블록으로는 Info, annotation, images, categories로 분류
- 기존 Labeling Tool에 폴리곤 작업 가능하게 업그레이드 개발
- 어노테이션 정보는 Coco-Data Set 형태로 구성하여 Polygon 라벨링

$$10KB(Annotaion - Data : AVG) \times 482{,}605(\text{이미지수}) \fallingdotseq 5.5\ GB$$

* File Informations(Segmentation)

기본 템플릿 : json

```
{
        "images" : [],
        "categories" :[],
        "annotations":[]
}
```

<Image 정보 포맷>

```
{
        "id": (정수)[이미지 ID Key값]
        "dataset_id": (정수)[해당 이미지가 참조 되어 있는 dataset ID key],
        "path": "(문자열)[이미지 참조 경로]",
        "width": (정수) [이미지 가로 pixel 길이],
        "height": (정수) [이미지 세로 pixel 길이],
        "file_name": "(문자열)[이미지 파일명] "
}
```

<categories 정보 포맷>

```
{
        "id": (정수)[물체 Key],
        "name": "(문자열)[물체명]",
        "supercategory": "(문자열)[물체의 항목]",
        "color": "(문자열 RGBCode)[Annotation시의 색상]",
        "metadata": "(배열)[기타 필요 메타 데이터 입력 필드]"
}
```

<annotation 정보 포맷>

```
{
        "id": (정수)[Annotation ID],
        "image_id": (정수)[참조 이미지 ID,
        "category_id": (정수)[Category 즉 물체 정보],
        "segmentation": (2차원 실수)[polyline의 좌표 정보
```

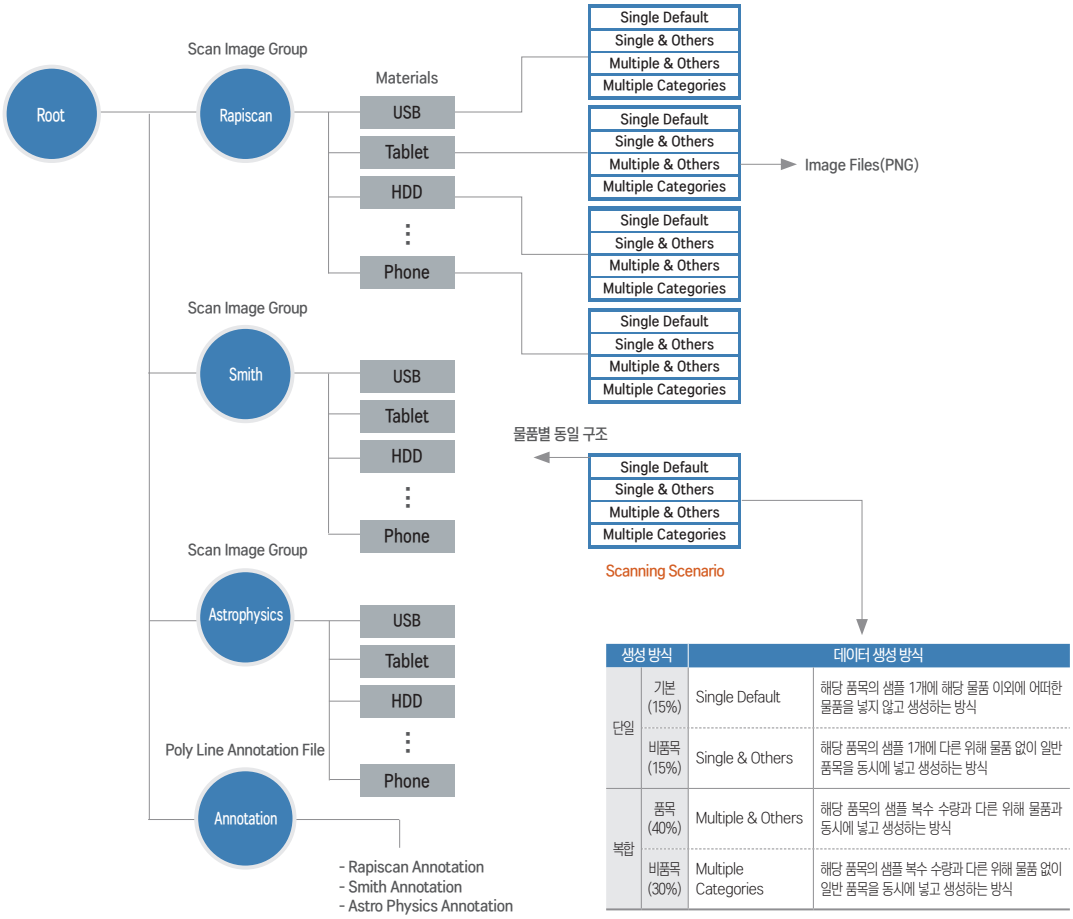
```
                                    (각 점에 대해 x,y를 반복하여 입력)],
                "area": [segmentation 보조정보],
                "bbox": (정수 or 실수)[x,y,width,height],
                "iscrowd": (bool)[복잡도 표시]],
                "color": "[annotation color 정보]",
                "metadata": "(배열)[기타 필요 메타 데이터 입력 필드]"
        }
```

* File Informations(Pascal VOC)

```
<?xml version="1.0" encoding="utf-8"?>
<annotation>
  <folder>ModifyData</folder> "(문자열) [folder name]"
  <filename>20180626093415599.png</filename> "(문자열)[Image File name]"
<path>G:₩XrayData₩Source₩98_ServerData₩ModifyData₩20180626093415599.png</path>
"(문자열)[Image File Location Path]"
  <source>
    <database>Unknow</database>
  </source>
  <size>
    <width>1024</width> "(정수) [Image Width Size]"
    <height>768</height> "(정수) [Image Height Size]"
    <depth>3</depth> "(정수) [Image 색차 정보 RGB]"
  </size>
  <segmented>0</segmented>
  <object>
    <name>USB_Metal</name> "(문자열) [물품 이름]"
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>184</xmin> "(정수) [Object Location]"
      <ymin>381</ymin>"(정수) [Object Location]"
      <xmax>202</xmax>"(정수) [Object Location]"
      <ymax>410</ymax>"(정수) [Object Location]"
    </bndbox>
  </object>
</annotation>
```

5.3 데이터셋 저장 위치 설명

공개 서비스를 위한 데이터셋의 다운로드 저장 위치는 아래와 같은 저장 위치를 구성한다.



생성 방식		데이터 생성 방식	
단일	기본 (15%)	Single Default	해당 품목의 샘플 1개에 해당 물품 이외에 어떠한 물품을 넣지 않고 생성하는 방식
	비품목 (15%)	Single & Others	해당 품목의 샘플 1개에 다른 위해 물품 없이 일반 품목을 동시에 넣고 생성하는 방식
복합	품목 (40%)	Multiple & Others	해당 품목의 샘플 복수 수량과 다른 위해 물품과 동시에 넣고 생성하는 방식
	비품목 (30%)	Multiple Categories	해당 품목의 샘플 복수 수량과 다른 위해 물품 없이 일반 품목을 동시에 넣고 생성하는 방식

Root를 중심으로 Layer1은 X-Ray 장비 벤더 별 폴더가 구성되어 Peudo-6컬러의 상이점이 다른 데이터의 집합을 분리함.

Layer2는 각각의 벤더별로 대상 물품에 대한 분리를 하여 폴더링이 되어 있음.

Layer3는 물품의 생성기준에 맞게 4개의 Catagory로 구성하여 폴더링됨.

학습 메타 정보는 폴더링 구조에 맞는 이미지의 경로와 이름을 기준으로 Segmentaion정보와 Bounding Box 정보를 포함한다.

6. 데이타 정제

● 6.1 이미지 정제

이미지 수집을 통해 생성된 데이터의 정제는 4가지 사항에 대한 정제를 진행한다. 이미지에 대한 육안 검출 및 이미지의 정보에 따른 방식을 동일하게 적용한다.

데이터-셋 네이밍 규칙"에 맞는 파일명으로 생성되었는지에 대한 검증으로 위배시 이미지 명 및 DB에 수정 적용하여 정제.

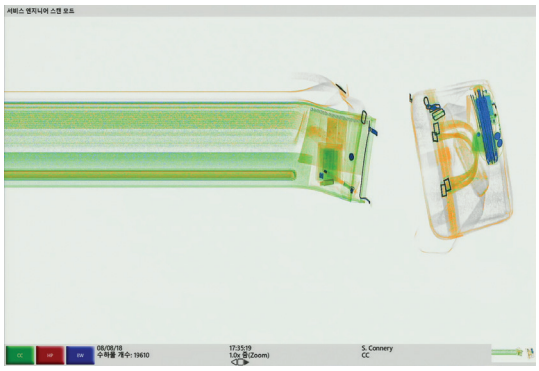
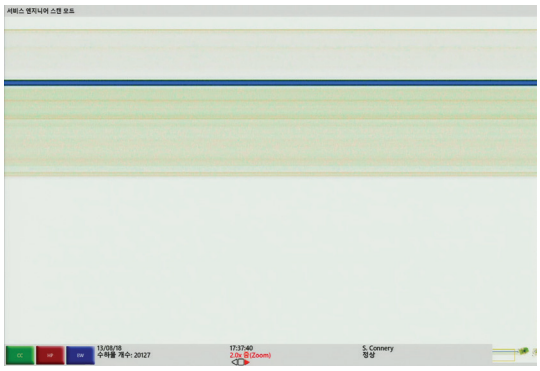
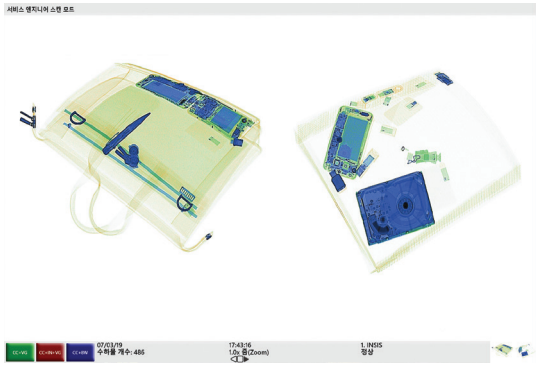
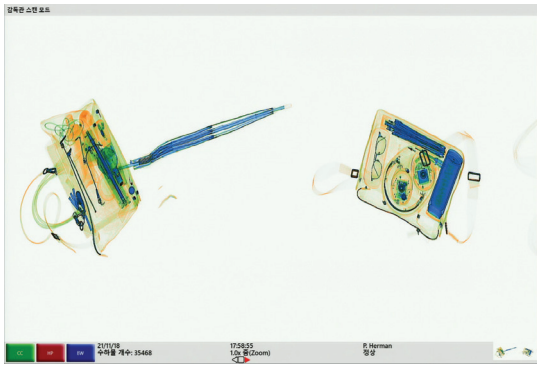
가로 픽셀, 세로 픽셀, 비트, 이미지 사이즈를 검토 하여 표준 데이터 셋과 다른 경우 데이터 정제 작업 진행.

[그림 9] 위해물품 엑스레이 이미지 데이터-셋 내역

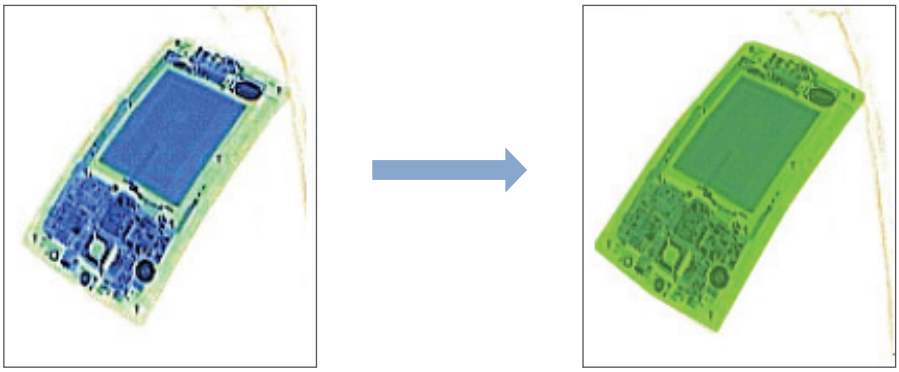
① 선정된 물품을 X-Ray Scanner에 투과하여 Scan 작업을 통해 투과를 시킨다.
사전에 정의된 데이터 생성방식을 기준으로 작업이 진행 되며 이때 X-Ray Scanner 화면 해상도는 Full-HD의 화질(PNG File)로 사전에 설정을 하여 진행 합니다.

② Capture & Inspection PC의 단계는 데이터를 생성하는 과정으로 X-Ray Scan
영상 중 X-Ray 내부에서 물체 밀림 현상으로 왜곡된 이미지를 찾아내어 제외 시키고 기타 비정상으로 생성된 데이터를 정제 합니다.

* 이미지 정제 대상 리스트

물품 늘어짐 현상	2. 장비 노이즈 현상
X-Ray Scanner 내부에 물품(Packing)이 걸려 컨베이어와 동일하게 움직이지 않고 밀리는 경우 발생하는 현상으로 학습용 데이터로 부적합하여 데이터-셋에서 제외.	X-Ray Scanner의 제너레이터 수명 또는 기타 전기/전자적이 사유로 생성되는 노이즈, 제너레이터 교체 또는 엑스레이 교체 필요, 해당 데이터의 경우 학습 데이터로 부적합하여 데이터-셋에서 제외
	
3. 물품 이미지 회절 현상	**4. 이미지 왜곡 현상**
X-Ray Scanner 이미지 중 물체의 모서리 부분의 회절이 발생하는 현상으로 Annotation 작업에 동일 물체에 대한 오차 발생 가능성 높음. 엑스레이 설정 변경 필요, 해당 데이터의 경우 학습 데이터로 부적합하여 데이터-셋에서 제외	물품의 원래 형태와 다르게 생성되는 현상으로 이미지 스캔 과정에서 물품이 이동하여 생기는 현상으로 정제 후 재 생성 필요, 해당 데이터의 경우 학습 데이터로 부적합 하여 데이터-셋에서 제외.
	

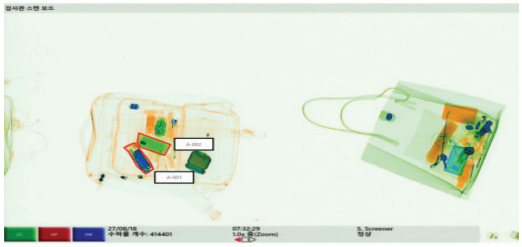
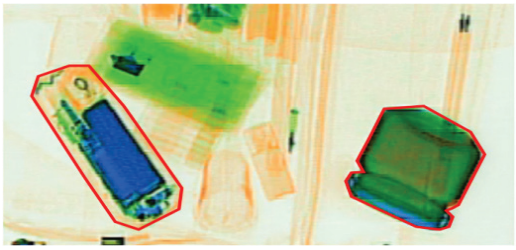
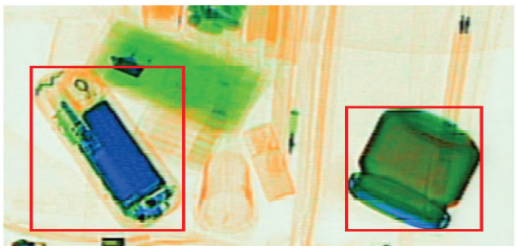
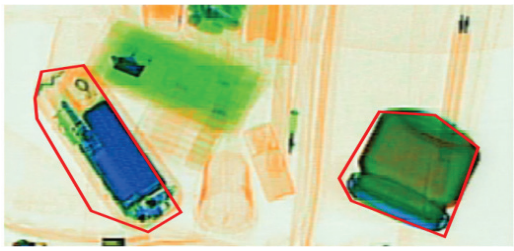
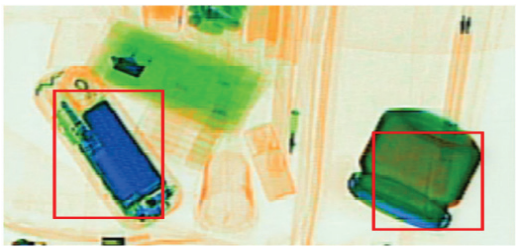
: 지나친 경우를 제외하고 Xray Scanner의 특성상 어느정도의 왜곡 현상은 발생하기 때문에 아래와 같은 사진 수준에 의한 물체의 굴절은 포함하여 Annotation을 진행 중.



7. 데이터 가공

● 7.1 Annotation 정보 생성

“6.1 이미지 정제” 단계를 통해 정제된 이미지 데이터의 Annotation 과정에 대한 단계로 1차 Poly Line Labeling작업을 라벨링 엔지니어를 통해 진행하고 Application을 통해 해당 이미지별 Tag정보를 Read하여 자동으로 Rectangle Labeling을 생성한다.

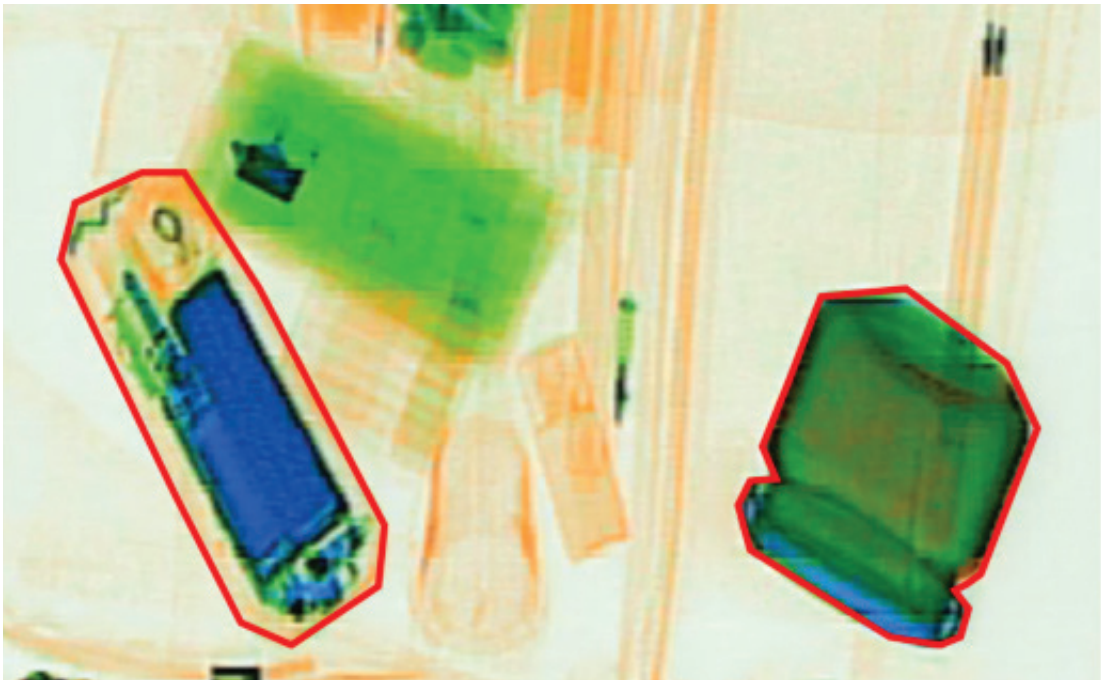
Poly Line Labeling		Rectangle Labeling
	>> Auto	
		
확대 보기		확대보기
	>> Auto	
라벨링 오류 케이스 • 물체의 경계를 정확하게 처리 하지 않음 • 정밀하게 물체의 라벨링 하지 않음		라벨링 오류 케이스 • 물체를 명확하게 바운딩 하지 않음 • 다른 위치에 바운딩 되어 있음

위 라벨링에 대한 물품 경계에 정확한 작업이 필요하여 적어도 2인 이상 동일한 라벨링을 통해 상호 교차 검증을 진행 한다.교차 검증은 소프트웨어 툴에 의해 진행이 되며 검증을 통한 레포트는 통합 관리 되어 관리자가 라벨링 검증을 검토 한다.

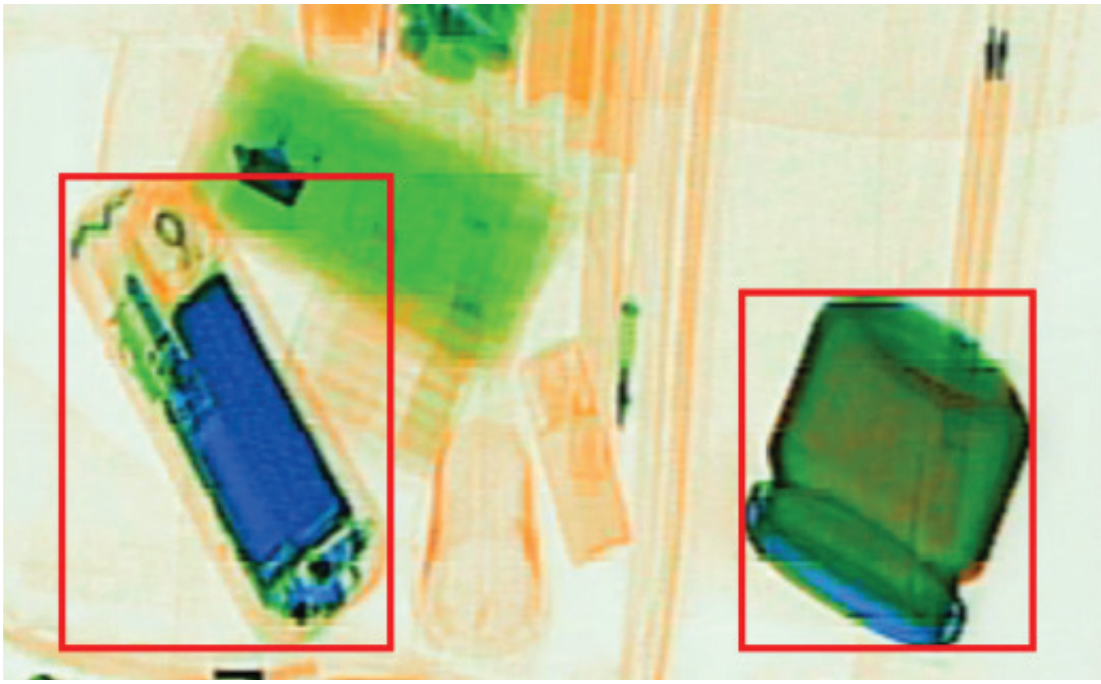
Rectangle Labeling 자동 변환 이후 Annotation 정보와 Image를 Over-Wrap하여 그 결과를 검토 한다. 해당 방법은 다음 장에 상세히 정리 한다.

7.2 Poly Line Labeling의 Rectangle Labeling 자동 변환

Annotation 자동 변환 절차 및 방법



1. Poly Line 라벨링의 Top(Min), Left(Min), Bottom(Max),Right(Max) 값을 Annotation 정보에서 Read 한다.
2. Top, Bottom은 수평으로 연장 Left, Right는 수직으로 연장하여 교차점을 생성



3. 수평, 수직 교차점에 대한 Rectagle 정보를 Annotation 정보로 저리 한다.
4. XML 형태의 데이터로 저장

● 7.3 데이터-셋 네이밍 규칙

○ 7.3.1 HS Code 정의

HS는 Harmonized System의 약자로서,국제통일상품분류체계에 따라 대외 무역거래 상품을 총괄적으로 분류한 품목분류 코드이다.국제협약에 따라 HS코드는 10자리까지 사용할 수 있으며, 6자리까지는 국제 공통으로 사용하는 코드이며 7자리부터는 각 나라에서 세분화하여 부여하는 숫자이다. 현재 한국은 10자리를 사용하고 있으며 관세통계통합품목분류표 (HSK)에 의하여 당해 물품의 관세율 및 수출입 통관시 이행하여야 할 요건을 확인할 수 있다.

○ 7.3.2 데이터 이미지 생성 규칙

데이터의 관리를 위해 이미지의 생성 규칙을 아래와 같이 정의 한다.

물품종류	물품	HS CODE	Sample Number	Image Sequence	Data-ID
위해물품(H)	탄약	9306.30-9000	001	0001	H_930630-9000_001_0001.PNG
위해물품(H)	가위	8213.00-1000	004	0002	H_821300-1000_004_0002.PNG
저장물품(S)	핸드폰	8517.70-1029	001	0011	S_851770-1029_001_0011.PNG
저장물품(S)	노트북	8471.30-0000	019	0455	S_847130-0000_019_0455.PNG

파일 처리시 오류 방지를 위해 HS-Code에서 “.”을 제거하여 사용한다. 위 표와 같이 HS-CODE를 기반으로 Sample Number, Image Sequence Number의 조합으로 유일한 데이터 네이밍을 생성한다. 특히 확장성을 고려하여 Sample Number는 3자리 조합으로 구성하면 Image Sequence Number도 4자리로 구성한다.

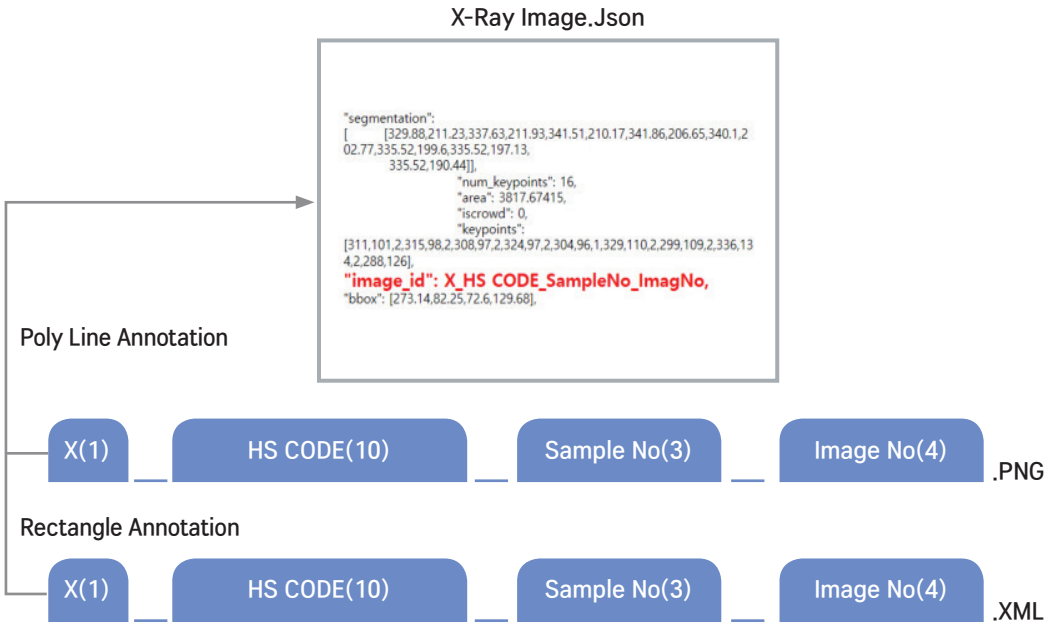
단일 품목 외 명명은 이미지 생성 규칙을 “_M”로 아래와 같이 정의한다.

물품종류	물품	HS CODE	Sample Number	Image Sequence	Data-ID
위해물품(H)	탄약	9306.30-9000	001	0001	H_930630-9000_001_0001_M.PNG
위해물품(H)	가위	8213.00-1000	004	0002	H_821300-1000_004_0002_M.PNG
저장물품(S)	핸드폰	8517.70-1029	001	0011	S_851770-1029_001_0011_M.PNG
저장물품(S)	노트북	8471.30-0000	019	0455	S_847130-0000_019_0455_M.PNG

품명	HS CODE	샘플수	품명	HS CODE	샘플수
탄약	9306.30-9000	15	핸드폰	8517.70-1029	15
폭축	9405.40-9000	25	USB	8544.42-2090	20
라이터	9613.80-0000	25	노트북	8471.30-0000	20
에어로졸	8481.80-1090	20	Tablet	8471.30-0000	10 (21~30)
휴대용가스	7321.11-0000	20	SSD	8523.51-1000	10
알코올	2905.12-2090	15	HDD	8471.70-2020	5
신나	3814.00-1090	15	외장하드	8471.70-2020	5(6~10)
라이터기름	9613.80-0000	15	끌	8205.30-0000	25
성냥	3605.00-9000	10	플라이어	8203.20-1000	20
고체연료	7321.89-0000	10	표창	9503.00-3493	5
건전지	8507.60-9000	15	전자담배(액상)	8543.70-4090	10
휴대용 배터리	8507.60-1000	30	전자담배(궐련)	8543.70-4090	10
손톱깎이	8214.20-0000	20			
칼	8211.91-0000	35			
톱	8202.39-2090	35			
가위	8213.00-1000	35			
송곳	8207.50-1090	25			
드라이버	8205.40-0000	25			
스패너	8204.20-0000	25			
액체(100ml이상)	3307.90-9000	25			
수갑	7326.90-9000	10			
도끼	8201.40-0000	20			
망치	8205.20-0000	20			
야구배트	9506.91-0000	10			
총기	8424.20-1000	20			
금속 파이프	7306.90-9000	25			
총기부품	8424.90-2000	25			

위 네이밍 정보를 기반으로 데이터를 생성한다.(단 노트북과 Tablet의 HS CODE가 동일하므로 Tablet의 Sample Number는 21번부터 시작, HDD와 외장하드의 HS CODE가 동일하므로 외장하드의 Sample Number를 6번부터 시작)

위 정의된 네이밍 규칙에 맞게 아래와 같이 어노테이션 정보를 생성한다.



● 7.4 데이터-셋 Ground-Truth 정의

아래 X-Ray 위해 물품에 모델 및 네트워크 적용에 대한 검지율에 대한 논문에 따르면

Model	Network	mAP	camera	laptop	gun	gun component	knife	ceramic knife
SWCNN	AlexNet	0.608	0.682	0.609	0.748	0.714	0.212	0.683
	VGG_M	0.634	0.707	0.637	0.763	0.731	0.246	0.719
	VGG_{16}	0.649	0.701	0.724	0.752	0.757	0.223	0.734
	$ResNet_{50}$	0.671	0.692	0.801	0.747	0.761	0.314	0.713
	$ResNet_{101}$	0.776	0.881	0.902	0.831	0.848	0.392	0.803
RCNN	AlexNet	0.647	0.791	0.815	0.853	0.582	0.188	0.658
	VGG_M	0.686	0.799	0.855	0.869	0.658	0.210	0.723
	VGG_{16}	0.779	0.888	**0.954**	0.876	0.832	0.304	0.819
F-RCNN	AlexNet	0.788	0.893	0.756	0.914	0.874	0.467	0.823
	VGG_M	0.823	**0.900**	0.834	0.918	0.875	0.542	0.869
	VGG_{16}	0.883	0.881	0.918	0.927	**0.938**	0.721	0.912
	$ResNet_{50}$	0.851	0.844	0.879	0.916	0.901	0.677	0.889
	$ResNet_{101}$	0.874	0.857	0.904	0.931	0.911	**0.732**	0.907
R-FCN	$ResNet_{50}$	0.846	0.894	0.928	0.932	0.918	0.506	0.896
	$ResNet_{101}$	0.856	0.887	0.906	0.942	0.925	0.556	**0.920**
YOLOv2	$Darknet_{288}$	0.810	0.821	0.861	0.914	0.904	0.551	0.814
	$Darknet_{416}$	0.851	0.888	0.883	**0.952**	0.924	0.605	0.851
	$Darknet_{544}$	**0.885**	0.896	0.894	0.943	0.933	0.728	0.913

TABLE V: Detection results of SW-CNN, Fast-RCNN (F-RCNN) [47], Faster RCNN (F-RCNN) [11], R-FCN [12] and YOLOv2 [13] for multi-class problem (300 region proposals). Class names indicates corresponding average precision (AP) of each class, and mAP indicates mean average precision of the classes.

다양한 모델 및 네트워크에 따라 인식율에 대한 차이가 일어 날 수 있습니다.

* 물품별 인식율은 Knife 0.188 ~ 0.728에 대한 편차가 있음. 실제 검지율이 높을수록 처리 속도에 대한 시간이 상품화 제약이 될 수 있는 부분이 있음 당 데이터 셋은 위 모델을 기반으로 mAP 0.500 ~ 0.700 정도의 인식에 대한 수치를 기준으로 진행
* 해당 수치는 일반적인 모델과 네트워크의 최고 인식율에 대한 평균 값으로 선정

8. 데이터 검증

● 8.1 이미지 검증

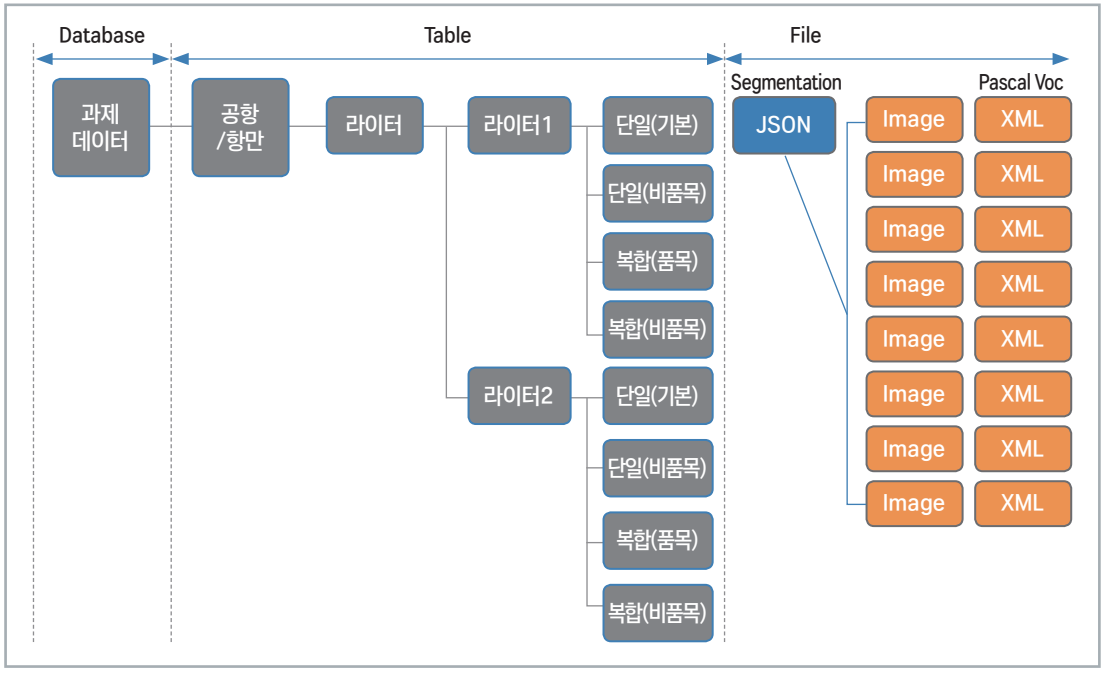
위해물품 엑스레이 데이터의 경우 공항/항만 등 보안의 문제로 외부에 공개가 불가능한 데이터라는 점에서 실데이터 구현에 대한 현업 프로세스 이해가 중요.이러한 문제점 해결을 위해 현업 프로세스 및 규정에 대한 사전 분석 및 보안 검색 업무 경력자를 채용하여 Know-How를 실데이터 구현에 반영. 항공보안 교육원(KASA)에서 인증하는 X-Ray 판독 교육 과정을 이수하고 해당 업무 5년 이상 경험자를 통한 실데이터화는 매우 중요한 사항입니다.

- X-Ray 단독 교육 과정 설명(KASA)

교과목	주요 교육 내용	교육시간(16시간)	비고
폭발물 등 위험물 식별 및 처리	• 폭발물의 이해 • 폭발물 등 의심물품 처리절차	2H	
보안검색실무	• 검색절차 : 검색장비 및 검색 요령 • 금속탐지기, 촉수검색 등 승객검색 • X-RAY 등 검색장비 운영 및 테스트	14H	
	• 위해물품 식별 • X-RAY 판독 시뮬레이션 훈련 (판독) • 테스트 등		

해당 업무는 폭발물 및 위험물 식별 및 처리를 위한 특별 교육으로 보안 검색 업무를 위해 필수 이수 항목입니다.

실제 데이터셋의 물리적인 구조는 아래와 같이 구성됨.



정보 활용자는 인증을 거쳐 시스템에 접속을 하게 되며 권한에 따른 검색 허가를 통해 간단한 조건을 기입하여 서비스 하며 내부적으로는 SQL Query를 통해 원하는 데이터제공 받을 수 있게 처리 합니다. 특히 Table의 상위를 선택하여 하위 데이터를 모두 가져온다거나 하위의 데이터를 선택 적으로 사용할 수 있는 UX적인 편의성도 제공 되어야 할 것입니다.
추가적으로 해당 데이터 구축 및 공개 이후 전담기관의 공개 정책에 따라 정보 활용자의 데이터 셋 추가 업무의 권한이 이루어 진다면 현 시스템에서 고도화 시킬 수 있는 빅데이터 구조로 전환이 용이하게 구성할 예정입니다.

● 8.2 학습데이터 검증

Dataset의 유효성은 실제 분석 알고리즘에 따라 매우 다를 수 있다. 이는 실제로 물체를 탐지하기 위한 알고리즘과 classification를 구분해야 할 수 있으며, 오픈된 자료에서는 아래와 같이 목록을 추릴 수 있다. (처리속도를 감안 하지 않은 아키텍쳐일 경우)

1) Detection 데이터 검증 학습 네트워크 : Faster RCNN
2) Classification Network : atrous convolution을 이용한 inception + resnet 모델 or Nasnet

일반적인 알고리즘을 사용 하더라도 각 알고리즘에 맞는 하이퍼 파라미터의 튜닝이 필요하다.

● 8.3 데이터 셋 검수 프로세스 및 작업

해당 데이터셋을 생성하고 정확도에 대한 검수를 위해 아래와 같은 3가지 Step의 작업을 진행하여 Ground Truth를 정의 한다.

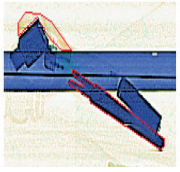
1) 자가 검증

- 작업자의 어노테이션을 자가 검증 하는 단계로 아래의 물품별 스텐다드를 기반으로 작업자의 육안 검증

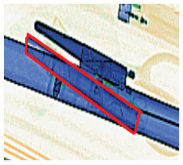
학습을 위한 어노테이션 가이드

Unknown (특징이 가리는 경우)

Unknown (중간이 잘리는 경우)



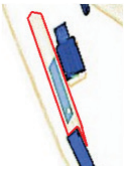
(2090_07_168) 겹쳤지만 특징이 많이 다르지 않는 경우



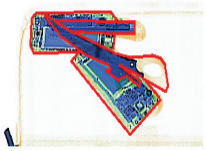
(2090_15_193) 겹쳤지만 특징 손실이 거의 없는 경우

2090_16_133(중간이 잘리는 경우)

Unknown 특징이 가리는 경우



S_8517.70-1029_01_164



2) 어노테이션 검증 전문 인원

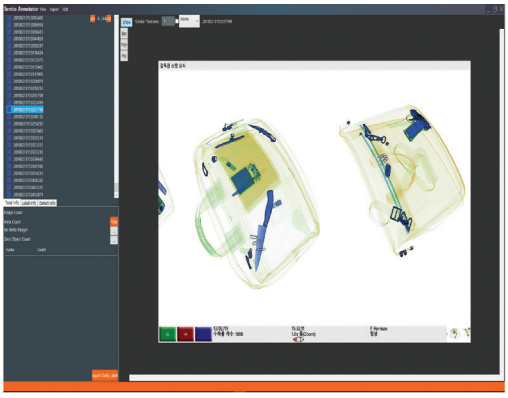
- 장비별 어노테이션 검증 인원을 통한 개인 작업자 어노테이션 확인 진행 주요 체크 내용
  - 잘못된 이름으로 Annotation이 진행 되는 경우
  - 중복 ID가 존재하는 경우
  - Annotation 누락 상황
  - Annotation 실수

  어노테이션 검증인원은 해당 사항을 체크하여 수정 작업하여 서버로 전송

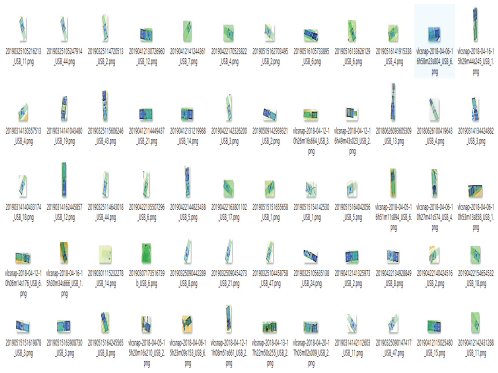
3) 데이터셋 관리자 검증

- 2차 검증이 완료된 데이터를 대상으로 물품번호 및 Annotation 데이터를 이용하여 원본 이미지에서 Crop하여 이미지를 복사 해당 폴더에 위 가이드에 맞지 않는 데이터를 최종 검증 하여 수정

어노테이션을 통한 물체 Crop



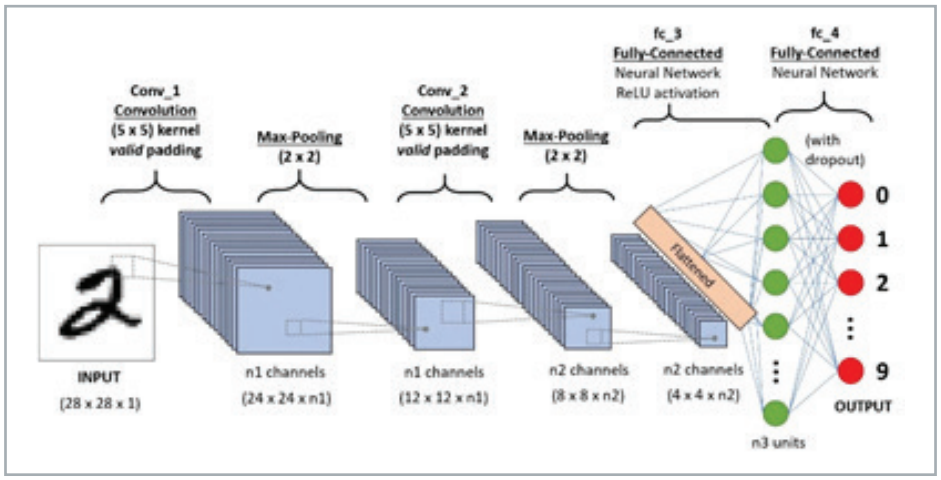
Crop된 물품 이미지셋



● 8.4 검증 소프트웨어에 및 검증 정책

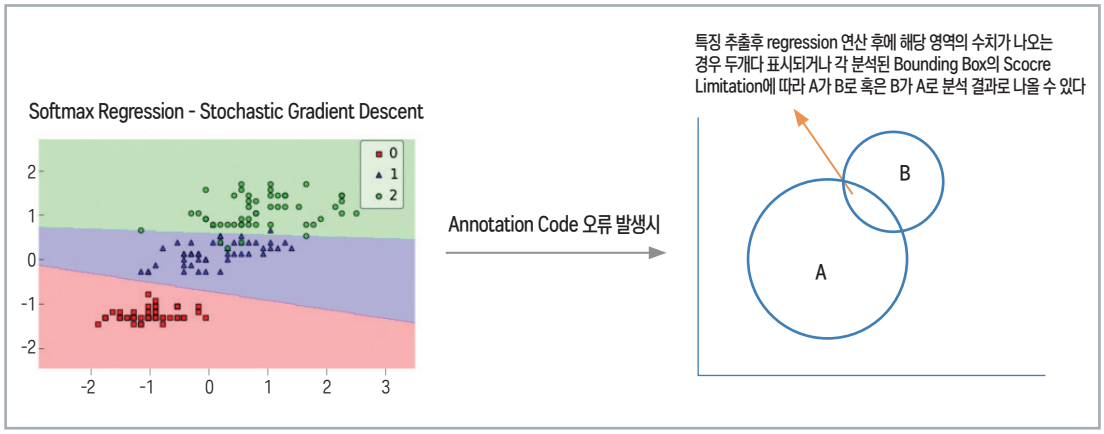
검증 정책은 어노테이션 유의 사항 3가지를 기반으로 정의 될 수 있다.

○ 8.4.1 최대한 분류하고자 하는 품목에 대한 간섭 배제



5x5 임의의 수로 이뤄진 Filter를 Sliding Window로 활용하여 이미지 특징 생성 Convolution은 기본적으로 nxn 사각형을 기준으로 하며 이를 Regression(분류를 위한연산)을 한다. 결국 Annotation 할때 주변 특정 물체의 주변 이미지는 Noise로 간주 될 수 있으므로 최대한 물체에 Noise가 없을 수 있도록 Annotation 하는 것이 분류 및 특징 추출에 유리하다.
단 Regression 연산과 Convolution Layer의 Activation Function에 따라 이는 변경 될 수 있다.

○ 8.4.2. Object의 타입코드 동일화(Annotation 정보와 실제 Image에 Annotation)



이는 위 8.4.1번 내용의 Convolution과 Regression시에 매우 영향을 많이 끼치는 부분이며, 오탐을 매우 많이 발생 시킬 수 있다. 서로 다른 A , B라는 물체가 있고, A와 B를 분류 하는 모델 및 데이터를 구축함에 있어 특징 추출및 분류시에 각 물체에 데이터적 특징이 겹치게 되면, 분류 수식 연산에서 내용이 겹칠 수 있다.

○ 8.4.3. 다양한 데이터-셋 보유

한가지 물체는 각도 및 빛에 따라 다양한 형상을 가질 수 있으므로, 최대한 다양한 데이터를 만드는 것이 좋다. Activation Function및 Regression에 따라 짤려있는 Object에 대한 Annotation도 훌륭한 데이터가 될 수 있다. (concatenated ReLU 등)

○ 8.4.4 IOU 설정

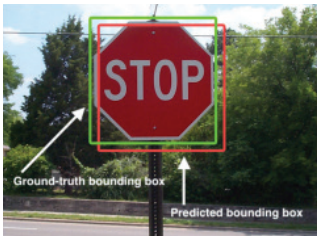
두개 이상의 영역에서 겹침 영역 이는 학습시에 예측된 Box와 실제 Annotation된 영역 즉 Ground Truth와의 겹침 영역을 검토한다.

기본적으로 Data에 대한 학습은 구축한 모델의 Layer(임의의 수치로 구성된)를 지나며, 마지막 결과와 최종 원본과의 데이터를 비교하고, 그 변경된 수치를 Backpropagation을 통해 임의의 수치로 구성된 각종 Layer의 가중치, 각 뉴런 노드의 가중치를 수정 (Regression)하며 학습한다.

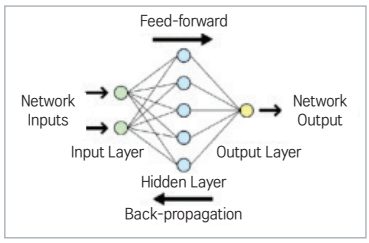
즉 IOU는 예측한 영역과 실제 Annotation된 영역의 비교를 통해 모델의 성능 지표를 만들며 물체 검색시에 얼마나 정답에 맞게 좌표값이 예측 되었는지 알수 있는 지표다.

따라서 보통 0.6정도를 사용 하는 것으로 알고 있고, 이 또한 여러 기술에 따라 달라 질 수 있다.

(Had Negative Mining 등)



IOU



Backpropagation

데이터-셋을 오픈 하는 것은 모델의 설계와 각종 산술 연산식의 설계를 가리기 위함이기 때문에 가장 중요한 8.4.2번 항목을 제외 하고는 8.4.1번과 8.4.3번 항목에 대해서는 명확한 기준을 가리기 힘들며, 각종 산술식의 임계치 설정 등에 따른다.

오픈된 데이터-셋을 이용하여 연구하는 경우 8.4.1번과 8.4.2번에 중점을 두고 Annotation 해야 한다. 만약 이를 상업적인 용도로 사용한다고 하면 그에 맞게 각 연구소에서 Annotation 정보를 설계된 모델과 산술 연산식에 맞게 분석하여 Annotation을 수정해야 한다.

9. 데이터-셋 활용

● 9.1 데이터 셋 제공 및 활용

AI용 Xray Scan 데이터는 현재 공공 차원에서 수집 및 활용을 하지 않는 것으로 유추 할 수 있습니다. 실제 X-Ray제조사의 Raw-Data를 외부에 공개하지 않고 있으며 파일 포맷도 제조사 별로 상이하기 때문에 표준화 되어 있는 파일이 아닙니다. 당사는 이번 제안을 통해서 표준화 및 공개를 통해 해당 분야의 사업의 활성화를 위해 기여 할 것입니다.

분류	내용
데이터 표준	• PNG File 포맷으로 표준화 • Full HD 해상도 데이터
데이터 공개	• 정보 공개 가능 기업 및 기관 전체 (전담 기관 인증을 통한 공개)
구축	• 당 과제 수행 후 시범서비스 구축 • 수요처 데이터 공개
보유기관	• 주관기관 1년간 서비스 운영 (전담기관 보유, 주관기관 운영/유지관리)
활용	• 관공서 및 기업 • AI 판독 시스템 개발 기업 등.

추가 활용 서비스 종류 및 예시

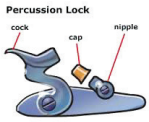
- 식품안전 분야 : 식품 제조 공정에서의 투과용 X-Ray상의 이물질 판독 서비스
  특정 원재료 탐지 서비스
- 제조산업 분야 : 비정상 제조 제품의 특성 판독 서비스

부록 I

항공기내 반입 금지 위해 물품 목록

물품명	객실 반입	위탁 수하물 반입	비 고
❶ 무기류 ※ 도검·무술호신용품·총기 등 무기류는 객실 반입금지, 위탁수하물 반입가능			
1. 도검류 가. 무술용 검, 펜싱용 검, 칼이 든 지팡이 등 검류  나. 부엌칼, 과도, 다용도칼, 사냥칼    다. 접이식칼, 맥가이버칼, 박스커터칼, 학용품용 칼    라. 면도칼, 외과용 메스, 식사용 나이프 등   	X	O	단, 플라스틱 칼, 둥근 날을 가진 버터칼, 안전날이 포함된 면도기, 안전면도날, 전기면도기 및 기내식 전용 나이프(항공사 소유에 한함)는 객실 반입가능   
2. 창류 가. 창, 작살, 작살총, 표창    나. 송곳날이 있는 다트, 새총, 투석기 등   	X	O	

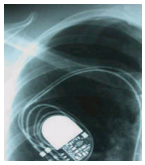
물품명	객실 반입	위탁 수하물 반입	비 고
3. 무술·격투용 무기류 (쌍절곤, 손가락에 끼워 공격하는 무기, 열쇠고리형 호신무기 등) 	X	O	
4. 곤봉·경찰봉 등 곤봉류 및 수갑류 	X	O	단, 범죄인 호송 등 공무 목적의 수갑류는 객실 반입가능
5. 호신용 스프레이류 (후추·고추·산성스프레이, 야생동물퇴치용 스프레이 등) 	X	O	단, 위탁수하물로 반입할 경우, 1인당 1개(100㎖이하)까지만 가능
6. 총기류 (권총, 연발권총, 엽총, 소총, 산탄총, BB총, 압축공기총, 라이플총, 볼베어링총 등) 	X	O	단, 위탁수하물로 반입할 경우, 해당 항공운송사업자에게 총기소지허가서 또는 수출입허가서 등 관련서류를 확인시키고, 총알과 분리한 후, 단단한 보관함에 넣은 경우에만 가능
7. 전자충격기(stun gun), 테이저건 	X	O	
8. 총기류 부품 	X	O	단, 총기류 부품 중 조준경은 객실 및 위탁수하물 반입가능
9. 발광/화염 신호총(flare gun), 출발 신호용 총(starter pistole) 	X	O	
10. 복제·모의 총기류 및 장난감 총 	X	O	

물품명	객실 반입	위탁 수하물 반입	비 고
11. 총알 (실탄, 공포탄 등) 	X	O	단, 위탁수하물로 반입할 경우, 해당 항공운송사업자의 승인이 필요하며, 「항공위험물 운송기술기준」(국토교통부 고시)에 적합한 경우에만 가능
❷ 폭발물류 ※ 폭발물류는 객실 및 위탁수하물 모두 반입금지			
1. 폭발물, 폭발장치 가. 다이너마이트, 수류탄, 지뢰 등 나. 뇌관(blasting caps), 기폭장치류(denotes), 신관, 도화선(fuses), 발파캡, 격발뇌관(percussion cap) 다. 화약류    라. 복제·모의 폭발물 또는 폭발장치 등 	X	X	
2. 불꽃·화염류 및 폭죽류 가. 조명탄, 신호탄, 화염탄 등  나. 불꽃놀이 및 장남감용 폭죽   다. 폭죽소리가 나는 장난감 총 등 	X	X	
3. 연막탄(smoke grenade) 및 연기를 발생시키는 캔이나 카트리지  	X	X	

물품명	객실 반입	위탁 수하물 반입	비 고
❸ 공구 및 생활용품류 ※ 흉기로 사용될 수 있는 공구류는 객실 반입금지, 위탁수하물 반입가능			
1. 공구류 가. 도끼, 손도끼, 얼음도끼, 망치, 장도리, 쇠지레 	X	O	
나. 산업용 못총·볼트총, 송곳, 얼음송곳, 휴대용 전동톱 등 톱류, 전동드릴 등 드릴류 	X	O	
다. 총길이 10㎝를 초과하는 렌치, 스패너, 펜치, 프라이어류 	X	O	
라. 날·심·끌의 길이가 6㎝를 초과하는 스크루드라이버, 드릴심, 끌, 정 및 가위류(조경용 가위 포함) 	X	O	
2. 전자충격용 봉 등 가축몰이용 봉 또는 막대, 동물도축용·살상용 도구 	X	O	
3. 휴대용 일반 소형 배터리 	O	O	
4. 리튬배터리 또는 리튬이온배터리가 포함된 개인용 휴대 전자장비 (시계, 계산기, 카메라, 캠코더, 휴대폰, 노트북컴퓨터, MP3 등) 	O	O	

물품명	객실 반입	위탁 수하물 반입	비 고
5. 일반 생활도구류 가. 금속제 포크·젓가락·숟가락 나. 끝이 둥근 금속제 병따개, 와인용 코르크 따개  다. 요리용 다지기, 감자칼, 무채칼, 믹스기 칼날  라. 손톱깍이, 손톱 줄, 족집게, 눈썹정리용 칼, 손톱정리용 가위  마. 대바늘, 코바늘, 바느질용 바늘 바. 제도용 콤파스 등 	○	○	
6. 액체·분무·겔 형태의 위생용품, 욕실용품 또는 의약품류 (화장품, 향수, 매니큐어, 매니큐어 리무버, 헤어스프레이, 헤어무스, 헤어젤, 염색약, 퍼머약, 방향제, 냄새제거제, 면도크림, 샴푸, 린스, 샤워젤, 치약, 콘택트렌즈용액, 몸에 뿌리는 벌레 기피제, 소염제, 의료용 소독 알콜, 내복약, 외용연고 등) 	○	○	단, 객실 반입(국제선에 한함)할 경우, 개별 용기당 100㎖ 이하로 1인당 1ℓ 비닐지퍼백 1개에 한해 반입이 가능하며, 의약품인 경우에는 의사 처방전 등 관련 증명서를 보안검색요원에게 제시하고 적정하다고 판단될 경우 비행중 필요한 용량에 한해 객실 반입가능 위탁수하물(국제선에 한함)로 반입할 경우에는, 개별 용기당 500㎖(0.5kg)이하로 1인당 2kg(2ℓ) 이하까지만 가능
❹ 스포츠 및 레저용품류 ※ 무기로 사용될 수 있는 스포츠용품은 객실 반입금지, 위탁수하물 반입가능			
1. 방망이, 노, 봉, 검, 날, 촉 형태의 스포츠 장비 가. 야구·소프트볼·크리켓* 등 배트류 나. 카약·카누·보트 등 노류 			

물품명	객실 반입	위탁 수하물 반입	비 고
다. 하키·라코르스** 등 스틱류 라. 골프채, 당구큐 등    마. 목검류, 볼링공, 아령 등 바. 빙상용 스케이트    사. 활, 화살, 양궁, 석궁, 국궁 등 활류  	X	O	단, 테니스·배드민턴·스쿼시·탁구 등 라켓류 및 인라인스케이트, 스케이트보드는 객실 반입가능
2. 공기가 1/3이상 주입된 축구공 등 스포츠용 공류 및 풍선류  	X	X	단, 공기가 주입되지 않는 야구공·골프공 등의 경우 객실 또는 위탁수하물 반입가능
3. 등산 장비 (등산용 스틱, 텐트 폴, 등반용 못, 고리, 아이젠 등)    	O	O	단, 산악용 망치는 객실 반입금지, 위탁수하물 반입가능 
❺ 의료·구조용 물품 ※ 의료용 물품은 객실 및 위탁수하물 모두 반입가능하나, 일부 위험가능 물품은 제한될 수 있음			
1. 주사바늘, 한방용 침류, 수지침  	O	O	
2. 소형 의료용 수은체온계  	O	O	단, 객실 또는 위탁수하물로 반입할 경우, 보호케이스에 담은 경우에만 가능

물품명	객실 반입	위탁 수하물 반입	비 고
3. 보행 보조도구 가. 노약자·장애인·환자가 사용하는 지팡이, 목발 등  	O	O	
나. 노약자·장애인·환자·유아가 사용하는 휠체어, 유모차  	O	O	단, 휠체어 및 유모차를 객실에 반입할 경우, 해당 항공운송사업자의 승인이 필요
다. 인공수족에 포함된 비인화성 가스 실린더 (여행기간 동안 필요한 여분의 실린더 포함)  	O	O	단, 객실 또는 위탁수하물로 반입할 경우, 해당 항공운송사업자의 승인이 필요
라. 장애인·노약자·환자가 사용하는 전동휠체어 또는 전동이동 보조기구(배터리 포함)   	X	O	단, 위탁수하물로 반입할 경우, 해당 항공운송사업자의 승인이 필요하며, 「항공위험물 운송기술기준」(국토교통부 고시)에 적합한 경우에만 가능
4. 휴대용 전자의료장비 (자동제세동기, 인공호흡기, 호흡기 치료기, 천식용 호흡기 등)  	O	O	단, 여분 배터리는 위탁수하물로 반입이 금지되며, 보관함에 안전하게 넣은 경우에 한해 2개까지만 객실반입 가능
5. 인체에 이식된 인공심박기, 이식 전자의료장치  	O	-	
6. 의료용으로 사용되는 5kg이하의 소형 산소통 또는 공기실린더 	O	O	단, 객실 또는 위탁수하물로 반입할 경우, 해당 항공운송사업자의 승인이 필요하며, 「항공위험물 운송기술기준」(국토교통부 고시)에 적합한 경우에만 가능

물품명	객실 반입	위탁 수하물 반입	비 고
7. 구명조끼에 포함된 비인화성·무독성 이산화탄소가스 실린더 1쌍 (여분 실린더 1쌍 포함)  	O	O	
8. 눈사태용 구조배낭(1인당 1개) 	O	O	
❻ 기타 인화성·화학성·유독성 물질 ※ 인화성물질 등은 원칙상 객실 및 위탁수하물 반입이 모두 금지되나, 『항공위험물 운송기술기준』(국토교통부 고시)에서 정한 안전요건을 충족하는 경우에 한해 가능			
1. 성냥 또는 라이터 가. 소형 안전성냥(safety match) (1인당 1개) 	O	X	단, 객실 반입할 경우, 자체발화 방지를 위하여 여행용가방 등에 보관하지 말고 직접 소지할 것
나. 딱성냥(마찰성냥) (strike anywhere match), 대형 곽성냥 * 아무데나 마찰시 발화되는 성냥  	X	X	
다. 휴대용 담배라이터(1인당 1개)  	O	X	단, 객실 반입할 경우, 자체발화 방지를 위하여 여행용가방 등에 보관하지 말고 직접 소지할 것 단, 선물용으로 포장되고 연료가 포함되어 있지 않은 라이터의 경우에는 위탁수하물로 반입가능
라. 총기모양 라이터  	X	O	단, 위탁수하물로 반입할 경우, 선물용으로 포장되고 연료가 포함되어 있지 않은 라이터에 한해 가능
마. 토치 라이터 및 비흡수성 액체연료로 작동하는 담배라이터 	X	X	

물품명	객실 반입	위탁 수하물 반입	비 고
2. 인화성 가스·인화성 액체 가. 부탄가스, 프로판가스, 탄화수소카트리지, 석유, 휘발유, 등유, 경유, 에탄올, 라이터 연료 등    나. 스프레이페인트, 인화성페인트, 인화성 광택도료, 페인트시너(Paint Thinner), 테레빈유(Turpentine) 등    다. 인화성 살충제, 인화성 다리미풀 등  	X	X	
3. 탄화수소가스(Hydrocarbon gas)로 작동하는 헤어컬 (1인당 1개) 	O	O	단, 가열부분에 안전커버가 씌워져 있는 경우에만 반입 가능하며, 여분의 가스리필은 객실 및 위탁수하물 반입이 모두 금지 
4. 자체적으로 연소·불꽃·폭발이 일어날 수 있는 물질	X	X	
5. 강한 열을 발생시키는 배터리로 작동하는 납땜장치(납땜인두), 수중랜턴(수중다이빙용 램프) 	O	X	단, 객실 반입할 경우, 반드시 배터리와 분리할 것
6. 기상용 수은기압계 및 수은온도계  	O	X	단, 공공용 목적으로 사용되는 물품으로 해당 항공운송사업자의 승인을 받고, 수은이 유출되지 않도록 내부가 봉인된 단단한 보관함에 담은 경우에만 객실반입 가능
7. 알코올성 음료 가. 70도(%) 이상의 알코올성 음료	X	X	
나. 24도(%) 이상 70도(%)미만 알코올성 음료 (1인당 5ℓ에 한함) 다. 24도(%) 미만 알코올성 음료	O	O	

물품명	객실 반입	위탁 수하물 반입	비 고
8. 최루가스(tear gas)	X	X	
9. 독극물·쥐약·농약 등 독성물질	X	X	
10. 방사능 물질 (의료용·상업용 동위원소)	X	X	
11. 부식성, 표백성, 산화성 물질 (염소, 표백제, 산화제, 수은, 하수구 청소제제 등)	X	X	
12. 전염성·생물학적 위험물질	X	X	
13. 소화기 	X	X	
14. 드라이아이스 (1인당 2.5kg) 	O	O	단, 객실 또는 위탁수하물로 반입할 경우, 해당 항공운송사업자의 승인이 필요하며, 「항공위험물운송기술기준」(국토교통부 고시)에 적합한 경우에만 가능

1. 상기 별표에서 정하지 아니한 물품 또는 반입을 허용하는 물품이더라도 해당 공항의 보안검색감독자 또는 항공운송사업자가 항공기 안전 및 승객·승무원에게 위해(危害)를 줄 수 있다고 판단하는 경우에는 항공기 내 반입을 금지할 수 있다.
2. 상기 별표에서 정한 물품 중 「항공위험물 운송기술기준」(국토교통부 고시) 제210조에서 정한 위험물에 대해서는 해당 고시에서 정한 방식에 따라 반입·운송하여야 하며, 그 외의 항공위험물에 대해서는 해당 고시에 따라 객실 또는 위탁수하물로 반입해서는 아니된다.
3. 항공기 내로 반입코자 하는 물질이 화학성·유독성 물질인 경우에는 승객이 「산업안전보건법」에 따른 물질안전보건자료를 해당 항공운송사업자에게 제시하고 안전한 물질로 판단되는 경우에 한해 위탁수하물로 반입이 가능하다.
4. 국제선 항공기 객실 내에 반입되는 액체·분무·겔류의 경우에는 「액체·분무·겔류 등 항공기내 휴대반입금지 물질 운영기준」(국토교통부 고시)에서 정한 기준에 적합한 경우에만 객실반입이 가능하다.

09

질병 진단 이미지 AI 데이터 구축 가이드라인

국립암센터
NATIONAL CANCER CENTER

	성명	소속	직위
가이드라인 제안	황보율	국립암센터	팀장
가이드라인 작성	조용연	국립암센터	연구원
	박현우	국립암센터	연구원
	이재혁	국립암센터	연구원
	정민영	국립암센터	연구원
	신현아	건양대병원	연구원
전문가 검토 위원	황영배	KETI	팀장
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 국립암센터/ 참여: 필라테크, 루닛, 헬쓰허브, 건양대병원, 인피니그루, AIMS

서 문

1. 가이드라인의 목적

방대하게 수집된 의료 데이터를 체계적으로 분석하여 질병의 발병 후 치료뿐만 아니라 질병을 예측할 수 있는 인공 지능 기술이 반드시 필요한 시점에 다가오고 있다. 반면, 우리나라는 아직 인공지능 벤처기업의 활동이 미약하고 인공지능기반 기술 및 데이터 활용이 제한적이며 산업적 기반도 부족한 상황이다. 따라서 본 표준은 공공 및 민간의 지능정보기술 개발 촉진과 지능정보산업 육성을 위한 의료 분야에 인공지능 학습용 데이터(지식베이스) 구축 및 이용환경을 제공하며, 이를 적용하기 위한 가이드라인을 제공하기 위하여 작성되었다.

2. 주요 내용 요약

본 표준은 임상적 활용을 위해 모델을 공개하고 활용 방법에 대한 지침을 서술한다. 학습된 모델의 임상적 활용을 위한 모델을 공개하여 수집된 데이터에 관한 기술과 활용 방법에 대한 예시를 제공하고, 모델 강화를 위해 추가 데이터를 가공하기 위한 저작도구를 공개하여 시스템 환경 구성방안을 제시한다.

3. 인용 문서와의 비교

- 해당 사항 없음

---

1 Purpose of Standard

Artificial intelligence technology that can predict disease as well as treatment after the onset of disease is coming to a point where systematic analysis of massively collected medical data is necessary. In Korea, on the other hand, activities of artificial intelligence venture companies are still weak, artificial intelligence based technology and data utilization are limited, and the industrial base is insufficient. Therefore, the standard provides the environment for building and using artificial intelligence learning data (knowledge base) in the medical field for promotion of public and private intelligence information technology development and intelligent information industry development, and to provide guidelines for applying it .

2 Summary of Contents

The standard describes the model for clinical use and provides guidance on how to use it. It provides an example of how to describe and utilize the data collected through model disclosure for the clinical use of the learned model, and We present authoring tool to process additional data to suggest system environment configuration for model enhancement.

3 Relationship to Reference Standards

None.

인공지능 기술 육성 및 활성화를 위한 의료영상 공개사업 가이드라인

Guidelines for medical image data open project for developing and advancing artificial intelligence technology

09

1. 적용 범위

암은 2011년 이후 국내에서 발생이 감소하는 추세를 보이고 있지만 유독 여성에게 발생하는 유방암만은 지속적으로 뚜렷한 증가하는 추세를 보인다. 유방암의 증가는 전 세계적인 현상으로 세계암통계에 의하면 연령표준 발생률에서 남녀를 모두 합쳐서도 유방암이 가장 많이 발생하는 암종이며, 이를 토대로 국내의 유방암의 증가는 앞으로도 지속될 것으로 판단된다.

또한, 2006년 진행한 메타 분석 연구에 따르면, 미국에서 각종 질병에 의해 발생된 실명으로 인한 경제적 손실은 연간 35억 달러에 이른다고 추정하고 있으며, 눈과 관련된 다양한 연구 중에 최근 가장 주목받고 있는 것이 바로 안저영상이다. 이러한 질병으로 인한 사회 경제적 손실을 줄이고, 환자의 건강상태를 증진하기 위해서는 관련 질환의 예방과 조기진단이 매우 중요하다.

본 표준의 적용은 유방암과 안저질환에 대한 의료 지식베이스 구축 및 효용성 검증으로 제한한다. 두 질환에 한하여 의료 인공지능 판독 기술 개발에 공통적으로 활용 가능한 이미지, 영상 등 범용성이 높은 의료 지식베이스 구축·공개하고 이에 대한 의료 지식베이스의 표준 지침을 서술한다.

2. 인용 표준

- 해당 사항 없음.

3. 용어 정의

3.1 어노테이션(Annotation)

- 주석처럼 코드에 달아 해당 코드에 특수한 의미를 부여해주는 데이터의 전처리 과정을 뜻하며, 이미지 데이터의 경우 이미지 내 각종 사물을 알아보고 경계선을 구분짓는 것

3.2 비식별화

- 개인정보가 포함된 데이터를 활용할 때, 정보집합물(데이터 셋)에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법

— 가명처리(Pseudonymisation): 개인정보 중 주요 식별요소를 다른 값으로 대체

— 총계처리 혹은 대체(Aggregation, Replacement) : 데이터의 총합 값을 보임으로서 개별 데이터의 값을 보이지 않도록 함

— 데이터 값(가치) 삭제(Data Reduction): 데이터 공유, 개방 목적에 따라 데이터 셋에 구성된 값 중에 필요없는 값 또는 개인식별에 중요한 값을 삭제

— 범주화(Data Suppression): 데이터의 값을 범주의 값으로 변환하여 명확한 값을 감춤

— 데이터 마스킹(Data masking): 공개된 정보 등과 결합하여 개인을 식별하는데 기여할 확률이 높은 주요 개인식별자가 보이지 않도록 처리하여 개인을 식별하지 못하도록 함

3.3 데이터 증강

- 학습 시 필요한 훈련 데이터의 양을 늘려 과적합(Overfitting)의 문제를 피하는 방법으로 기존의 데이터에 대한 특징은 살린 상태에서 인위적인 변경을 통해 훈련 데이터의 양을 증폭

— 유방촬영영상 및 안저촬영영상 자동 판독 모델을 구축하는 과정에서 최적화의 대상이 되는 모델 파라미터를 충분히 학습시키고, 특정 데이터(영상)에 종속되지 않게 판독 모델을 생성하기 위한 학습 데이터를 제공하기 위함

— 대표적인 방법

(1) 픽셀 수준에서의 데이터 증강 (photometric augmentation) ex) 밝기 조정

(2) 좌표계 수준에서의 데이터 증강 (affine transformation) ex) 기울이기

(3) 관심영역 (region-of-interest) 수준에서의 데이터 증강

3.4 안저

- 동공을 통해 안구의 안쪽을 들여다보았을 때 보이는 부분. 즉 망막 및 망막혈관 시신경유두 맥락막(포도막을 구성하는 막)을 종합하여 이르는 용어

— 안저 질환으로는 당뇨망막병증, 황반변성, 망막혈관폐쇄 등이 존재

3.5 PACS (Picture Archiving and Communication System; 의료영상저장전송 시스템)

- CT, MRI 같은 디지털 의료영상 장비를 사용하여 획득된 디지털 의료영상이미지를 DICOM (Digital Imaging and Communications in Medicine)이라는 국제표준규약에 맞게 저장, 가공 저장하고 네트워크를 통해 병원 내·외 단말로 전송이 가능한 시스템

3.6 Core biopsy (핵생검; 조직검사)

- 피부에 부분 마취를 하고 병변에 바늘을 여러 번 질러 조직 혹은 세포를 채취하는 시술법으로, 세포흡입보다 정확하며 입원을 하지 않고 시행할 수 있는 장점

3.7 BI-RADS (Breast Imaging Reporting and Data System)

- 디지털 유방촬영술에서의 유방암 이미지에서 유방암 단계를 분류별(category)로 표현

— BIRADS Ⅰ: 정상

— BIRADS Ⅱ: 양성으로 판명하며, 양성이 될 가능성이 100%

— BIRADS Ⅲ: 일반적으로 양성 또는 양성일 가능성으로 분류되며, 양성이 가능성이 매우 높지만, 악성 종양 위험은 0.3-2%

— BIRADS Ⅳ: 악성 확률이 높으며, 생검을 권장

— BIRADS Ⅴ: 악성 종양이 95% 이상인 유방암의 특징적인 형태를 보임

— BIRADS Ⅵ: 최종 치료 전에 생검으로 입증된 암 환자

3.8 병리검사

- 인체에서 발생하는 모든 질병에 대해 질병이 의심되는 부위의 조직을 떼어내어 슬라이드로 만든 후 광학현미경으로 관찰 판독하는 것으로서 질병에 대한 최종진단
- 수술도중에는 동결조직병리 검사를 응급으로 시행함으로써 수술범위 결정에 도움을 주며 수술 후에는 떼어낸 조직에서 질환에 의한 조직변화를 현미경으로 세밀히 관찰하고 분류하며, 필요한 경우 다양한 추가 병리검사를 시행하여 최종진단

3.8 OCT 검사 (Optical coherence tomography; 빛간섭단층촬영)

- 시신경 유두, 망막 및 각막을 포함한 안구의 구조를 파악하고 분석하는 거사로, 10개의 망막층을 높은 해상도로 촬영하여 이상 유무를 확인하고 진단하여 망막에 발생하는 질환을 조기에 발견하기 위한 검사방법

4. 약어

- **BI-RADS** Breast Imaging-Reporting and Data System
- **CR** Computed Radiography
- **DICOM** Digital Imaging and Communications in Medicine
- **DR** Digital Radiography
- **EMR** Electronic Medical Record
- **LCC** Left CranioCaudal
- **LMLO** Left MedioLateral Oblique
- **PACS** Picture Archiving Communication System
- **RCC** Right CranioCaudal
- **RMLO** Right MedioLateral Oblique

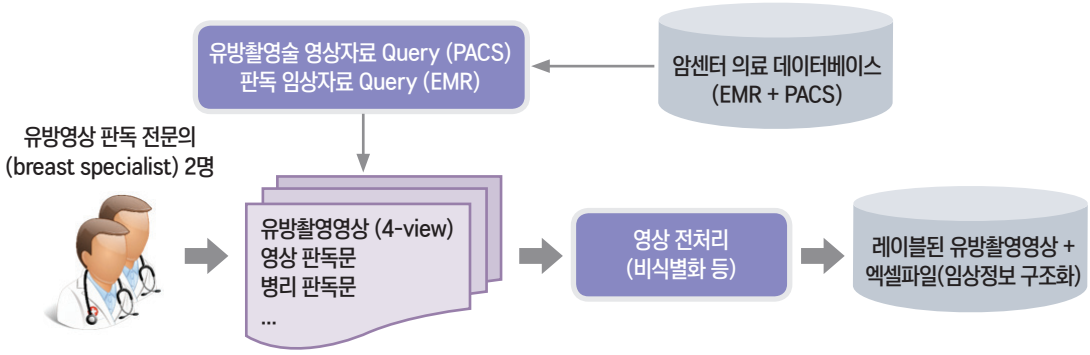
5. (1 세부) 유방촬영영상 인공지능 학습 데이터 구축

● 5.1 수집 데이터 상세명세서

○ 5.1.1 데이터 수집 경로

- 국립암센터에서 운영하고 있는 의료데이터베이스 (EMR, PACS)에서 데이터베이스 질의를 통해, 유방암 진단을 위해 촬영한 유방촬영영상 및 임상정보 등을 수집
  - 각 의료기관 마다 의료 영상을 다루는 장비와 시스템이 상이하여 의료 영상 이미지 마다 서로 다른 resolution으로 생길 수 있는 bias를 방지하기 위해서 각 의료 기관별 고유의 PACS만을 사용
  - 개인정보 보호법 등의 이유로 개인 또는 타 기관에서 각 기관의 의료데이터베이스에 직접적인 접근은 불가능함. 또한 데이터베이스 접근을 위해 반드시 IRB 승인이 선행되어야하며, 승인된 연구자에 한하여 의료데이터베이스에 접근가능

[그림 5-1] 데이터 수집 절차



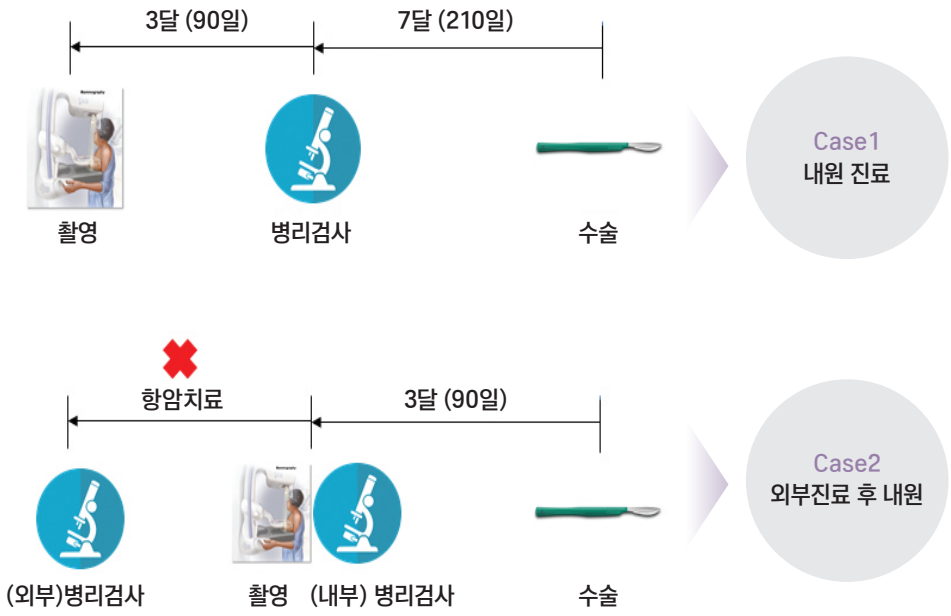
  - 기존의 유방암 관련 공개 DB들은 비교적 적은 샘플 수를 기반으로 임상정보를 제외하고 dicom 영상 이미지만을 공개하고 있는 한계점을 보이고 있으며, 이러한 기존 공개 DB들의 한계점을 표 5-1에 정리하였음
  - 이러한 한계점을 해결하기 위하여 본 연구에서는 7,500명의 정밀하고 세계 최대 규모의 유방촬영영상 이미지와 더불어 임상 데이터를 포함하여 공개함으로써 다양한 연구에 활용가능 하도록 지원

[표 5-1] 기존 공개 DB 동향

데이터명	수집기관	샘플 수	한계점
DDMS: Digital Database for Screening Mammography	University of South Florida	• 2,620 cases - 정상: 695 - 양성: 870 - 악성: 914	• 적은 수의 샘플 수
INbreast Database	Breast Centre, located in a University Hospital	• 115 cases - 정상: 90 cases, 410 images - 악성: 25 cases, 50 images	• 샘플 수가 아주 적음
CBIS-DDSM	DDMS 이미지를 포함한 업그레이드 버전	• 6,675 studies (10,239 images)	• DICOM 이미지 파일만 공개 • 각 환자들의 임상정보 미포함
VICTRE	FDA (VICTRE project)	• 8,749 studies (2,994 patients, 217,913 image)	• DICOM 이미지 파일만 공개 • 각 환자들의 임상정보 미포함

- 유방촬영영상의 진단 결과는 악성 (malignant), 양성 (benign), 정상 (normal) 유방촬영영상으로 구분할 수 있으며, 이 중 양성과 정상은 비악성으로 구분
  - 1) 유방암 세부 기준: 조직검사 결과 악성으로 확진된 환자의 유방촬영영상 이미지이며, 악성의 경우 조직검사를 기준으로 선정하는 것이 일반적(ScreenPoint Medical BV, 2018)이며, 내원에서 진료받은 경우 수술일 기준으로 210일 안에 병리검사를 하고, 300일 안에 유방촬영영상을 확보한 경우와 외부에서 진료 후 내원한 경우 내원하기 전 외부에서 병리검사 후 항암치료로 진단결과의 bias가 수술일 기준으로 90일 이내에 병리검사와 유방촬영영상을 확보한 이미지를 수집
  - 2) 양성 세부 기준: 영상 판독상 양성(BI-RADS category 2,3)이며, 1년 이상 추적검사 결과 여전히 양성인 환자의 유방촬영영상
  - 3) 정상 세부 기준: 영상 판독상 정상(BI-RADS category 1)이며, 1년 이상 추적검사 결과 여전히 정산인 환자의 유방촬영영상
  - 양성, 정상의 경우 판독 당시 오진이 있을 수 있으므로 추후 검사를 통해 변화가 없는지를 확인하는 것이 필요하며, 일반적으로 at least 1year follow-up 검사에서 이상 소견이 없는지를 확인하여 판단
- 유방암 수집절차 세부 기준 절차는 아래 그림 5-2와 같으며, 전체 데이터 수집의 상세 내역은 아래 표 5-2와 같음

[그림 5-2] 악성 유방암 수집 절차 세부기준



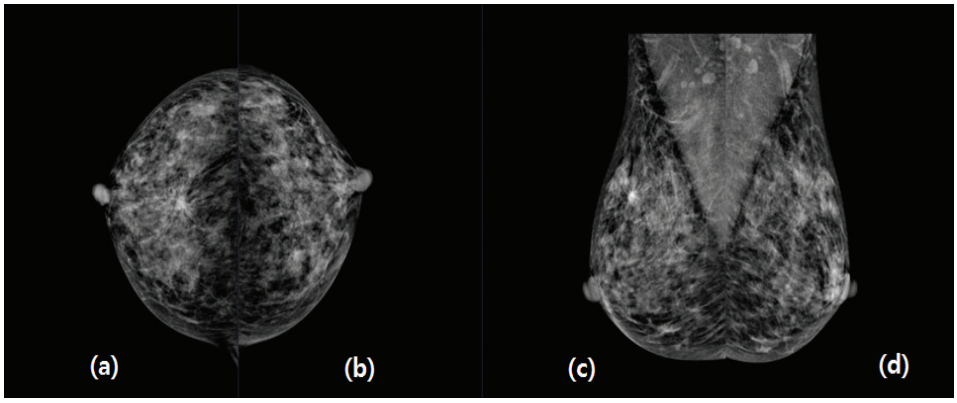
[표 5-2] 데이터 수집 절차

수집 절차	상세 내역	산출물
1. 국립암센터 의료 데이터 베이스 (EMR, PACS)에 질의	만 19세 이상 성인 여성환자의 유방암 진단을 위해 촬영한 유방촬영 이미지와 임상 정보 수집	
2. 4-view paired 유방촬영 영상 자료 수집	RCC, RMLO, LCC, LMLO 유방촬영영상 수집	
3. 유방촬영영상 판독	영상 판독문 기준으로 유방촬영영상의 악성, 양성, 정상으로 구분	
3.1 유방촬영영상 악성 판정	1. 유방촬영영상에서 악성 병변이 발견되어 악성이 의심되는 의심 환자의 추적 조직검사를 수행 2. 추적 조직검사를 통해 유방암으로 판명된 유방암 환자의 유방촬영영상 데이터 수집	• 유방암 환자의 유방촬영영상 데이터 • 조직검사 및 병리 검사보고서 (pathology report) • BI-RADS 4, 5, 6
3.2 유방촬영영상 양성 판정	1. 유방촬영영상에서 양성 병변 확인 2. 1년 이상의 지난 시점에서 양성으로 판명된 환자의 유방촬영영상 데이터 수집	• 양성 유방촬영영상 데이터 • BI-RADS 2, 3
3.3 유방촬영영상 정상 판정	1. 유방촬영영상에서 음성으로 판단 2. 1년 이상의 지난 시점에서 여전히 음성으로 판명된 정상군의 유방촬영영상 데이터 수집	• 정상 유방촬영영상 데이터 • BI-RADS 1
4. 영상 비식별화 전처리	영상 데이터 내에 포함된 개인정보(환자번호, 성별, 나이 등)의 비식별화 처리	• 비식별화된 유방촬영영상 데이터
5. 비식별화된 영상 데이터와 임상 데이터의 결합	레이블(악성, 양성, 정상) 된 유방촬영영상 데이터와 관련한 임상 데이터의 결합	• 두 데이터가 결합된 최종 데이터

○ 5.1.2 영상 데이터

- 만 19세 이상 성인 여성환자들을 대상으로 유방암 1,500명 (6,000장), 양성종양 1,500명 (6,000장), 정상유방 4,500명 (18,000장) 유방촬영영상 데이터 수집
- 4-view (RCC, RMLO, LCC, LMLO) paired 유방촬영 영상으로 한 환자당 4개의 이미지 데이터 수집하며 그림 5-3과 같은 영상촬영 영상 데이터 수집

[그림 5-3] 악성 유방촬영영상 (a) LCC (b) RCC (c) LMLO (d) RMLO



○ 5.1.3 임상 데이터

- 영상 자료 이외의 해당 영상의 환자 임상정보 수집하며, 수집하는 임상 데이터의 목록은 표 5-3과 같음

[표 5-3] 임상 데이터 수집 목록

데이터명	상세내역
인구학적 정보	• 촬영시점의 나이, 성별 등
유방촬영영상의 판독보고서 (radiology report)	• 영상 촬영 일시 • 유방촬영 영상의 형태 (4-view paired 여부) • 유방촬영 영상의 기기 정보 (모델명, 제조사명) • BI-RADS 판독 결과
악성 유방촬영 영상의 조직검사 및 병리 검사 보고서 (pathology report)	• 조직 검사 일시 • 유방암 조직 검사 결과 • 조직 검사가 수행된 병소의 유방촬영영상에서의 위치
양성 및 정상 유방촬영영상의 추적 유방촬영영상 판독 보고서	• 추적 영상 촬영 일시 • BI-RADS 판독결과 (final assessment)

○ 5.1.4 데이터 구조

- 유방촬영 영상의 판독후 악성, 양성, 정상으로 판명된 유방촬영 영상 데이터와 해당 환자의 임상정보를 결합하여 최종 데이터 구조는 표 5-4과 같음

[표 5-4] 수집 데이터 구조

데이터명	상세내역
영상자료	• 4-view paired 유방촬영영상 (RCC, RMLO, LCC, LMLO의 dcm (다이콤 파일), png (pixeldata))
레이블	• 악성, 양성, 정상 • 환자 단위 레이블 • 유방 단위 레이블 (RCC, RMLO 유방촬영영상에서 병변이 관찰된 경우, right) • 뷰 단위 레이블 (각 RCC, RMLO, LCC, LMLO 뷰 단위)
BI-RADS 판독 정보	• 환자 단위의 BI-RADS 판독 카테고리 정보(1, 2, 3, 4a, 4b, 4c, 5)
악성 유방촬영영상의 병리 판독 정보	• 조직검사 (core-biopsy) 결과
대상자 정보	• 인구학적 정보 (예: 촬영시점의 나이, 성별)
기타 정보	• 영상 촬영 일시, 유방촬영영상의 기기정보(모델명, 제조사명)

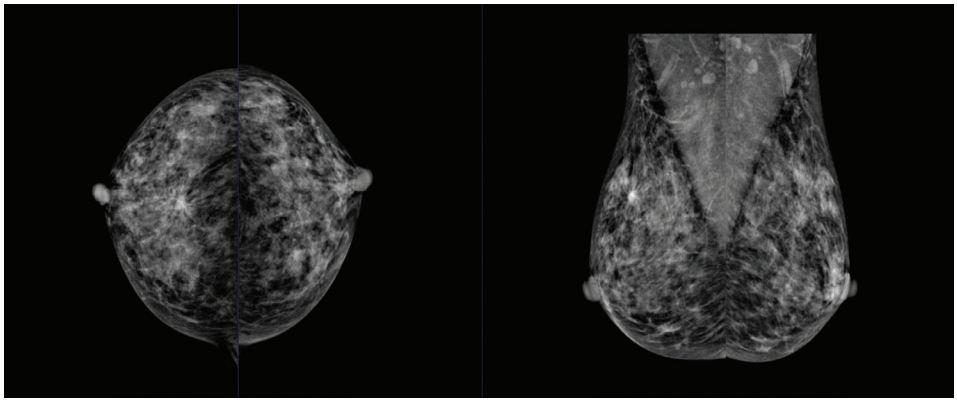
5.2 데이터 정제

- 한 환자당 최적의 4-view paired 영상자료를 보장하기 위하여 최종 4개 (RCC, RMLO, LCC, LMLO) view를 선별
  - 4-view paired (RCC, RMLO, LCC, LMLO) 중 하나라도 없는 유방촬영 영상은 제거
  - 각 4개의 view 중에서 영상이 2장 이상 존재하는 경우, 유방암 판독에 적합한 영상을 제외하고 중복되는 영상은 제거 (예: spot compression 영상, magnification view 영상 등)
- DR (digital radiography) 영상자료 수집
- CR (computed radiography) 영상 제거
- Film scan 영상 제거

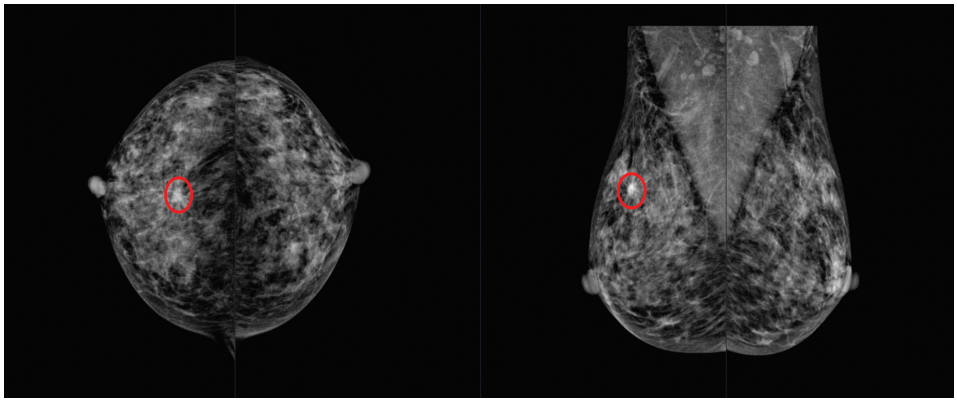
5.3 데이터 가공 (어노테이션)

- 유방촬영영상에서의 악성 병변 어노테이션을 위해 pixel-level segmentation 형태이기 때문에, 악성병변의 다양한 형태를 고려한 annotation이 가능
  - backend database: dicom 영상
  - frontend: web browser (chrome)
  - annotation: mouse draw
  - output: contour 좌표를 포함하는 json file
- 원본 영상의 좌표계 (예: 2000x3000 영상은 [0,0] ~ [1999,2999] range) 에서의 annotation polygon 의 좌표의 list (예: [(100,100), (103, 102), (105, 108), ... , (232, 303)] 형태
  - 꼭지점이 200개인 polygon이라면 200개의 좌표의 list로 구성) 로 contour를 정의할 수 있으며, 해당 contour가 어떤 study의 어떤 view (예: RCC, LCC, RMLO, LMLO)를 참조하는지를 포함하도록 함

[그림 5-4] 악성 유방촬영영상



(a) 원본 악성 유방촬영영상



(b) 어노테이션 (segmentation 형태)

5.4 최종 데이터 구조

- 최종 어노테이션이 완료되어 공개하는 데이터의 수량과 종류는 아래와 같음
  — 최종 목표 데이터 수량은 아래 표 5-6과 같음

[표 5-5] 최종 목표 데이터 수량

	환자수 (명)	영상이미지 (개)
정상유방	4,500	18,000
양성종양	1,500	6,000
유방암	1,500	6,000
합계	7,500	30,000

5.4.1 영상 DICOM

- 각 케이스는 4장의 dicom 영상, 즉 RCC, LCC, RMLO, LMLO 영상으로 구성

- DICOM 파일: 태그가 있는 이미지 파일. 이미지와 이미지에 대한 데이터 정보가 모두 들어있음. DICOM 이미지 파일의 데이터는 개별 요소의 순서로 저장되며, 각 요소에는 이미지 또는 이미지 자체에 대한 정보 항목이 tag형식으로 존재
  - Tag: 일반적으로 16진수 형식 (XXXX,XXXX)으로 속성을 식별하며, DICOM의 그룹 번호와 요소 번호로 더 나눌 수 있음
  - DICOM Value Representation (VR): 속성 값의 데이터 유형 및 형식
  - Name: 해당 Tag의 이름
  - DICOM tag의 표준 속성 구조는 아래 주소에서 확인 가능 (https://www.dicomlibrary.com/dicom/dicom-tags/)

[그림 5-5] DICOM tags

Tag	VR	Name	
(0002,0000)	UL	File Meta Information Group Length	
(0002,0001)	OB	File Meta Information Version	
(0002,0002)	UI	Media Storage SOP Class UID	
(0002,0003)	UI	Media Storage SOP Instance UID	
(0002,0010)	UI	Transfer Syntax UID	
(0002,0012)	UI	Implementation Class UID	
(0002,0013)	SH	Implementation Version Name	
(0002,0016)	AE	Source Application Entity Title	
(0002,0017)	AE	Sending Application Entity Title	
(0002,0018)	AE	Receiving Application Entity Title	
(0002,0100)	UI	Private Information Creator UID	
(0002,0102)	OB	Private Information	
(0004,1130)	CS	File-set ID	
(0004,1141)	CS	File-set Descriptor File ID	
(0004,1142)	CS	Specific Character Set of File-set Descriptor File	
(0004,1200)	UL	Offset of the First Directory Record of the Root Directory Entity	
(0004,1202)	UL	Offset of the Last Directory Record of the Root Directory Entity	
(0004,1212)	US	File-set Consistency Flag	
(0004,1220)	SQ	Directory Record Sequence	
(0004,1400)	UL	Offset of the Next Directory Record	
(0004,1410)	US	Record In-use Flag	
(0004,1420)	UL	Offset of Referenced Lower-Level Directory Entity	
(0004,1430)	CS	Directory Record Type	
(0004,1432)	UI	Private Record UID	
(0004,1500)	CS	Referenced File ID	
(0004,1504)	UL	MRDR Directory Record Offset	Retired
(0004,1510)	UI	Referenced SOP Class UID in File	
(0004,1511)	UI	Referenced SOP Instance UID in File	
(0004,1512)	UI	Referenced Transfer Syntax UID in File	
(0004,151A)	UI	Referenced Related General SOP Class UID in File	
(0004,1600)	UL	Number of References	Retired
(0008,0001)	UL	Length to End	Retired
(0008,0005)	CS	Specific Character Set	

DICOM tag들이 가지는 값의 정의는 value representation (VR)은 아래 표 5-6과 같이 정리

[표 5-6] DICOM Value Representations (VR)

VR	Name	Length of Value
AE	Application Entity	16 bytes maximum
AS	Age String	4 bytes fixed
AT	Attribute Tag	4 bytes fixed
CS	Code String	16 bytes maximum
DA	Date	8 bytes fixed
DS	Decimal String	16 bytes maximum
DT	Date Time	26 bytes maximum
FL	Floating Point Single	4 bytes fixed
FD	Floating Point Double	8 bytes fixed
IS	Integer String	12 bytes maximum
LO	Long string	64 chars maximum
LT	Long text	10240 chars maximum
OB	Other Byte String	see Transfer Syntax definition
OD	Other Double String	bytes maximum
OF	Other Float String	bytes maximum
OW	Other Word String	see Transfer Syntax definition
PN	Person Name	64 chars maximum per component group
SH	Short String	64 chars maximum
SL	Signed Long	4 bytes fixed
SQ	Sequence of Items	not applicable
SS	signed Short	2 bytes fixed
ST	Short Text	1024 chars maximum
TM	Time	16 bytes fixed
UI	Unique Identifier (UID)	64 chars maximum
UL	Unsigned Long	4 bytes fixed
UN	Unknown	Any length
US	Unsigned Short	2 bytes fixed
UT	Unlimited Text	bytes maximum

— 유방촬영영상 이미지에서 비식별화 처리 후에 DICOM tag는 표 5-7로 정리

[표 5-7] 비식별화 후 DICOM tag 정보

Tag (Group, Element)	VR	TAG Description	Value (예시)
(0002,0000)	UL	FileMetaInformationGroupLength (파일의 메타 정보 그룹 길이)	188
(0002,0002)	UI	MediaStorageSOPClassUID (SOP 클래스의 고유식별자)	1.2.840.10008.5.1.4.1.1.1.2
(0002,0003)	UI	MediaStorageSOPInstanceUID (SOP 인스턴스의 고유식별자)	1.2.840.113681.2887049286.14 91896736.4980.135780
(0002,0010)	UI	TransferSyntaxUID (데이터셋 인코딩에 사용되는 전송 구문의 고유식별자)	1.2.840.10008.1.2.4.91
(0002,0012)	UI	ImplementationClassUID (구현된 클래스의 고유식별자)	1.2.410.200003.2020819.5.2.1
(0008,0008)	CS	ImageType (이미지 타입)	DERIVED₩PRIMARY
(0008,0016)	UI	SOPClassUID (SOP 클래스의 고유식별자)	1.2.840.10008.5.1.4.1.1.1.2
(0008,0018)	UI	SOPInstanceUID (SOP 인스턴스의 고유식별자)	1.2.840.113681.2887049286.14 91896736.4980.135780
(0008,0060)	CS	Modality (양식)	MG
(0008,0070)	LO	Manufacturer (제조사)	XXXXXX, Inc.
(0008,1090)	LO	ManufacturerModelName (제조사 모델명)	XXXXX Dimensions
(0028,135A)	CS	SpatialLocationsPreserved (공간 위치 보존 여부)	YES
(0010,0020)	PN	PatientID (환자 ID)	10001
(0010,0030)	DA	PatientBirthDate (환자 생년월일)	19000101
(0018,1164)	DS	ImagerPixelSpacing (픽셀 간격)	0.070000₩0.070000
(0018,5101)	CS	ViewPosition (이미지 대상의 방향)	CC
(0020,0062)	CS	ImageLaterality (이미지의 부위 위치: 왼쪽, 오른쪽)	L
(0028,0002)	US	SamplesPerPixel (픽셀당 샘플 크기)	1
(0028,0004)	CS	PhotometricInterpretation (광도 해석)	MONOCHROME2
(0028,0010)	US	Rows (행)	3328

Tag (Group, Element)	VR	TAG Description	Value (예시)
(0028,0011)	US	Columns (열)	2560
(0028,0030)	DS	PixelSpacing (픽셀 간격)	0.065238₩0.065238
(0028,0100)	US	BitsAllocated (할당된 bit 단위)	16
(0028,0101)	US	BitsStored (저장된 bit 단위)	12
(0028,0102)	US	HighBit (픽셀 최상위 bit 값)	11
(0028,0103)	US	PixelRepresentation (픽셀 표현방식)	0
(0028,1050)	DS	WindowCenter (창 중심값)	2047
(0028,1051)	DS	WindowWidth (창 너비값)	4096
(0028,1052)	DS	RescaleIntercept (저장된 값과 출력 사이에 있는 값)	0
(0028,1053)	DS	RescaleSlope (재조정 스케일)	1
(0028,1054)	LO	RescaleType (재조정 유형)	US
(0040,0244)	DA	PerformedProcedureStepStartDate (수행된 프로세스 시작 날짜)	19000101
(0040,0245)	TM	PerformedProcedureStepStartTime (수행된 프로세스 시작 시간)	0
(0008,0100)	SH	CodeValue (코드 값)	R-10242
(0008,0102)	SH	CodingSchemeDesignator (코딩 스키마 지정자)	SNM3
(0008,0104)	LO	CodeMeaning (코드 의미)	cranio-caudal

5.4.2 임상 메타정보

- 임상정보: (정상, 양성) 영상판독보고서 및 임상정보, (악성) 영상판독보고서 및 병리판독보고서, 임상정보 xlsx 파일로 정리

[표 5-8] 메타데이터 정보

데이터명	변수명	내용
공통추출변수 (정상/ 양성/악성)	grp	• 양성/악성/정상에 대한 그룹 정보
	recvid	• 비식별후에 대체번호로 부여된 환자 ID
	manufacturer	• 검사장비
	manufacturerModelName	• 검사장비 모델
	both1	• BIRADS category (양쪽)
	left1	• BIRADS category (좌)
수술후 병리 변수 (악성만 해당)	spr_loca_cmnt	• 악성 종양의 위치 (ex. breast left/right)
	invs_size_vl	• 악성 종양의 크기
	htlg_type_cmnt	• 조직형 진단 (IDC, ILC, DICS, others)
	path_t_stag	• 종양의 병기 (ex. 1, 2, 3, 4)

5.4.3 어노테이션 정보

- annotation 결과: 앞서 언급한 대로, pixel-level annotation pologon (segmentation) 의 coordinate의 list를 포함하며, 해당 annotation이 어떤 영상을 참조하는지를 포함한 json 파일

[표 5-9] 어노테이션 json 정보

컬럼명	상세내역
user_id	수행 영상판독 전문의 ID
case_id	해당 유방촬영 이미지 ID
view_location	유방촬영영상 위치 (RCC, RMLO, LCC, LMLO)
lesion_id	악성병변 ID
contour coordinate	악성병변 coordinates list
discard_yn	해당 영상에 문제가 있어서 사용하지 못하게 될 케이스임을 표시하는 용도

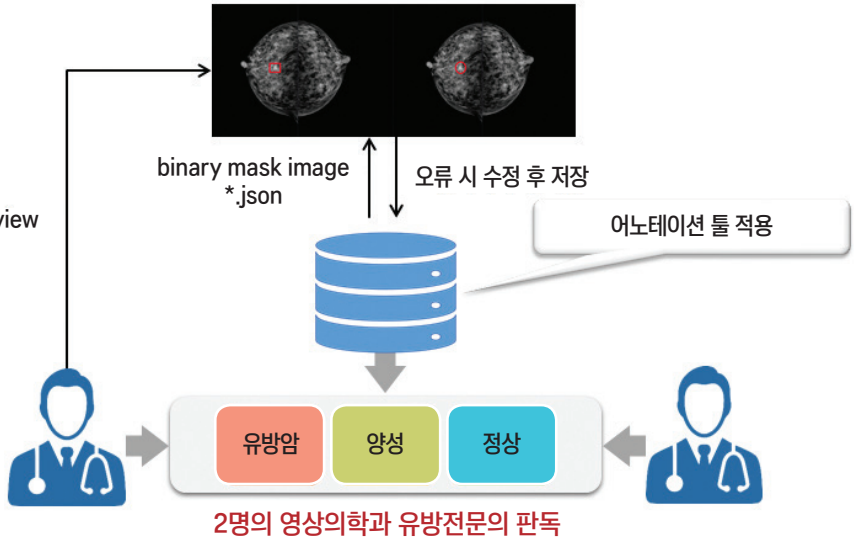
[그림 5-6] 어노테이션 json 파일 depth별 구성 예시

```
{
    "user_id": 1,      #annotation 전문의에게 부여되는 Tool 로그인 ID
    "case_id": 100001, # 해당 Mammography 케이스 ID
    "contour_list": {     # 각 View에 그려지는 Lesion에 대한 정보가 List for coordinates 형태로 저장
            "cancer": {   # Caner case에 대한 lesion 정보가 View (rcc, rmlo, lcc, lmlo)별로 저장
                "rmlo": {  # rmlo view에 대한 lesion 좌표 정보들
                    "lesion01": [{"y": -78, "x": 195}, {"y": -78, "x": 191}, ... ,
                    {"y": -68, "x": 239}, {"y": -78, "x": 195}],
                    "lesion02": [{"y": -78, "x": 195}, {"y": -78, "x": 191}, ... ,
                    {"y": -68, "x": 239}, {"y": -78, "x": 195}]
                },
                 "rcc": {  # rcc veiw에 대한 lesion 좌표 정보들
                    "lesion03": [{"y": -62, "x": 354}, {"y": -62, "x": 340}, ... ,
                    {"y": -105, "x": 412}, {"y": -62, "x": 354}]
                }
            }
    },
    "discard_yn": 0  # 해당 영상에 문제가 있어 사용이 불가능할 경우 제외 여부 yes or no 표기
}
```

5.5 데이터 검수

- 악성 유방촬영영상 선정기준에 부합하는 악성 영상자료에 대해, 추적 조직검사가 수행된 악성 병변의 위치를 radiology report 및 pathology report를 참고하여 전문가 (영상의학과 전문의)가 어노테이션 수행하여 검수
- 악성 유방촬영영상의 조직검사가 수행된 악성병변 이외에 악성이 의심되는 병변이 존재할 경우, 숙련된 전문가 (유방판독 전문의)의 판단하에 해당 영역 어노테이션을 추가로 수행
- 영상판독문 또는 병리판독문이 존재하는 영상에 대해서 새로운 두 명의 유방영상 판독 전문의가 교차판독을 통하여 두 명의 전문의의 소견이 100% 일치를 통해 데이터의 검수를 진행
- Option 1: 저작도구에 어노테이션 review 기능 구현:
  - 지금까지 진행된 어노테이션에 대해, 데이터베이스에 저장된 결과파일 (json 혹은 binary mask image)을 영상자료상에 로드하여 해당 병변이 제대로 어노테이션 및 저장됐는지 확인
  - 오류가 있을 경우 저작도구를 활용하여 수정하고, 해당 내용에 대해 기록을 남겨 2명의 전문의의 합의하에 최종 판단
- Option 2: 저장된 어노테이션 결과 데이터를 포함한 영상을 별도로 생성하여 검수
  - 지금까지 진행된 어노테이션에 대해, 데이터베이스에 저장된 결과파일을 영상자료상에 contour 형태로 표현하여, annotated 영상을 새로 생성하여 해당 병변이 제대로 어노테이션 및 저장됐는지 확인
  - 오류가 있을 경우 해당내용에 대해 기록을 남겨 2명의 전문의의 합의하에 최종 판단하고, 이 데이터를 별도로 모아서 최종 어노테이션 수정

[그림 5-7] 데이터 검수 과정

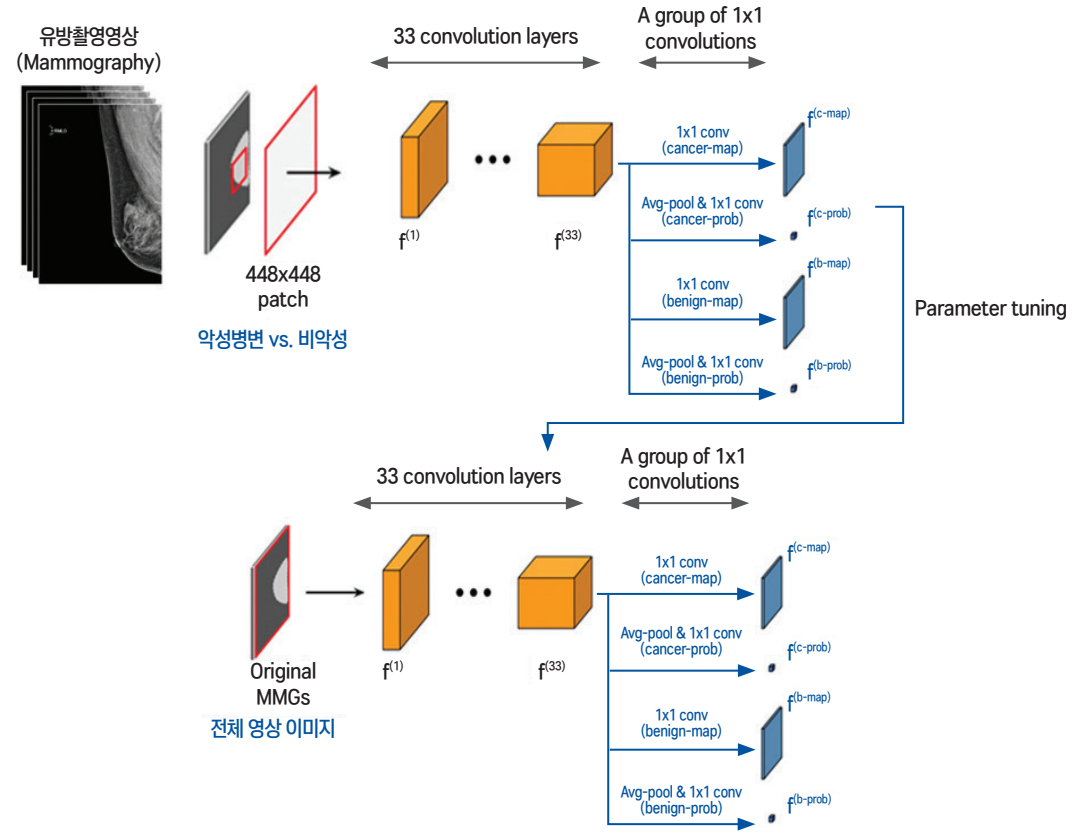


- 영상 선별 기준에 따르면,
  - 가장 중요한 악성 영상의 경우, 영상판독으로 악성을 정의하는 것이 아니라, 조직검사결과 유방암 확진된 환자의 영상으로 정의하기 때문에 100% ground-truth 임
  - 정상 및 양성 영상은 영상 판독만으로 정의할 수밖에 없으나, 통상적으로 at least 1 year of follow-up imaging으로 변화를 관찰하여 정의함
  - 이 과제에서는 위의 기준으로 영상을 수집하기 때문에, standard를 따른다고 볼 수 있음

5.6. 인공지능 판독 시범 모델 개발 및 데이터 검증

- 다양한 영상 분석에서 높은 성능을 보여주며, 영상 분석 딥러닝 알고리즘 중 가장 널리 사용되고 잇는 CNN (convolutional neural network) 기반의 ResNet-34 (Kaiming et. al., CVPR 2016)를 고려하며 인공지능 판독 시범 모델을 개발할 예정이며, 세계적인 이미지넷 이미지 인식대회인 ILSVRC (ImageNet Large Scale Visual Recognition Challenge)를 통해 현재 공개된 이미지 인식에 대한 인공지능 알고리즘 기술이 기존의 컴퓨터비전 또는 머신러닝 모델보다 훨씬 향상된 성능이 검증되었기에, 머신러닝 기법 보다는 다른 영상 분석 딥러닝 알고리즘을 고려할 예정
- 이미 기존에 임상데이터(EHR, EMR)를 활용한 암 재발예측 및 위험예측 연구가 활발히 이루어지고 있으며, 영상 데이터뿐만 아니라 수집된 임상데이터를 함께 활용함으로써 예측 정확도를 높이는데 사용
- 학습과정 1단계는 패치기반으로 악성 병변의 특징과 비악성 영역을 극명하게 대조시키면서 악성 병변의 특징을 학습
- 학습과정 2단계는 양싱기반으로 학습은 영상단위 정보를 전반적으로 활용하여 1단계에서 악성 병변에 특화하여 학습한 네트워크 모델의 파라미터를 기반으로 2단계 학습 모델의 파라미터 세부 튜닝 (tuning)
  - 학습된 악성 병변의 특징을 잃지 않도록, 출력 맵 손실함수를 추가
  - 영상단위의 특징을 추가적으로 학습할 수 있도록, 분류 손실함수를 추가

[그림 5-8] ResNet-34 기반의 학습모델 예시

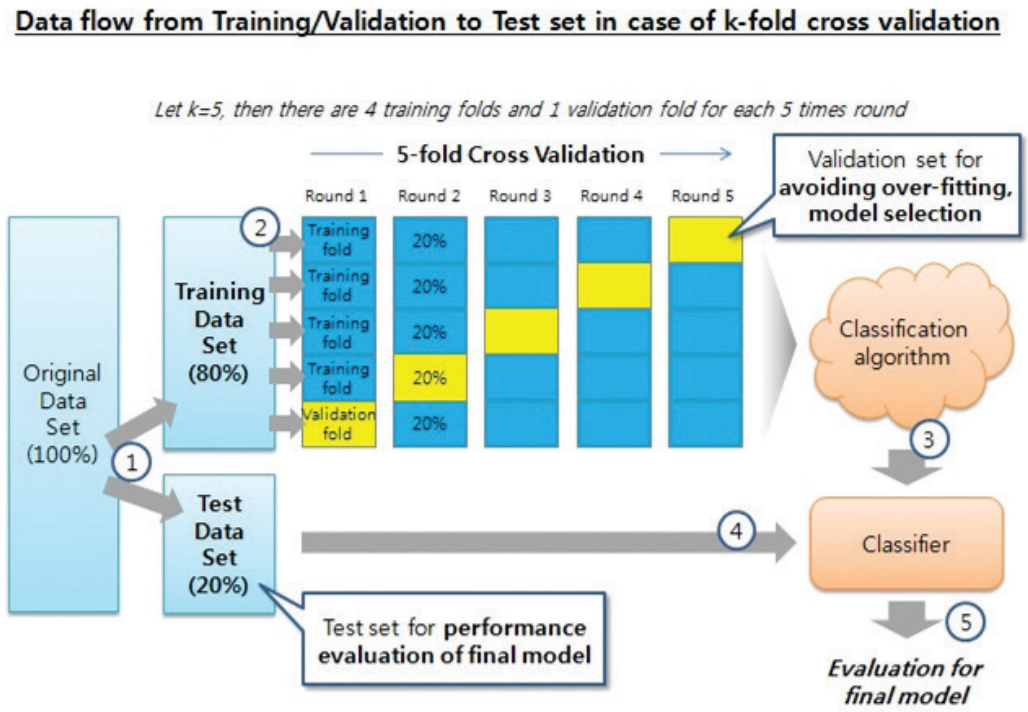


- 유방촬영영상의 총 데이터를 학습(training), 검증(validation), 테스트(test)로 나누어서 유방촬영영상의 유방암 판독 학습모델을 개발
- 유방암 판독 학습모델을 검증하기 위하여 n-fold cross validation을 수행하며, 예시로 전체 데이터를 학습용 (80%), 테스트용 (20%)로 나누고, 학습용 데이터에서 n개의 데이터 셋으로 random split
- n-1개 set으로 모델을 학습하고, 1개의 set으로 검증하며 판독 모델의 검증을 위한 data flow는 아래 그림 5-8과 같음
- 인공지능 판독 모델의 성능을 다양한 척도를 기반으로 평가하며, 환잔 단위 영상자료 기반의 민감도 (sensitivity), 특이도 (specificity), 정밀도 (precision), ROC curve에서의 AUC를 측정
- 최근 FDA 허가 완료 (Rodríguez-Ruiz, et al., 2018)된 딥러닝 기반 유방촬영영상 판독보조 소프트웨어에서는 약 18,000케이스의 유방촬영영상 이미지를 학습하여 14명의 방사선전문의와 AI 기술 기반의 판독 비교 결과, 판독의 성능이 향상
- 해당 과제에서는 최근 연구에서의 약 40% 수준 (7,500 cases)의 학습데이터로 AUC:0.85-0.87를 목표로 하며, 이는 적은 수준의 케이스로 유사한 결과를 도출하는데 있어서 과제 목표로 적절하다고 사료됨

[표 5-10] 학습모델 목표달성 지표 예제

데이터명	Radiologist	AI Support	목표치
민감도 (Sensitivity)	0.83 (0.81-0.85)	0.86 (0.84-0.88)	≧ 0.80
특이도 (specificity)	0.77 (0.75-0.79)	0.79 (0.77-0.81)	≧ 0.80
AUC	0.87 (0.83-0.90)	0.89 (0.85-0.92)	≧ 0.85 - 0.87

[그림 5-9] 판독 모델 검증 data flow 예시



출처: http://rfriend.tistory.com

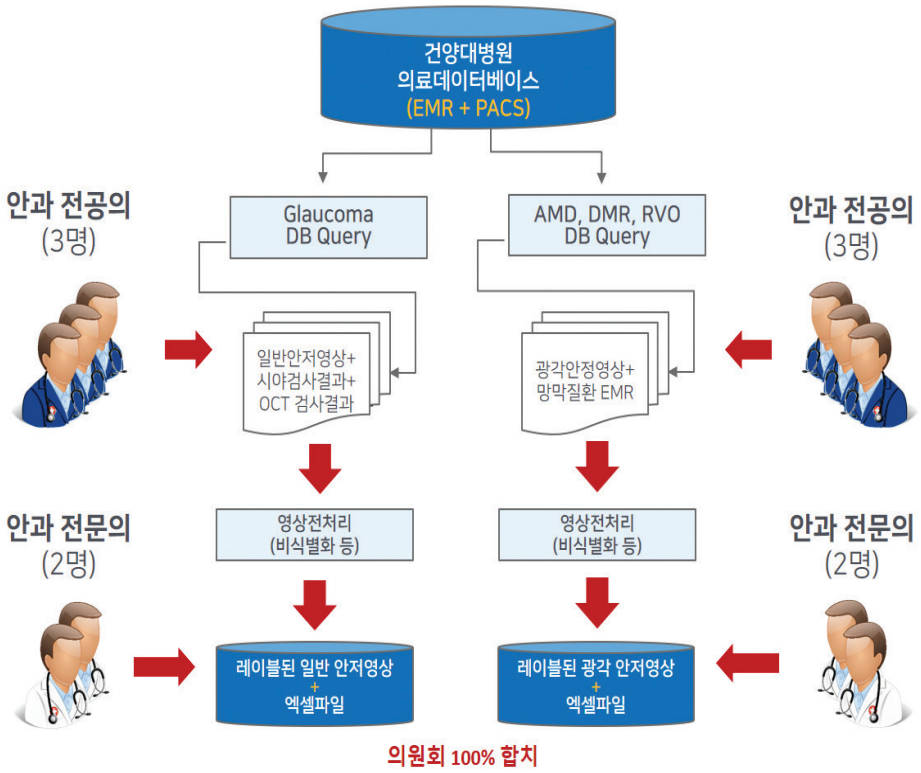
6. (2 세부) 안저 영상 인공지능 학습 데이터 구축

● 6.1 수집 데이터 상세명세서

○ 6.1.1 데이터 수집 경로

- 건양대학교 병원에서 운영하고 있는 의료데이터베이스(EMR, PACS)에서 데이터베이스 쿼리를 통하여 녹내장의 진단을 위한 일반 안저 영상과 망막질환(황반변성, 당뇨성망막병증, 망막정맥폐쇄 등)의 진단을 위한 광각 안저 영상을 추출함
  - 각 의료기관 마다 의료 영상을 다루는 장비와 시스템이 상이하여 의료 영상 이미지 마다 서로 다른 resolution으로 생길 수 있는 bias를 방지하기 위해서 각 의료 기관별 고유의 PACS만을 사용
  - 개인정보 보호법 등의 이유로 개인 또는 타 기관에서 각 의료 기관의 의료데이터베이스에 직접적인 접근은 불가능하며, 제한된 인원만이 의료기관과 협의를 하고 IRB 승인절차 필요
- 데이터 수집 절차는 아래 그림 6-1과 같음

[그림 6-1] 데이터 수집 절차



  - 기존의 안저질환 관련 공개 DB들은 매우 적은 샘플 수의 당뇨성망막병증에 대하여 공개된 DB가 존재하지만, 본 연구에서와 같이 녹내장 일반 안저 영상과 망막질환(황반변성, 당뇨성망막병증, 망막정맥폐쇄)을 고려한 연구결과가 없음

[표 6-1] 기존 안저질환 공개 DB 현황

데이터명	수집기관	샘플 수	한계점
DRIVE: Digital Retinal Images for Vessel Extraction	diabetic retinopathy screening program in The Netherlands	• 400 당뇨환자 샘플 • 정상: 33 • 초기 당뇨성망막병증: 7 • 당뇨성망막병증: 360	당뇨성망막병증의 데이터만 포함되어 있음
High-Resolution Fundus (HRF) Image Database	Friedrich-Alexander University Erlangen-Nuremberg (Germany)	• 정상 : 15 • 당뇨성망막병증: 15 • 녹내장: 15	매우 적은 수의 샘플과 황반병성 및 망막정맥폐쇄 관련 이미지는 없음

- EMR 데이터에서 환자의 임상데이터를 활용하여 관련 질환 확진된 환자를 대상으로 안저영상을 추출하며, 추출시 사용한 임상데이터는 영상 데이터셋 구축에 더 이상 필요로 하지 않음
- 레이블된 안저촬영영상과 엑셀파일 형태에 (좌/우, 촬영장비 등)으로 변환이 이루어짐

6.1.2 영상 데이터

- 안저촬영 장비의 기본 표준인 JPG 영상을 기준으로 사용
  — 일반/광각 안저촬영영상
- 데이터 판독 기준
  — 녹내장 : 1) C/D ratio 0.5 이상, 2) optic disc hem orrhage, 3) RNFL defect
  — 당뇨성망막병증 : 1) 망막출혈, 2) 망막의 exudate, 3) 미세동맥류
  — 황반변성 : 1) 황반출혈, 2) 황반의 drusen
  — 망막정맥폐쇄 : 1) 망막의 화염상 출혈, 2) 망막 혈관의 thrombosis
- 수집 목표: 총 4,500장
  — 망막촬영(일반안저영상): 정상소견 1,500장, 녹내장소견 1,000장
  — 망막촬영(광각안저영상): 정상소견 600장, 당뇨성망막병증 700장, 황반변성 500장 정맥폐쇄 200장
- 수집 분량의 적정성
  — 2018년 수집된 광각 안저영상 데이터 (정상 495장, 황반변성 473장, 당뇨성망막병증 280장, 정맥폐쇄 250장)을 활용한 판독 학습모델 시범 개발 및 성능검증의 결과에 비춰 이번에 제시된 목표 데이터 수집 분량은 적절한 것으로 판단할 수 있음. 아래 표 6-2와 6-3으로 해당 내용 확인할 수 있음

[표 6-2] 2018년 4종 안과 질환 의료 지식베이스 구축 결과

질환	수량 (장)		타입
	질환	정상	
나이관련황반변성(AMD)	473	495	광각
당뇨성망막병증(DMR)	825		광각
망막정맥폐쇄(RVO)	250		광각
녹내장(Glaucoma)	750	750	일반

[표 6-3] 2018년 녹내장 진단 모델 성능 평가표

	평균	최소	최대
정확도	0.936	0.920	0.955
민감도	0.925	0.886	0.955
특이도	0.948	0.909	1.000
AUROC	0.977	0.950	0.991

○ 6.1.3 데이터 구조

[표 6-4] 안저영상 메타데이터 구조

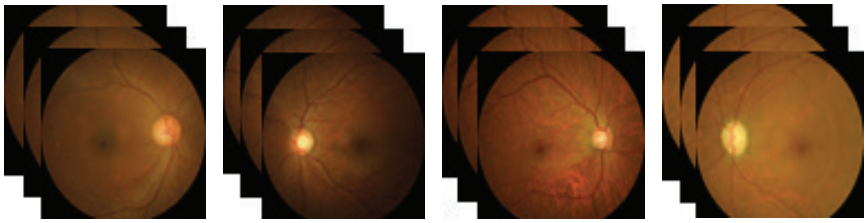
일련번호	이미지파일명	일반/광각	정상/질환코드	레이저치료유무	좌/우	촬영장비

□ 일련번호 : 일반(G) 광각(W), 질환코드 시작 알파벳, 4자리 숫자
- 일반 정상 : G_N_0001
- 일반 녹내장 : G_G_0001
- 광각 정상 : W_N_0001
- 광각 황반변성(AMD) : W_A_0001
- 광각 당뇨성망막병증(DMR) : W_D_0001
- 광각 망막정맥폐쇄(RVO) : W_R_0001

□ 이미지파일명 : 일반(General) 광각(Wide), 질환코드, 4자리 숫자, L/R
- 일반 정상 : General_Normal_0001_L
- 일반 녹내장 : General_Glaucoma_0001_L
- 광각 정상 : Wide_NOR_0001_L
- 광각 황반변성(AMD) : Wide_AMD_0001_L
- 광각 당뇨성망막병증(DMR) : Wide_DMR_0001_L
- 광각 망막정맥폐쇄(RVO) : Wide_RVO_0001_L

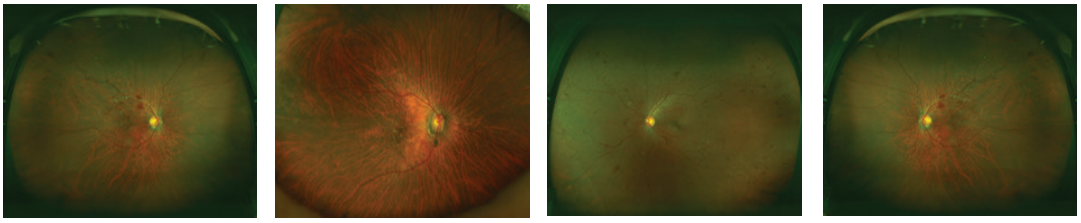
□ 일반/광각 : 일반, 광각
□ 정상/ 질환코드 : 정상, 녹내장, AMD, DMR, RVO
□ 레이저치료유무
□ 좌/우 : 좌, 우
□ 촬영장비

[그림 6-2] 안저촬영영상 데이터 예시



정상	녹내장 의심

(A) 일반 안저촬영영상



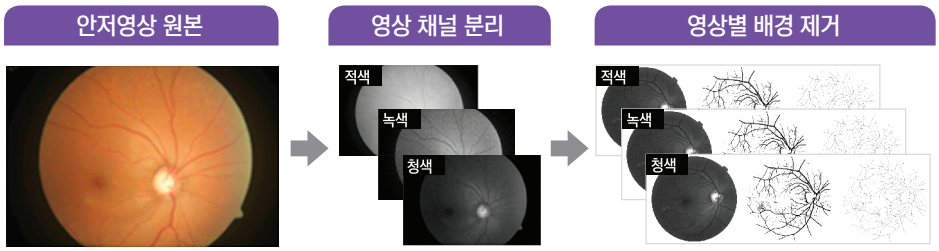
정상	AMD	DMR	RVO

(B) 광각 안저촬영영상

6.2 수집 데이터의 정제

- 개인정보 비식별화 처리된 안저 이미지는 3채널의 컬러 이미지로 전처리 방법에 따라 이미지 분석 및 분류 모델의 정확도와 신뢰도에 차이가 클 수 있으므로, 안질환에 따른 최적의 전처리 방법을 선택해야 됨
- 최적의 분류모델을 구현하기 위한 영상채널별 전처리
  - 3채널 컬러사진을 적색, 녹색, 청색으로 분리하여 각 영상에서 특이적 이미지의 콘트라스트, 선명도를 개선시킴. 혈관의 경우, 녹색 채널에서 혈관의 분포를 확인하기 용이하며, 적색의 경우, 황반변성 부분에 확인이 용이
- 효율적 학습/분석을 위한 영상내 배경 제거
  - 4K 초고해상도 안저촬영영상을 3개의 채널로 분리 후, 학습/분석시 상당한 시간과 연산능력이 요구됨. 이를 개선하기 위해 영상내 명백히 불필요한 부분을 제거하여 학습/분석시 효율성을 향상시키는 기술임

[그림 6-3] 안저촬영영상 전처리 기술 개념도 및 주요사항



안저영상 전처리 과정의 기술적 의의

- 이미지 픽셀 값 정규화
- 영상번호, 촬영일 등 마킹정보 제거
- 영상 채널 분리
- 영상별 배경 제거
- 안저 중심점 위치 정보

→

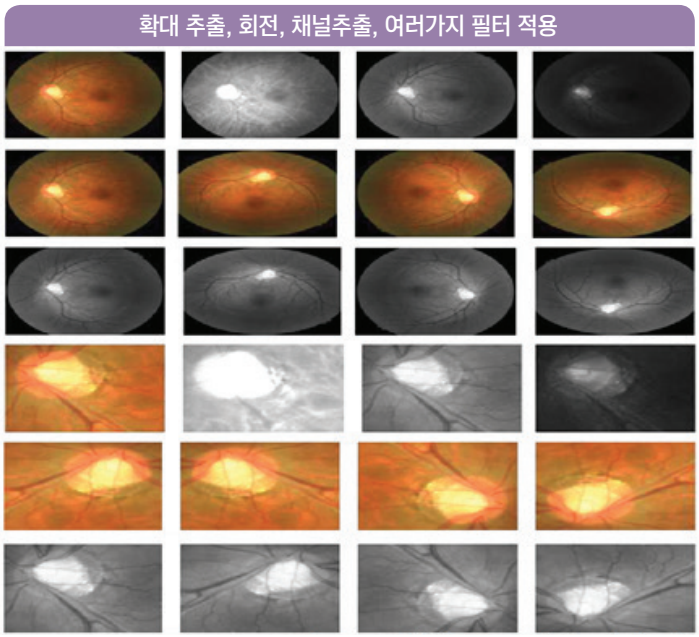
- 학습 데이터 효율성 개선
- 개인정보 보호
- 영상 분석 시간 단축
- 안저영상 정확도 개선

6.3 데이터 가공

6.3.1 데이터 증강

- 데이터 증폭의 목적은 안저촬영영상 자동 판독 모델을 구축하는 과정에서 최적화의 대상이 되는 모델 파라미터를 충분히 학습시키고, 특정 데이터(영상)에 종속되지 않게 판독 모델을 생성하기 위한 학습 데이터를 제공하기 위함
- 안저촬영영상 데이터의 증폭을 위해 다양한 필터적용, 확대추출, 채널추출, 회전등의 방업을 사용하여, 24개의 이미지로 데이터를 증폭하여 복수의 개별학습 모델에 적용하고자 함

[그림 6-4] 안저촬영영상 데이터 증강 처리결과 예시



- 데이터 증강에 대한 내용은 아래 표 6-5 와 같음

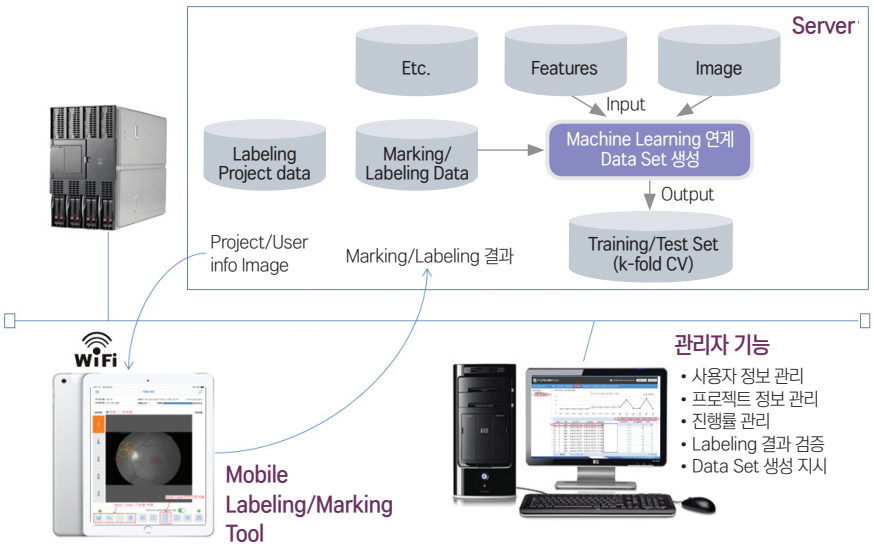
[표 6-5] 데이터 증강 방식

다양한 개별 모델의 생성을 위한 요인 변수 및 적용 방법 설계			
기준	코드	설명	비고
영상 필터 종류	B	Bilateral 필터 적용 영상 사용	6단계
	G	Gaussian 필터 적용 영상 사용	
	H	Histogram EQ 필터 적용 영상 사용	
	M	Median 필터 적용 영상 사용	
	O	원본 영상 그대로 사용	
	S	Sharpening 필터 적용 영상 사용	
영상 확대 여부	O	원본 영상 그대로 사용	2단계
	U	원본을 125% 확대한 후 중심점 기준 299x299 크기로 잘라낸 영상 사용	
영상 회전 여부	O	원본 영상 그대로 사용	4단계
	R	원본을 90, 180, 270도 회전한 영상 사용	
영상 컬러	C	컬러 영상 사용	2단계
	G	회색 영상 사용	

6.3.2 어노테이션

- 안저영상의 경우 어노테이션은 라벨링과 중의적인 표현으로 사용됨
- 어노테이션 저작도구는 서버와 모바일용, 그리고 관리자가 관리 및 분석하는 관리자 시스템으로 구성

[그림 6-5] 어노테이션 저작도구 시스템 구상도



저작도구 - 서버

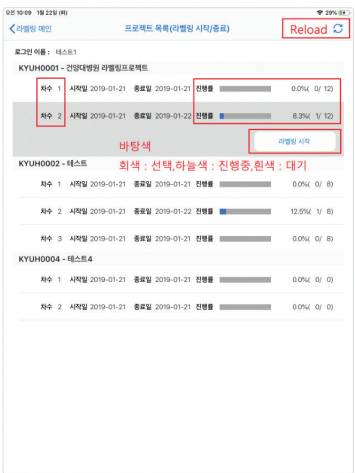
- DB 구성, 데이터 저장 및 저작도구 및 관리자 기능에서 요청하는 응답 처리 프로세스로 구성되며, 관리자 기능에는 각종 프로젝트에 대한 기준 정보, 이미지 정보 및 결과 분석 등의 관리자 프로그램으로 구성
  - 모바일 어노테이션 저작도구로부터 각종 정보에 대한 요청을 받고 관련 정보를 회신하며, 저작도구로부터 생성된 어노테이션 결과를 수신하여 DB에 저장하는 내부 프로시저와의 통로 역할을 수행
  - 서버에서 가동되는 메인 프로시저는 기준 정보 요청 및 결과 저장 요청 등의 정보를 파싱하여 프로젝트 목록 정보, 라벨 목록 정보 응답 처리, 어노테이션 결과 저장 등 해당하는 세부 프로시저를 호출하고 각각의 프로시저는 해당 기능을 수행하도록 구성

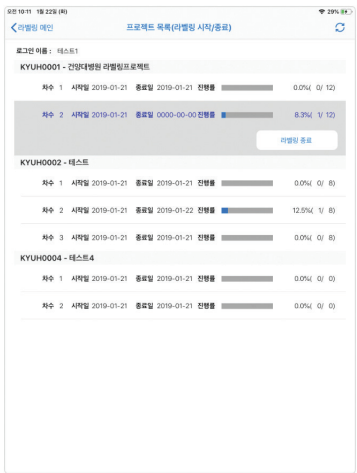
저작도구 - 모바일

- 의료 데이터에 대한 질환 판별을 수행하는 임상 전문의들의 공간 이동에 따른 어노테이션 작업의 편의성을 고려한 모바일 어노테이션 앱
  - 서버에 어노테이션 결과를 실시간으로 저장하여, 사업 관리자 또는 프로젝트 관리자가 어노테이션의 진행률이나 어노테이션 결과의 교차 검증 등 분석이 가능하도록 시스템의 토대를 마련
  - 어노테이션 프로젝트 목록
  - 서버에 저장된 어노테이션 대상 프로젝트 목록을 서버에 요청하여 화면에 출력
  - 표시된 프로젝트 중에서 어노테이션 할 프로젝트를 선택한 후 시작 버튼을 터치하여 어노테이션 작업 모드로 전환
  - 하나의 프로젝트에 대해서 어노테이션 작업을 수회 수행할 경우(교차 검증에 활용)를 대비하여 어노테이션 차수를 표시하고 해당 차수를 선택하여 어노테이션 메인 화면으로 전환
- 어노테이션 메인 화면
  - 서버로부터 어노테이션할 이미지를 다운로드하여 화면에 출력함.
  - 좌우 어노테이션이 필요한 경우를 대비하여 라벨 목록(Target 값)을 좌우에 표시할 수 있도록 함

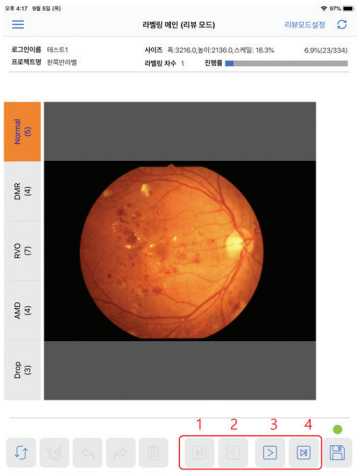
[그림 6-6] **안저촬영영상 저작도구 개발화면**

(A) 프로젝트 목록 조회 화면





(B) 메인 어노테이션 작업 화면



• 어노테이션 절차 및 방법

1) 관리자 시스템을 이용하여 각종 마스터 등록

① 프로젝트 정보 등록

② 프로젝트의 라벨 마스터 등록(Target 분류 코드)

③ 사용자 정보 등록

④ 프로젝트에 사용자 권한 부여

⑤ 서버에 어노테이션 대상 이미지 업로드 후 프로젝트에 등록된 이미지 매핑 작업

2) 어노테이션(Labeling) 앱 사용

① 로그인

② 사용자에 해당하는 프로젝트 목록 조회

③ 어노테이션을 하기 위한 프로젝트 선택

④ 어노테이션 실시

Ⅰ. 서버로부터 하나의 Source 이미지(안저 이미지)가 다운로드된 후 화면에 표시됨(참조 이미지를 확인할 수도 있음)

Ⅱ. 서버에 등록된 라벨 목록이 화면에 표시됨

Ⅲ. 라벨 목록에서 하나를 선택하여 라벨링

Ⅳ. Labeling(Annotation) 결과를 서버에 저장

⑤ (ⅰ)~(ⅳ) 항을 모든 이미지 annotation이 완료될 때까지 반복

3) 관리자 시스템을 통하여 어노테이션 결과 확인

① Labeling 결과 확인 프로그램에서 진도율 확인

② 이미지 ID별 Label 결과 확인

③ 동일 프로젝트에 대한 사용자별 Label 결과 비교(Cross checking)

※ 상이한 것들은 재 검토를 하거나 학습에서 제외되도록 정제 가능

4) Source 이미지, Label 결과 데이터를 이용하여 기계학습에 활용

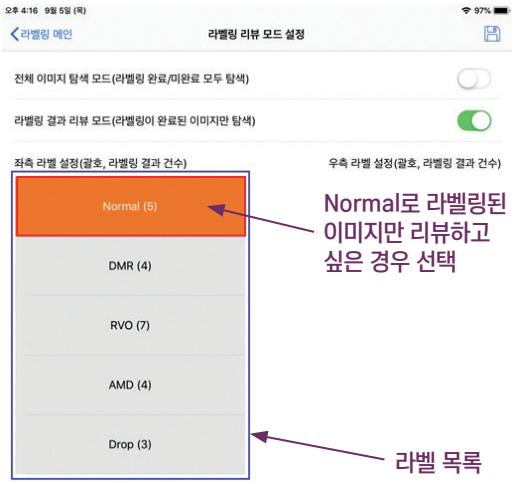
6.4 데이터 검수

- 건양대병원 EMR 데이터에서 안저질환 확진 환자를 대상으로 추출한 영상에 대해 안과 전공의들이 1차 판독을 함. 1차 판독 완료된 이미지에 한해 비식별화를 진행하고, 비식별화 완료된 데이터만을 저작도구를 사용해 2차 판독 및 검수
- 안저영상과 OCT 영상을 함께 참조하여 저작도구를 통해 녹내장 전문의 2인과 망막 전문의 2인이 영상을 판독 및 검수하고, 의원회에서 100% 합치된 영상만을 데이터셋으로 활용하기 때문에 100% ground truth로 볼 수 있음
- 안저촬영영상 진단 정보의 효율적 어노테이션을 위한 저작도구 및 교차 검증 지원
  - 서버에 Labeling 결과를 실시간으로 저장하여, 사업 관리자 또는 프로젝트 관리자가 Labeling의 진행률이나 Labeling 결과의 교차 검증 등 분석이 가능하도록 시스템의 토대를 마련
- 라벨 결과 확인
  - 모든 데이터는 프로젝트 단위로 수집되며, 수집된 라벨링 결과를 토대로 영상을 기계학습에 활용
  - 라벨별 검토 기능을 통해 데이터 재검토 역할을 수행

[표 6-6] Labeling 결과 테이블

필드명	데이터 타입	NULL 여부	비고
PROJECT_CD	CHAR	No	프로젝트 코드
USER_ID	CHAR	No	사용자 ID
LABELING_ORD	NUMBER	No	라벨링 차수
IMAGE_ID	VARCHAR	No	이미지 ID
LABEL_CD	NUMBER	No	라벨값
REG_DT	DATE	Yes	등록일시

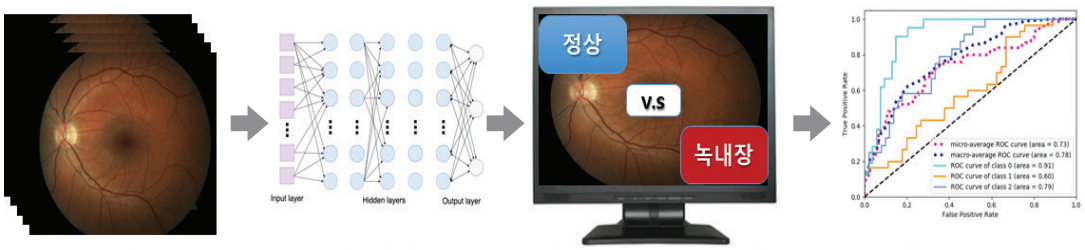
[그림 6-7] 리뷰모드 (라벨별 재검토)



6.5. 인공지능 판독 시범 모델 개발 및 데이터 검증

- (안저) 영상 처리에 적절한(이미지 처리 분야에서 성능 향상을 기대할 수 있는) 전통적 머신러닝 기법이 없음
  - 영상 matrix를 PCA나 기타 차원 축소 방법을 통해서 일정한 feature 수로 감소 시킨 뒤에, SVM등을 사용하는 방법이 존재할 수도 있으나, 이 또한 성능이 보장되지 않고 (만약 성능이 어느 정도 나왔다면 지금의 영상분야에서 딥러닝을 사용하지 않았을 것) 머신러닝을 사용해 분석 성능을 향상 시켜야 하기 때문에 더 많은 시간과 노력이 필요함
  - 해당 과제에서 유효성 검증은 Kfold Cross Validation이 적합하다 판단됨
  - 구축된 안저 영상데이터를 학습용과 검증용 그리고 테스트용 데이터로 분리시켜, 각각 80%, 10%, 10%로 나누어서 판독 학습모델을 개발함. 학습용 데이터에 검증용 데이터를 포함시키지 않도록 하여, 이후 유효성 평가하여 임상 의사결정지원 지식베이스 활용성 검증
  - 안저영상 판독 학습모델을 검증하기 위해 n-fold cross validation을 수행하며 이때 전체데이터를 학습용 90%, 테스트용 10%로 나누고 학습용 데이터에서 n개의 데이터 셋으로 랜덤하게 분리하여, n-1개 셋으로 모델을 학습하고, 1개의 셋으로 검증함
  - 이미지 분석에서 높은 성능을 보여주고 있는 딥러닝 알고리즘 중 ResNet, Inception ResNet v2, Xception, ResNeXt, DenseNet 등을 기본 모델로 사용하여 인공지능 기반의 광각 안저영상 판독 시범모델을 개발할 예정임
  - 기타 방법론으로는 squeezed excitation, cutout / erase (전처리), shortcut and concatenation 등을 활용하려고 함
  - 각 안저질환별 학습 모델의 목표달성 지표 예제는 아래 표 6-7과 같음

[그림 6-8] 유효성 검증 구조도 (안저촬영영상 예시)



안저 영상 | 딥러닝 모델링 | 지능형 녹내장 선별검사 실증실험 | 녹내장 선별검사 성능평가

[그림 6-7] 질환별 학습모델 목표달성 지표 예제

데이터명	녹내장	당뇨성망막변증	황반변성	정맥폐쇄
민감도	≥85%	≥85%	≥85%	≥85%
특이도	≥85%	≥85%	≥85%	≥85%
정확도	≥85%	≥85%	≥85%	≥85%
AUROC	≥0.85	≥0.9	≥0.9	≥0.9

7. 인공지능 학습용 데이터 공개·개방 방안

7.1 의료 데이터 제공을 위한 법 제도 현황

- 공공데이터의 제공 및 이용 활성화에 관한 법률에 정의된 바로 정해진 대상, 절차, 방법을 따라야만 의료 데이터를 제공할 수 있음
- 생명윤리법에 의거하여 비식별화된 의료 데이터 제공에 대해서는 판례, 법률, 가이드라인을 종합하여 준수해야함
- [판례] 서울중앙지방법원 2017. 9. 11. 선고 2014가합508066, 2014가합538302 판결
  - 적절한 비식별화 조치가 이루어졌는지 여부는 다음과 같은 점을 종합적으로 고려하여야 한다고 판시함
  - 원본 데이터의 특성
  - 비식별화된 정보가 사용된 특정한 맥락이나 상황
  - 비식별화 조치에 활용된 기법?세부기술의 수준
  - 비식별화된 정보를 제공받은 자의 이용목적 및 방법, 이용기간, 전문지식이나 기술력 경제력에 따른 재식별화 능력
  - 비식별화된 정보를 제공받은 자가 재식별화로 얻을 수 있는 이익의 유무
  - 비식별화된 정보를 제공받은 자의 개인정보 보호 수준
  - 비식별화된 정보와 외부정보 사이의 결합 가능성
  - 비식별화된 정보를 제공한 자와 제공 받은 자의 관계
  - 비식별화된 정보에 대한 접근권한 관리 및 접근통제 등
    (출처)http://www.legalline.co.kr/news/view.php?no=220

- [법률]생명윤리 및 안전에 관한 법률(2019.10.24.)
  - 제18조(개인정보의 제공)
    ① 인간대상연구자는 제16조제1항에 따라 연구대상자로부터 개인정보를 제공하는 것에 대하여 서면동의를 받은 경우에는 기관위원회의 심의를 거쳐 개인정보를 제3자에게 제공할 수 있다.
    ② 인간대상연구자가 제1항에 따라 개인정보를 제3자에게 제공하는 경우에는 익명화하여야 한다. 다만, 연구대상자가 개인식별정보를 포함하는 것에 동의한 경우에는 그러하지 아니하다.
    * 고유식별정보는 생명윤리법에 해당 사항 없음
- [가이드라인]개인정보 비식별 조치 가이드 라인(2016.7.1, 정부부처합동) / (단계 : 사전검토 -> 비식별조치 -> 적정성 평가 -> 사후관리)
  * 사후관리 = 기술적, 물리적, 관리적 보안 대책 준수(국립암센터 정보보안 체계안에 있어 준수 되고 있음)

7.2 데이터 내 개인정보 비식별화

- 인공지능 학습용 데이터 내 개인정보 비식별화 처리를 통해 재식별 할수 없는 형태로 공개하며, 비식별화 방안 예시는 아래 그림 7-1과 같음
  - 1단계: 추출 조건 확인 및 DICOM 헤더 내 포함된 환자번호, 성별, 나이 등 관련 정보를 제거하여 1차

영상 추출

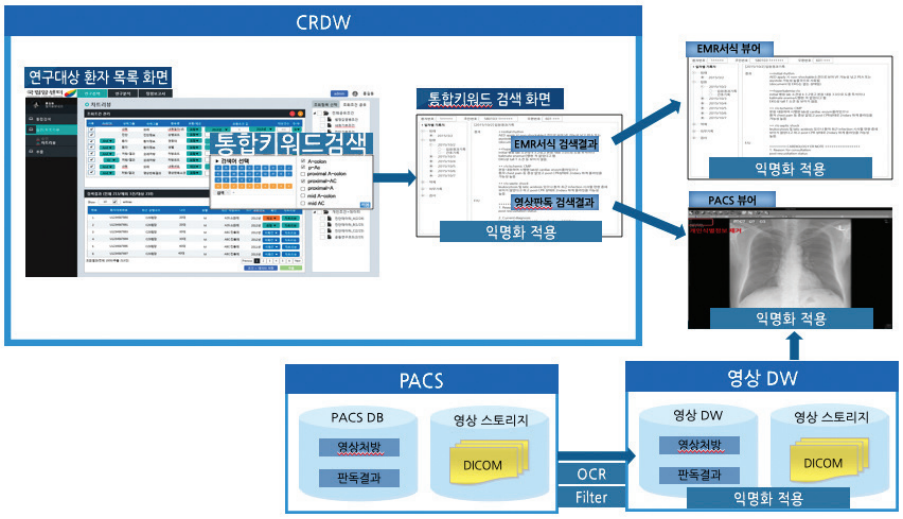
– 2단계: 추출 대상 파일 이관 (예: 운영기→영상처리 워크스테이션)

– 3단계: 비식별화 프로그램을 통한 DICOM 헤더 내 모든 식별자 (준식별자 포함) 제거

– 4단계: 이미지 자체에 포함되어 있는 텍스트 유관확인 및 블랙마킹 처리

– 5단계: 기관 보안 담당자를 통한 무작위 샘플링 유관 확인 및 공개

[그림 7-1] 비식별화 방안 예제



7.3 의료 지식베이스 제공

- 공공 및 민간의 의료 인공지능 산업 육성을 확장하기 위해서 유방암과 안저 질환을 주요 질환으로 선정하고 두 질환의 영상촬영 데이터를 인공지능 학습용 데이터(지식베이스)로 구축하여 범용성이 높은 의료 지식베이스 공개를 목표
- AI 오픈 이노베이션 허브 (http://aihub.or.kr)를 통해 메타 데이터 정보만 제공하며, 의료 지식베이스를 활용하여 분석을 희망하는 경우 사전에 신청을 하고 내외부 전문가 위원회를 구성하여 적정성 평가를 평가 후 승인된 사용자에게 한하여 폐쇄된 영역에서 분석 후에 분석결과만 반출

[그림 7-2] AI 오픈 이노베이션 허브

AI 데이터 구축 가이드라인

10

이상행동 CCTV 영상 AI 데이터 구축 가이드라인

MINDs Lab

	성명	소속	직위
가이드라인 제안	백훈	마인즈랩	전무
가이드라인 작성	이혜정	마인즈앤컴퍼니	이사
	김정욱	마인즈랩	연구원
	서상원	마인즈랩	팀장
	송혜원	마인즈랩	연구원
	박지원	마인즈랩	연구원
전문가 검토 위원	이승재	KISA	수석연구원
	여명훈	한국디지털 CCTV 연구조합	과장
	배동석	TTA	책임
	김송이	TTA	선임

※ 주관: 마인즈랩/ 참여: 수원시, 마인즈앤컴퍼니

서 문

1. 가이드라인의 목적

공공의 안전을 위해 CCTV 역할은 점점 더 커지고 있다. 우리나라 전체 229개 지자체 중 90%가 넘는 208개 지자체에서 CCTV 통합관제센터를 구축하고 운영하고 있다. 현재 지자체 통합관제센터 운용은 주로 사람(육안)에 의존하고 있다. 24시간 교대로 모니터 요원을 운용하고 있으나, 개인별 숙련도 차이나 집중력 저하로 인한 사건 검출에 실패하는 경우가 많다. 따라서 이상행동 발생에 즉각 대응 및 집중 모니터링 가능한 지능형 CCTV 도입이 점점 더 필요해지고 있다. 하지만 현재 지능형 CCTV를 구현하기 위해 학습에 필요한 데이터셋이 턱없이 부족한 상황이다. 현재 공개된 데이터셋은 대부분 해외의 데이터셋인데, 양이 많지 않을 뿐더러 우리나라의 일상적인 패턴과는 다른 영상 데이터(총격, 폭발 등)이므로 국내 AI 연구자 및 개발자들이 적용하기에 적합하지않다. 따라서 본 과제는 지능형 영상 인식의 학습 개발에 활용하기위한 데이터셋을 구축하고 공개하여 이를 바탕으로한 선순환 생태계를 조성하며, 이를 적용하기 위한 가이드라인을 제공하기 위함이다.

2. 주요 내용 요약

이 가이드라인은 인공지능 학습용 데이터셋을 구축하고 검증, 그리고 이를 공개하고 활용하는 방법에 대한 지침을 서술한다. 인공지능 학습용 데이터셋을 구축에 관한 기술과 구축한 데이터셋에 대한 검증, 그리고 활용에 대한 방안을 제공하고, 선순환 생태계 조성을 위해 데이터셋 플랫폼을 제공하여 시스템 환경 구성방안을 제시한다.

3. 인용 문서와의 비교

- 해당 사항 없음

이상행동 CCTV 데이터 구축 사업 가이드라인

Guidelines for Build abnormal event CCTV video dataset

10

1. 적용 범위

지금은 데이터를 기반으로 AI 기술을 활용하여 새로운 제품과 서비스를 창출하는 경제 시대이다. AI와 데이터는 전체 사업에서의 혁신 성장을 가속화 할 수 있는 중요 요소이다. AI와 데이터는 4차 산업혁명 시대의 중요한 생산요소이며, 미래의 경쟁은 대규모 데이터 확보와 AI 역량 강화가 핵심일 것이다. 경쟁력을 갖추기 위해서는 민·관 협동하여 대규모 데이터를 구축하고 개방하는 것이 필요하다. 미국, 유럽 등 AI 선도국에서는 정부와 민간 협업과 투자로 대규모 데이터를 구축하여 공개하고 있다. 국내 중소·벤처기업들은 AI 학습용 원천 데이터의 확보가 어렵고, 데이터의 자체 구축과 가공에 많은 시간과 비용을 소요하고 있다. 따라서 AI 선도국과 기술 격차를 줄이기 위한 노력이 필요하다. 데이터, AI 분야별 육성전략과 융합촉진을 위한 정책을 통해 다양한 영역에서 데이터와 AI 발전을 도모해야 한다.

그 중 CCTV 데이터는 안면인식, 의료 등과 더불어 규제가 특히 심한 부문이다. 개인정보보호법 및 시행령, 시행규칙과 지방자치단체의 CCTV 설치 및 운영 규정, 공공기관 CCTV 관리 가이드라인 등에 의거하여 CCTV 화상 정보의 수집·처리 관련 행위가 엄격하게 제한된다. 정보·공공 주도 하에 공개 가능한 CCTV 영상 데이터의 구축 및 공개·유통은 민간기업의 AI 기술 발전을 위해 큰 의의를 지닌다.

공개된 해외 CCTV 영상은 총격, 폭발, 상점털이 등의 이상행동이 주를 이뤄 한국적 환경에 적용하기 어려움이 있다. 따라서 실제 지자체 CCTV 환경과 유사하고 한국 생활에 밀접한 데이터가 필요하고, 정부 차원에서 관리하는 5대 중대범죄 중심의 이상행동 정의 및 그에 따른 AI 학습용 CCTV 영상 데이터가 필요하다.

본 가이드라인의 적용은 CCTV 또는 보안 카메라로부터 입력되는 동영상에서 이상행동을 탐지하는 지능형 영상 인식 AI의 학습 개발에 활용하기 위한 데이터셋을 구축하고 시범 모델을 개발하여 데이터셋을 검증하는 것이 범위이고, 최종적으로 인공지능 학습을 위한 최적의 이상행동 CCTV 데이터셋 구축 및 공개를 목표로 한다.

2. 인용 표준

- 해당 사항 없음.

3. 용어 정의

● 3.1 이상행동(Abnormal Event)

이상행동(Abnormal Event)이란 평소 상태나 상황과는 다른 어떠한 이상이 발생했을 경우 혹은 상황을 이루는 사람이나 객체가 이상한 상태일 때를 뜻한다. CCTV 영상 데이터의 관점에서의 이상행동은 범죄 혹은 사고 상황이다.

● 3.2 레이블링(Labeling)

레이블링의 사전적 의미는 제품의 상표·품질·내용 따위를 포장(包裝)에 인쇄하거나 또는 부표(附票)를 붙여서 그것을 표시하는 일이다. 기계학습의 데이터 관점에서는 레이블링은 학습을 위해 각 데이터에 원하는 방향으로 정답을 주석으로 달아주는 것이다. 데이터들은 레이블링 작업을 통해 학습용 데이터셋으로 거듭날 수 있다.

● 3.3 크로마키(Chroma key)

크로마키(Chroma key)는 두 개의 영상을 합성하는 기술이다. A와 B 이렇게 두 개의 영상이 있으면, B 영상으로부터 '좁은 범위의 색'을 제거하거나 투명하게 만들어서 뒤의 A 영상이 비치게 할 수 있다. 이 기술은 또한 컬러 키, 색 분리 오버레이, 그린스크린, 블루스크린이라고도 일컬을 수 있다. 보통 기상 예보 방송에서 쓰이는데, 예보자가 큰 지도 앞에서 서 있는 것처럼 보여 주지만 실제로는 파란색 배경이나 초록색 배경의 스튜디오 안에 있는 것이다.

4. 약어

- **MIL** Multiple Instance Learning
- **GAN** Generative Adversarial Network
- **C3D** 3D ConvNet
- **FPS** Frames per Second
- **DB** Database

5. 이상행동 CCTV 영상 데이터 구축

이상행동 인식을 위한 학습용 CCTV 영상 데이터는 데이터 수집, 데이터 정제, 데이터 가공, 데이터셋 검수, 데이터 활용의 5단계를 거쳐 구축된다.

● 5.1 이상행동 CCTV 영상 데이터 구축 절차

[표 5-1] 데이터 구축 절차

구축 단계	세부 절차	설명
1. 데이터 수집	이상행동 및 정상행동 개념 정의	CCTV 촬영 환경 하에서 사람(1명) 혹은 사람과 사람(여러 명) 사이에 일어나는 여러 가지 동작과 행동을 구체적으로 기술함
	이상행동 및 정상행동 데이터 분량 분류	이상행동 및 정상행동 개념 정의 내용에 따라 데이터셋을 어떻게 구성할지 계획함
	이상행동 및 정상행동 시나리오 수립	이상행동 및 정상행동을 연기하여 촬영하기 위해 시나리오를 수립하여 작성함
	업체를 이용한 실내 촬영	실내 스튜디오에서 이상행동 및 정상행동을 연기하고 촬영하여 데이터를 수집함
	업체를 이용한 실외 촬영	시나리오에서 수립된 실외 장소에서 이상행동 및 정상행동을 연기하고 촬영하여 데이터를 수집함
2. 데이터 정제	영상 데이터 정제	불필요하거나 무의미한 영상 데이터를 제거하는 과정을 통해 데이터를 정제함
3. 영상 데이터 가공	크로마키 영상 합성 툴을 이용한 영상 편집	크로마키 영상 합성 툴(애프터 이펙트, 프리미어 프로)을 이용하여 추후 배경 합성이 용이하도록 영상을 편집함
	레이블링 프로세스 및 태깅을 정의	촬영한 CCTV 동영상에서 레이블링 방법 및 절차를 정의하고, 어떤 정보를 레이블링할 지에 대해 정의함
	동영상 레이블링 인력 모집	촬영한 CCTV 동영상을 레이블링할 작업자를 모집함
	레이블링 도구 제작	촬영한 CCTV 동영상을 작업자가 효율적으로 태깅하기 위해 레이블링 도구를 제작함
	동영상 레이블링	작업자가 레이블링 도구를 이용하여 CCTV 동영상을 레이블링하고 데이터셋을 구축함
	레이블링 데이터 검수	3.4단계에서 레이블링된 결과물에 대해 오류가 없는지 검수함
4. 데이터셋 검증	데이터셋 품질 검증	TTA에서 제공한 품질 산출물을 작성하여 품질을 검증함
	이상행동 검출 알고리즘을 통한 데이터셋 유용성 검증	구축한 데이터셋을 최신 딥러닝 기반의 이상행동 검출 알고리즘을 이용하여 학습하고 그 성능을 평가하여 데이터셋의 유용성을 검증함

5.1.1 데이터 수집

- 이상행동 검출을 위한 CCTV 데이터셋 구축을 위해 먼저 동작(action)과 이상행동(abnormal event)에 대해 정의하는 것이 필요
- 동작(action): 사람의 움직임 등을 구체적으로 정의하는 것
- 이상행동(abnormal event): 여러 가지 동작들이 모여서 이루어지는 행동 가운데 정상이 아닌 행동으로 정의되는 행동
- 동영상에서의 동작 인식을 다루는 여러 가지 논문과 오픈 데이터셋을 분석하고 동작의 후보들을 정리
- 동영상에서 이상행동 인식을 다루는 여러 가지 논문을 분석하고 이상행동의 후보들을 정리
- 동작과 이상행동을 객관적 및 실제적으로 정의하기 위해, 참여기관인 수원시에서 관리중인 CCTV를 모니터링하는 관제부서에서 정의하는 이상행동을 정리
- 참고 논문 리스트는 <표 5-2>와 같음

[표 5-2] 참고 논문 리스트

No	이상행동 검출 분야 참고 논문 리스트
1	Sultani, Waqas, Chen Chen, and Mubarak Shah. "Real-world anomaly detection in surveillance videos." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
2	Sanz, D., et al. "BOSS: intelligent embedded video surveillance system with real-time transmission of video to the control centre." 8th World Congress on Railway Research, Seoul, Korea. 2008.
3	AL-DHAMARI, A. H. L. A. M., RUBITA SUDIRMAN, and NASRUL HUMAIMI MAHMOOD. "ABNORMAL BEHAVIOR DETECTION IN AUTOMATED SURVEILLANCE VIDEOS: A REVIEW." Journal of Theoretical & Applied Information Technology 95.19 (2017).
4	Wang, Jun, and Limin Xia. "Abnormal behavior detection in videos using deep learning." Cluster Computing: 1-11
5	Sabokrou, Mohammad, et al. "Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes." Computer Vision and Image Understanding 172 (2018): 88-97.
6	Ko, Kwang-Eun, and Kwee-Bo Sim. "Deep convolutional framework for abnormal behavior detection in a smart surveillance system." Engineering Applications of Artificial Intelligence 67 (2018): 226-234. in a smartsurveillance system
7	Coşar, Serhan, et al. "Toward abnormal trajectory and event detection in video surveillance." IEEE Transactions on Circuits and Systems for Video Technology 27.3 (2016): 683-695.
8	Luo, Weixin, Wen Liu, and Shenghua Gao. "A revisit of sparse coding based anomaly detection in stacked rnn framework." Proceedings of the IEEE International Conference on Computer Vision. 2017
9	Liu, Wen, et al. "Future frame prediction for anomaly detection-a new baseline." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
10	Arifoglu, Damla, and Abdelhamid Bouchachia. "Activity recognition and abnormal behaviour detection with recurrent neural networks." Procedia Computer Science 110 (2017): 86-93.
11	Xu, Honghua, et al. "Movement Human Actions Recognition Based on Machine Learning." International Journal of Online Engineering 14.4 (2018).
12	Almeida, Aitor, and Gorka Azkune. "Predicting human behaviour with recurrent neural networks." Applied Sciences 8.2 (2018): 305.
13	Han, Yamin, et al. "Going deeper with two-stream ConvNets for action recognition in video surveillance." Pattern Recognition Letters 107 (2018): 83-90.

- 이상 여러 소스에서 취합된 동작 및 이상행동에 대한 전체 리스트(full list)는 다음 <표 5-3>와 같음

[표 5-3] 동작/이상행동 전체 리스트

소스	동작/이상행동 리스트
[동작]	
논문 및 오픈 데이터셋	걷기, 뛰기, 서있기, 던지기, 제스처, 운반하기, 줍기, 조깅하기, 다가가기, 만나기, 나눠지기, 손 흔들기, 박수치기, 줄넘기하기, 위로 뛰기, 구부리기, 안기, 밀기, 차기, 주먹으로 치기, 악수하기, 가리키기, 허리 굽히기, 손 흔들기, 박수치기, 불안정한 동작, 서성거림 등
[이상 행동]	
논문	학대, 체포, 방화, 폭행, 교통사고, 절도, 폭발, 싸움, 강도, 총격, 절도, 들치기, 공공기물파손, 추격, 기절, 미행, 배회, 자동차 파손 등
KISA	배회, 침입, 유기, 쓰러짐, 싸움, 방화, 침투, 폭발 등
수원시	강력범죄, 교통사고, 절도, 성범죄, 화재, 재산손괴, 경찰공조, 주취자, 오인범죄, 데이트폭력, 자살미수 등

- 논문, KISA, 수원시 CCTV 자료 등을 기준으로 이상행동 15종 선정

[이상행동 (Abnormal event) 15종]
① 방화, 화재, 폭발 ② 폭행 ③ 싸움 ④ 교통사고 ⑤ 절도 ⑥ 기물파손 ⑦ 실신 ⑧ 배회
⑨ 주취행동(비틀거림, 주저앉음, 음주운전) ⑩ 침입, 월담 ⑪ 유기, 무단투기 ⑫ 자살 ⑬ 강도 ⑭ 납치 ⑮ 미행

- 이상행동 내 후보 동작 17종 선정

[이상행동 내 후보 동작 (Action) 17종]
① 걷기 ② 뛰기 ③ 밀기 ④ 안기 ⑤ 위로 뛰기 ⑥ 서있기 ⑦ 던지기 ⑧ 놓기 ⑨ 줍기 ⑩ 주먹질 ⑪ 발길질 ⑫ 쓰러짐
⑬ 운반하기 ⑭ 다가가기 ⑮ 웅크리기 ⑯ 두리번거리기 ⑰ 불안정한 동작

• 최종 이상행동(Abnormal event) 도출

[표 5-4] 최종 이상행동 리스트

No	이상행동	출처	객체	객체수	이상행동 내 동작									
1	폭행	논문, 수원시	사람	2인 이상	휘두르기	차기	밀치기	당기기	던지기	쓰러짐	위협하기	찌르기		
2	싸움	논문, 수원시, KISA	사람	2인 이상	휘두르기	차기	밀치기	당기기	던지기	쓰러짐	위협하기	찌르기		
3	절도	논문, 수원시	사람	1인 이상	줍기	운반하기	당기기							
4	기물파손	논문, 수원시	사람, 기물	1인 이상	휘두르기	차기	밀치기	던지기						
5	실신	논문, 수원시, KISA	사람	1인 이상	쓰러짐	비틀거림								
6	배회	논문, KISA	사람	1인 이상	맴돌기	가다 서다 반복								
7	침입	수원시 KISA	사람	1인 이상	월담	당기기	걷기							
8	투기	수원시	사람, 쓰레기, 가방	1인 이상	던지기	떨어뜨림	운반하기							
9	강도	논문	사람	1인 이상	휘두르기	차기	밀치기	당기기	던지기	위협하기	찌르기	쓰러짐	빼앗기	
10	데이트 폭력 및 추행	수원시	사람	1인 이상	휘두르기	차기	밀치기	당기기	안기	만지기	위협하기	찌르기	쓰러짐	던지기
11	납치	논문, 수원시	사람	1인 이상	휘두르기	차기	밀치기	당기기	안기	위협하기	찌르기	쓰러짐		
12	주취행동	수원시	사람	1인 이상	비틀거림	주저앉음								

• 납치, 침입, 강도는 국내 CCTV 현황에서 다양한 예시는 없었으나 공공안전을 위하여 분류체계에서 확인된 시나리오의 분량을 촬영함

1) 실사(실내/실외) CCTV 동영상 촬영

[표 5-5] 실사 CCTV 동영상 촬영 시간

실내 (100시간=1200컷)*카메라 3대 = 300시간)	3600컷
실외 (200시간=2400컷)*카메라 2대 = 400시간)	4800컷
총 700시간	8400컷

* 한 컷당 평균 5분
* 컷 = 시나리오

[표 5-6] 실사 CCTV 동영상 촬영 세부 조건 및 분류

NO	이상행동	비율(수량)	촬영 장소	특징
1	폭행	10%	실내/실외	도구 (각목, 쇠파이프, 야구방망이, 식칼, 커터칼, 술병(깨진 술병), 의자)
2	싸움	13%	실내/실외	도구 (각목, 쇠파이프, 야구방망이, 식칼, 커터칼, 술병(깨진 술병), 의자)
3	절도	11%	실내/실외	소품 (공사장 자재, 핸드폰, 의류, 자동차, 오토바이, 자전거, 택배, 가방, 쓰레기, 지갑)
4	기물파손	8%	실내/실외	소품 (공사장 자재, 핸드폰, 의류, 자동차, 오토바이, 자전거, 택배, 가방, 쓰레기, 지갑)
5	실신	10%	실내/실외	도구/소품 없음
6	배회	10%	실내/실외	소품(베게, 지팡이 등)
7	침입	4%	실내/실외	도구/소품 없음
8	투기	8%	실내/실외	소품 (쓰레기, 폐수, 인형(동물, 아기 대체용), 대형폐기물(공사장 자재, 가구 등)
9	데이트폭력 및 추행	7%	실내/실외	도구/소품 없음
10	주취행동	13%	실내/실외	도구/소품 없음
11	납치	3%	실내/실외	차
12	강도	3%	실내/실외	도구(각목, 쇠파이프, 야구방망이, 식칼, 커터칼, 술병(깨진 술병), 오토바이, 지갑, 차(열쇠), 현금, 의자)

* 위의 비율은 컷 비율이며 실제 촬영 시간은 변동될 수 있음
* 납치와 강도의 경우 국내 사례가 없었기 때문에 비율을 조정
* 수원시 CCTV 영상 분류 기준에 따라 3~4%의 오차범위 적용될 수 있음
* 촬영 상황이나 주변 환경에 따라 실내촬영과 실외촬영의 비율 조정될 수 있음
* 컷 수는 700시간을 기준으로 변동될 수 있음

[표 5-7] 촬영 시나리오 예시

이상행위	세부이상행위	플롯
폭행	가해자(한 명 또는 여러 명)가 본인의 신체(팔, 다리 등)를 사용하여 일방적으로 피해자(한 명)를 구타함	여A가 여B의 발로 하체를 찬다 놀란 여B는 도망가지만 여A는 쫓아가 손으로 머리를 계속 때린다. 여3명이 그들을 말린다.
싸움	도구(흉기, 각목 등)를 활용하여 서로에게 폭력을 가함	남A와 남B가 골목에서 언쟁을 벌이고 있다. 남A가 남B를 밀친다. 남B가 남A에게 주먹을 휘두른다. 남A 계속 맞다가 남성B를 밀치고 흉기를 꺼내서 휘두른다. 주변 남성 2명이 남성B를 말리고 주변 여성 2명은 소리를 지른다.
절도	타인 모르게 또는 허락 없이 금전적 가치가 있는 물건(돈, 핸드폰, 지갑 등)을 주워서 빠르게 도망치거나 물건을 주워서 태연하게 걸어감	바닥에 주취자의 가방이 떨어져있다. 지나가던 남1이 그것을 발로 찬다. 주취자와 가방이 멀리 멀리 떨어졌다. 남1이 그것을 주워 들고 사라진다.
기물파손	신체를 사용하여 기물을 파손함	남A가 학원을 마치고 집에 가는 길이다. 누군가와 통화를 한다. 화가 나있는 상태이다. 그러다 갑자기 편의점 앞베너를 발로 찬다. 베너가 넘어진다.
실신	걷다가 혹은 앉아있다가 바닥으로 쓰러짐	노인이 걷다 두통을 호소하며 비틀거리다 쓰러진다. 지나가던 남자 두 명이 노인에게 달려온다.
배회	가다 서다를 반복	여자가 가방을 끌어안고 화면 오른쪽에서 왼쪽으로 걷는다. 잠시 후 멈춰선다. 다시 걷는다.
침입	벽을 기어올라 월담을 함	검은색 옷을 입은 남자가 골목에 진입한다. 주변을 확인한 후 대문 앞으로 다가가 초인종을 누른다. 대문 근처에서 잠시 기다리다 다시 초인종을 누른다. 응답이 없자 주변을 확인한 후 담을 기어올라 넘는다.
투기	쓰레기(봉투에 담긴 것, 가전, 가구 등의 생활용품), 폐수 등을 정해진 장소가 아닌 곳에 던지거나 버리고 감	남1이 지나가다 가방 속에서 쓰레기 봉지를 꺼내 쓰레기들이 쌓여 있는 곳에 던지고 간다.
강도	신체나 도구를 이용한 폭행이나 폭언따위로 심신에 위해를 가하거나, 가한다고 예고하며 주머니 가방 따위를 뒤져서 재물을 받음	남1이 여1의 뒤에서 목을 조르며 맨손으로 위협한다. 여1이 주머니에 있던 지갑과 차키 핸드폰을 남1에게 준다.
데이트폭력 및 추행	남성이 신체를 활용하여 여성을 구타하거나 (머리채를 잡거나, 뺨을 때리거나, 밀침 등) 혹은 도구(벽돌 등)를 사용하여 여성에게 위협함	남A와 여A가 나란히 길을 걷는다. 여A는 전화통화를 하고 있다. 전화를 끊자 남A가 여A에게 화를 낸다. 여A가 맞받아친다. 남A가 여A를 밀친다. 여A가 바닥에 나뒹군다. 여A가 일어나려고 하자 남A가 다시 밀어서 넘어뜨린다. 여A가 남A를 보며 마구 소리를 지른다. 그러자 남A가 여A를 발로 찬다.
납치	강제로 끌고 가서 차, 특정 장소에 감금	여자A는 주차되어 있는 차를 향해 걸어간다. 팔에 걸려있던 핸드백 안에서 차키를 찾는지 뒤적거린다. 차키를 꺼내 차문을 여는 순간, 뒤에서 흉기를 든 남자A가 여자A의 입을 막고 뒤로 당긴다. 뒷문을 연 남자A는 여자A를 차 안으로 밀어넣는다. 뒷문을 거칠게 닫고 남자A는 운전석에 탄다.
주취행동	주저앉음	여자 2명이 여1이 술에 취해서 주저앉아있는 걸 발견한다. 일면식도 없는 사람이지만 걱정을 한다. 가서 여성을 지켜보면서 어떻게 부축을 할지 이야기한다.

- 촬영 업체를 섭외하여 시나리오를 기반으로 실제 영상 내용을 참조하여 사람 동작과 이상행동을 재현하여 직접 제작
- 최대한 현실과 근접한 영상 제작을 위해 실제 CCTV 영상을 참고하여 시나리오를 제작하여 촬영

[표 5-8] 피촬영자 분류 예시

회차	씬 번호	연기자 연령대	인원수	분류
1	24	30남2	2	폭행
	212	30남1		주취행동
	208	30남1		주취행동
2	191	30남1. 30여1	2	데이트폭행
	216	30남1. 30여1		주취행동
	209	30여1		주취행동
3	22	20여2	2	폭행
	233	20여2		주취행동
	245	20여1		실신
4	6	30남1. 50남1	2	폭행
	46	50남1		싸움

- 피촬영자는 촬영 업체에서 섭외한 전문 배우들로 구성되며 10대, 20대, 30대, 40대, 50대, 60대로 주취자, 치매노인, 임산부 등의 배역을 맡아 이상행동을 재현
- 현재 데이터 분류를 토대로 정의한 12가지의 이상행동의 종류 (폭행, 싸움, 절도, 기물파손, 실신, 배회, 침입, 투기, 데이트폭행 및 추행, 주취행동, 납치, 강도)로 총 8400개의 시나리오로 촬영 (이상행동의 종류 업데이트 중이므로 추후에 이상행동이 추가되거나 제외될 수 있음)
- 촬영 대상자에게 반드시 개인정보 활용 동의서 작성 등의 법적 절차를 거쳐 촬영 및 데이터를 공개하며, 법적 의무사항을 준수할 수 있도록 절차와 동의서 등에 대한 법률 자문을 구함

[그림 5-1] 개인정보제공동의서

㈜마인즈랩 (ver. 1.0, 2019.08)

개인 식별 데이터 공개에 대한 동 의 서

(이상행동 CCTV 학습용 데이터)

과제명 : 이상행동 CCTV 영상 AI 데이터 구축

1. 나는 본 과제를 통해 얻어진 본인의 데이터를 과제 결과 공유의 목적으로 공개하는 것에 대한 설명을 들었습니다.

2. 나는 이 촬영을 통해 얻어진 본인 식별 데이터를 연구목적으로 공개하는 것에 대하여 자발적으로 동의합니다.

3. 나는 이 촬영에서 얻어진 나에 대한 정보를 현행 법률과 연구윤리심의위원회 규정이 허용하는 범위 내에서 과제 수행기관이 수집하고 처리하는데 동의합니다.

4. 나는 과제 수행기관이나 위임 받은 대리인이 과제를 진행하거나 결과 관리를 하는 경우와 과제 수행 기관에서 실태 조사를 하는 경우에는 비밀로 유지되는 나의 개인 신상 정보를 직접적으로 열람하는 것에 동의합니다.

5. 나는 언제라도 이 촬영의 참여를 철회할 수 있고 이러한 결정이 나에게 어떠한 해도 되지 않을 것이라는 것을 압니다.

6. 나의 서명은 이 동의서의 사본을 받았다는 것을 뜻하며 촬영 참여가 끝날 때까지 사본을 보관하겠습니다.

촬영 대상자	성명:	서명:	서명일:
법정대리인 (필요시)	성명: 촬영 대상자와의 관계:	서명:	서명일:
입회인 (필요시)	성명:	서명:	서명일:

- 3 -

[그림 5-2] 촬영 장소 예시

〈실내〉

〈실외〉

2) CCTV 영상 활용

- 정의한 이상행동에 해당되는 CCTV 영상을 참고
- 공개 시 개인정보보호 이슈가 발생할 수 있으므로, 수집된 영상 데이터는 기관 내부적으로 검증용 데이터셋으로만 활용
- 개인정보보호법 등 관련 법규를 검토하고 결과 제출

※ 데이터 수집의 경우 실내 촬영과 실외 촬영이 정상적인 데이터 구축 과정이며, 수원시의 비식별화된 CCTV 데이터만 활용함 (CCTV 영상 활용은 마인즈랩과 수원시의 업무 협약으로 이뤄진 사항임)

※ 검증용 데이터를 기반으로 시나리오를 만들었기 때문에 학습용 데이터와 검증용 데이터 사이의 일관성에 문제 없음

○ 5.1.2 영상 데이터 정제

- 불필요하거나 무의미한 데이터 제거하는 과정을 통해 영상 데이터 정제
- 영상 데이터가 이상행동 시나리오에 맞게 촬영되었는지 확인
- 영상 데이터의 품질 검수 및 정제

○ 5.1.3 영상 데이터 가공

1) 영상 데이터 편집

- 추후 데이터를 공개했을 때 사용자들이 다양한 영상 배경을 합성할 수 있도록 크로마키 촬영 기법 활용
- 촬영 과정에서 배우가 스크린 앞에서 연기를 하고, 편집 과정에서 스틸컷 사진 등으로 배경을 바꾸면 바꾼 배경에서 연기한 것과 같은 효과를 낼 수 있음
- 충분한 크기의 스튜디오 임대 및 다양한 각도와 거리에서 촬영할 수 있도록 3대의 카메라 준비
- 스튜디오 전체에 크로마키 배경을 설치하고 그 앞에서 연기자들이 계획된 이상행동 시나리오대로 연기하는 것을 촬영
- 영상에 다양한 배경 데이터를 합성할 수 있도록 영상 속 인물의 키를 추출한 후 초록색 배경을 씌워서 편집

2) 데이터셋 제작

- 수집한 데이터에 대해 레이블링 작업을 효율적으로 수행하기 위해 자체적으로 레이블링 툴을 개발
- CCTV 영상 데이터를 툴을 이용하여 레이블링하고 학습용 데이터셋으로 구축
- 레이블링 오류를 최소화하고 좀 더 풍부한 레이블링 정보를 삽입하기 위해 레이블링 작업을 삽입 및 검수의 2단계로 구성하고 과정을 명확히 정의
- 삽입 단계: 레이블링을 위해 각 프레임을 확인하면서 이상행동에 대한 동작(action)과 객체(object) 정보를 삽입하는 단계
- 검수 단계: 삽입된 레이블 정보와 비디오 각 프레임을 하나씩 확인하면서 올바른 레이블 정보가 정확한 시간 위치에 존재하는지, 누락된 동작이나 객체가 없는지를 확인하는 단계

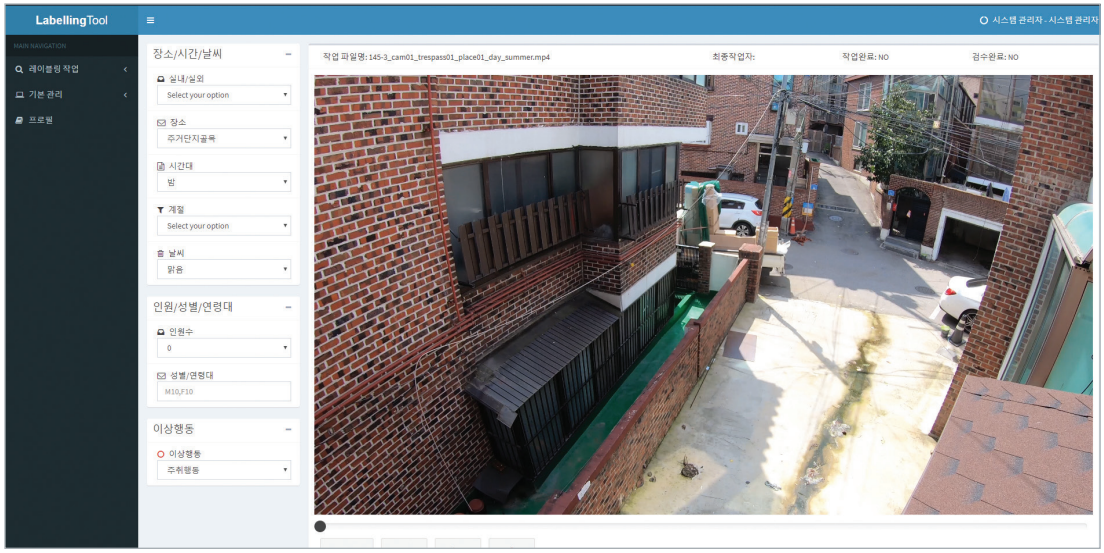
- 저장 단계: 개별 동영상에 대한 레이블링 작업이 종료되면, 데이터셋 구축 툴의 기능을 이용하여 정보 xml 파일을 생성하고, 데이터셋 DB에 저장

3) 레이블링 작업 방법

<Step 1> 레이블링 준비

- 개요: 레이블링 UI의 첫 화면으로 이 상태에서 레이블링 준비 작업을 진행함
- 작업 순서 및 방법
  - 작업자는 먼저 이상행동을 선택해야 함. 이상행동을 선택하면 그 이상행동을 포함하고 있는, 작업되지 않은 영상을 자동으로 불러옴
  - 영상을 선택하면 옆에 영상 미리보기가 떠서 대략 어떤 영상인지 확인할 수 있고, 영상에 대한 기본 정보(파일명, 해상도, 재생 시간, FPS 등)도 함께 출력
  - KISA 데이터셋과 같은 방식으로 파일명에 여러 정보가 들어있다면 파싱하여 부가적인 정보도 출력 가능함. 장소, 포함된 이상행동, 카메라 거리 등의 정보를 바로 보여줄 수 있고, 이상행동 명을 직접 입력하지 않아도 그대로 레이블링 될 수 있음
  - 이상행동에 대한 후보 동작 태그를 그 아래 보여주며, 추가/제거가 가능함: 이 기능이 필요한 이유는 프레임 별 세부 레이블링 작업 시에, 이상행동을 이루는 동작을 입력할 때 직접 타이핑하는 것이 아닌 태그를 선택하는 방식으로 업무의 효율성을 높이기 위함임
  - 시작 버튼을 누르면 다음 화면으로 진행

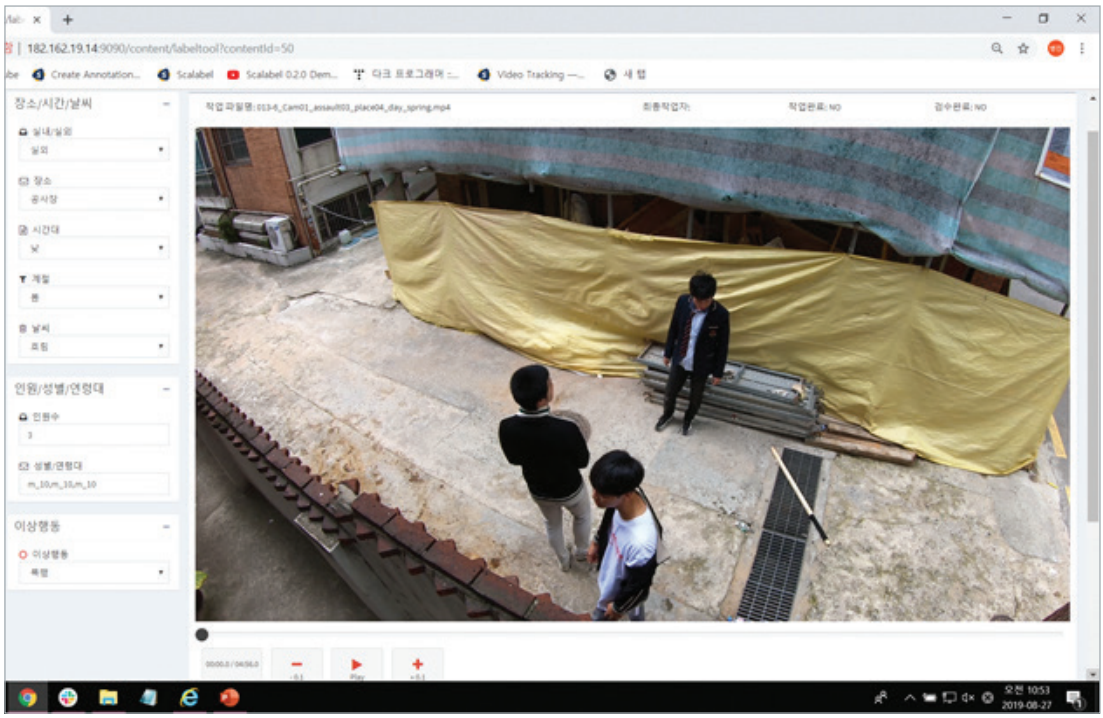
Step 1 : 레이블링 작업 준비



<Step 2> 레이블링 작업 1단계

- 개요: 해당하는 이상행동이 언제 발생했는지, 언제 종료됐는지를 레이블링 하는 단계임
- 작업 순서 및 방법
  - 기준은 정의해놓은 이상행동 정의에 대한 내용을 따름. 작업자는 객관적인 기준대로 레이블링을 진행
  - 영상을 재생하고 자유롭게 컨트롤이 가능. 프레임 컨트롤 바가 영상 바로 아래에 보여지고, 스크롤을 통해 영상을 넘기거나, 버튼을 통해 프레임 단위로 앞뒤로 넘길 수 있음
  - 재생을 정지한 상태에서 버튼을 통해 이벤트 시작 프레임과 종료 프레임을 레이블링 할 수 있음. 수정하고 싶다면 다른 위치에서 다시 버튼을 누르면 그 위치로 레이블링이 갱신됨
  - 다음 버튼을 누르면 진행한 레이블링 작업이 종료되고, 다음 세부 레이블링 작업 페이지로 넘어감

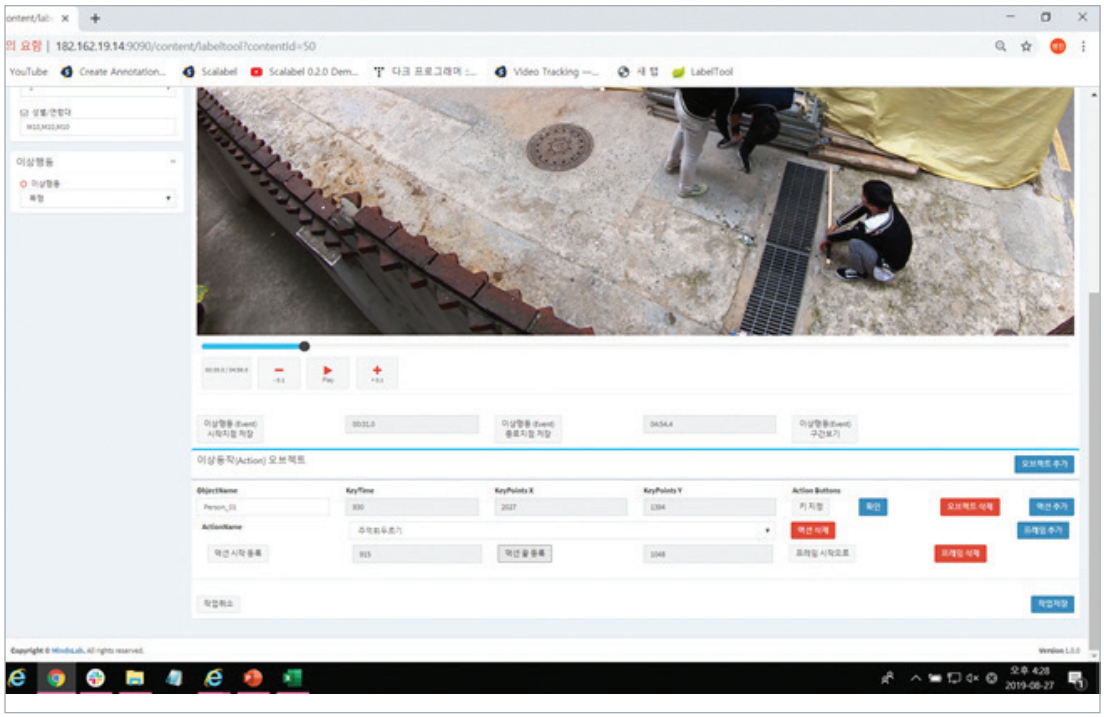
Step 2 : 레이블링 작업 1단계



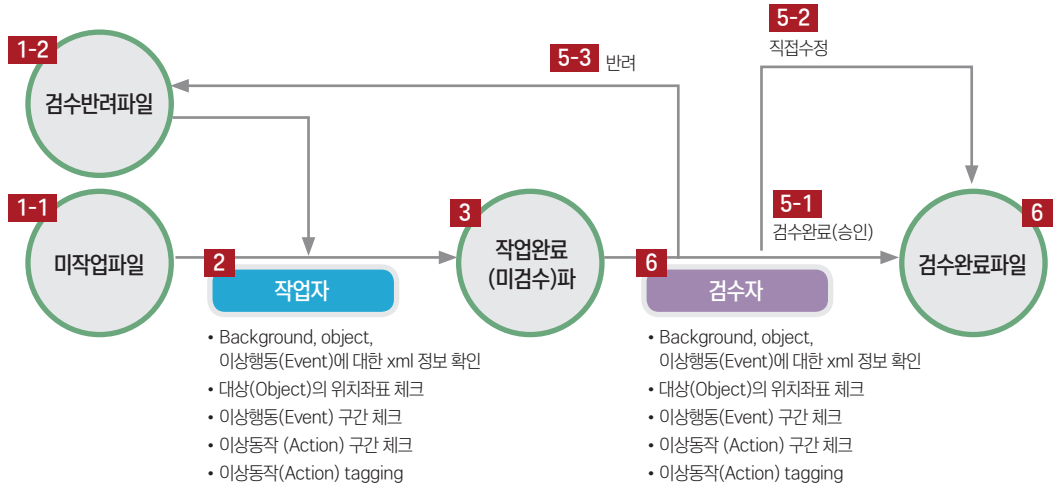
<Step 3> 레이블링 작업 2단계

- 개요: 앞서 설정한 이상행동 구간에 대해 프레임 별 세부 레이블링을 진행함 (발생된 구간 앞 뒤로 일정 시간(10~20초 정도) 포함)
- 작업 순서 및 방법
  - 프레임 별로 상세하게 레이블링 진행
  - 레이블링 해야 할 항목은 이벤트를 이루는 Object(사람)의 수, 이벤트를 이루는 Action
  - 버튼을 통해 프레임을 넘겨가며 작업함. 프레임을 넘기면 그 전 프레임에 설정한 레이블링 정보가 복사되며, 변화가 없으면 그대로 넘기면 레이블링 정보가 그대로 적용이 되고, 변화가 일어났을 시 레이블링을 수정함 : 레이블링 정보가 중복되는 프레임이 다수일 것이므로 이같은 작업 방식이 효율적임
  - 완료 버튼을 누르면 레이블링이 종료됨. 종료와 동시에 저장경로에 meta 정보를 담고 있는 xml 파일이 생성됨

Step 3 : 레이블링 작업 2단계



[그림 5-3] 데이터셋 검수 프로세스



- 작업자 작업 순서

  1) 작업자 로그인 후 검색을 통해 미작업 영상파일 호출

  2) 자동으로 생성된 메타정보 확인 (오류시 검수자에게 보고)

  3) 대상(Object)의 위치좌표 체크

  4) 이상행동(event) 구간표시

  5) 이상동작(action) 구간표시 및 annotation

  6) 저장

- 검수자 검수 순서

  1) 검수자 로그인 후 검색을 통해 작업완료 영상파일 호출

  2) 메타정보 확인

  3) 대상(Object)의 위치좌표 확인

  4) 작업한 이상행동(event) 구간확인

  5) 작업한 이상동작(action) 구간확인 / annotation 확인

  6) 저장

< 태깅 작업 매뉴얼 >

1. 영상 데이터 및 파일명

— 영상 데이터는 폭행, 싸움, 절도, 기물파손, 실신, 배회, 침입, 투기, 납치, 강도, 데이트폭력및추행, 주취행동 총 12가지로 이루어져 있습니다.

— 하나의 영상은 약 4분에서 5분정도이며, 정상행동 1분, 이상행동 2분, 정상행동 1분으로 구성되어 있습니다. (단, 이상행동별로 정상행동과 이상행동의 분량이 달라질 수는 있습니다.)

2. 메타 데이터 명칭

— 메타 데이터에 대한 명칭은 아래와 같습니다.

event	이상행동 12가지 - 폭행, 싸움, 절도, 기물파손, 실신, 배회, 침입, 투기, 납치, 강도, 데이트폭력및추행, 주취행동
action	세부이상행위 20가지 - 휘두르기, 차기, 밀치기, 당기기, 던지기, 위협하기, 찌르기, 쓰러짐, 줍기, 운반하기, 비틀거림, 맴돌기, 가다 서다 반복, 월담, 걷기, 떨어뜨림, 빼앗기, 안기, 만지기, 주저앉음
object	이상행동을 하는 사람을 오브젝트라고 합니다. 오브젝트 키 지정은 사람의 가슴을 기준으로 클릭합니다.
objectname	오브젝트 네임은 [person_숫자] 의 형식으로 태깅을 답니다.
person_1	이상행동을 가장 먼저 시작하는 사람입니다. (person_1이 화면 밖을 나갔다가 다시 들어오는 경우 새로운 오브젝트로 지정하지 않습니다.)
성별/연령대	M(남자),F(여자)10~60(연령대)로 적습니다. 여기서 M/F는 대문자로 작성합니다. 만약 오브젝트가 3명일 경우 M10,M10,M10의 형식으로 등장하는 모든 오브젝트의 성별과 연령대를 적습니다.

※ 작업은 자동 저장이 되지 않습니다. 작업을 열어둔 채로 태깅을 달지 않거나 혹은 저장하지 않은 상태로 1시간이 지날 경우 작업 내용은 사라지게 됩니다.

3. 이상행동 및 세부이상행동 xml 태깅 가이드

① 이상행동을 하는 사람(주체)의 액션만 태깅합니다.

② 객체(주체에 의해 당하는 사람)를 태깅하는 이상행동(event)은 폭행, 데이트폭행 및 추행, 납치, 강도 총 4가지이며 쓰러짐의 액션만 태깅합니다.

③ 명확한 액션이며, 이상행동(event)에 해당되는 세부이상행동(action)일 경우에만 태깅합니다.
(예 : 주취행동에서 쓰러지는 행동을 '주저앉음'으로 태깅하지 않습니다.)

④ 이상행동의 시작지점과 끝지점을 동일하게 맞춥니다.
예) 휘두르기 : 이상행동 시작(주먹을 올리기 전 차렷자세)→주먹 휘두르기→이상행동 끝(주먹을 휘두른 후 다시 차렷자세)

· 같은 동작 반복하는 사이에 다른 동작이 들어갈 경우 (예 : 주먹질 + 차기 + 주먹질)로 태깅합니다.

· 같은 동작을 반복하는 사이에 틀에 적혀있지 않는 행동이 들어갈 경우 끊어서 태깅합니다.

⑤ 오브젝트가 같은 동작을 짧은 시간 안에 연속적으로 반복할 경우 1번만 태깅합니다.
· 단, 휘두르는 여러 번의 액션 사이에 텀이 생길 때, 액션의 시간보다 텀이 길면 끊어서 태깅합니다.
ex) 휘두르기(3초) - 휴식(3초) - 휘두르기(3초) 일 경우 휘두르기를 한 번으로 태깅합니다.
ex) 휘두르기(3초) - 휴식(5초) - 휘두르기(3초) 일 경우, 중간에 쉬는 부분(5초)을 끊고 두 번 태깅합니다.

⑥ 장소
· 세워져 있는 차가 3대 이상일 경우 주차장으로 장소 태깅합니다.
· 영상과 장소의 구분이 애매한 경우 영상 파일명대로 장소 태깅합니다.
· 장소와 파일명의 장소가 명확히 다를 경우 피드백시트에 남겨주세요.

⑦ 연령대
1. 시나리오를 기준으로 메타정보 입력합니다.
2. 만약, 영상 속 등장인물과 시나리오의 정보가 다를 경우,

1그룹	10대
2그룹	20대, 30대, 40대
3그룹	50대, 60대

⑧ 연령대를 위와 같이 나누고, 영상 속 등장인물이 그룹 내에 포함되는 경우에는 시나리오대로 작성합니다.
그룹 내에 포함되지 않는 경우 피드백 시트에 남긴 후 수정합니다.
예 1) 시나리오 : 30대, 영상 속 배우 : 20대 일 경우 시나리오 대로 메타정보를 입력합니다.
예 2) 시나리오 : 50대, 영상 속 배우 : 10대 (초등학생, 유치원생)일 경우 영상 속 배우의 나이로 메타정보 입력합니다.
· 교복을 입은 사람은 10대로 메타정보 입력합니다.

⑨ 인물이 화면 밖으로 나가는 경우, 오브젝트의 몸이 50% 미만으로 보일 때 태깅을 종료합니다.
· 인물이 화면 안에서 다른 물체에 가려져 이동하거나 사라지는 경우, 오브젝트의 몸이 완전히 보이지 않을 때 태깅을 종료합니다.

4. 이상행동 액션 태깅 가이드라인

1. 폭행 : 1인 또는 다수의 사람이 일방적으로 상대방의 신체에 직접적인 유형력을 행사한 것

1) 이상행동(event별) 세부동작(action)
— 휘두르기(punching) : 손이나 도구를 사용하여 상대방의 신체에 휘두르는 행위
→ 손/도구를 상하좌우로 움직이는 모든 행위는 휘두르기로 태깅합니다.
→ 손/도구가 상대방의 신체에 닿았을 때 휘두르기로 태깅합니다.
(※ 단, 주체의 주먹/도구 등이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 휘두르기로 태깅합니다.)
— 차기(kicking) : 발을 사용하여 상대방의 신체를 차는 행위
→ 발이 상대방의 신체에 닿았을 때 차기로 태깅합니다.

(※ 단, 주체의 발이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 차기로 태깅합니다.)

— 밀치기(pushing) : 손, 어깨를 사용하여 상대방의 몸을 밀치는 행위
→ 상대방의 몸에 손/어깨 등을 댄 후 앞으로 밀어버리는 경우 밀치기에 해당합니다.

— 당기기(pulling) : 손, 팔을 사용하여 상대방의 몸을 잡고 세게 당기거나 흔드는 행위
→ 상대방의 몸을 잡는 행위(멱살잡기), 잡고 흔드는 행위 모두 당기기에 해당합니다.

— 던지기(throwing) : 도구(각목, 쇠파이프, 야구방망이, 벽돌)를 상대방에게 세게 던지는 행위
(※ 폭행에 사용하는 도구를 바닥에 던지거나 허공에 던지는 액션은 던지기로 태깅하지 않습니다.)

— 위협하기(threaten) : 직접적으로 상대방의 신체에 폭력을 가하지는 않으나 손을 머리 위로 들어올려 상대방을 때리는 듯한 제스쳐를 취하는 행위
→ 이상행동을 하는 주체의 주먹/발/도구 등이 상대방의 신체에 닿지 않았을 경우 위협하기로 태깅합니다.

— 찌르기(piercing) : 도구(흉기)를 사용하여 상대방의 신체를 찌르는 행위

— 쓰러짐(falldown) : 강하게 밀쳐져 바닥으로 넘어지거나, 쓰러지는 행위
→ 폭행에서의 쓰러짐은 객체만 태깅합니다.
→ 바닥에 완전히 쓰러지지 않더라도, 신체가 수직 상태에서 바닥을 향해 낙하하게 되면 그 과정을 태깅합니다. (단, 쓰러짐을 유지하는 상태는 태깅하지 않습니다)

2. 싸움 : 2인 이상의 사람이 서로 상대방의 신체에 직접적인 유형력을 행사한 것

1) 이상행동(event별) 세부동작(action)

— 휘두르기(punching) : 손이나 도구를 사용하여 상대방의 신체에 휘두르는 행위
→ 손/도구를 상하좌우로 움직이는 모든 행위는 휘두르기로 태깅합니다.
→ 손/도구가 상대방의 신체에 닿았을 때 휘두르기로 태깅합니다.
(※ 단, 주체의 주먹/도구 등이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 휘두르기로 태깅합니다.)

— 차기(kicking) : 발을 사용하여 상대방의 신체를 차는 행위
→ 발이 상대방의 신체에 닿았을 때 차기로 태깅합니다.
(※ 단, 주체의 발이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 차기로 태깅합니다.)

— 밀치기(pushing) : 손, 어깨를 사용하여 상대방의 몸을 밀치는 행위
→ 상대방의 몸에 손/어깨 등을 댄 후 앞으로 밀어버리는 경우 밀치기에 해당합니다.

— 당기기(pulling) : 손, 팔을 사용하여 상대방의 몸을 잡고 세게 당기거나 흔드는 행위
→ 상대방의 몸을 잡는 행위(멱살잡기), 잡고 흔드는 행위 모두 당기기에 해당합니다.

— 던지기(throwing) : 도구(각목, 쇠파이프, 야구방망이, 벽돌)를 상대방에게 세게 던지는 행위
→ 바닥에 던지거나 허공에 던지는 행위는 던지기로 태깅하지 않습니다.

— 위협하기(threaten) : 직접적으로 상대방의 신체에 폭력을 가하지는 않으나 손을 머리 위로 들어올려 상대방을 때리는 듯한 제스쳐를 취하는 행위

→ 이상행동을 하는 주체의 주먹/발/도구 등이 상대방의 신체에 닿지 않았을 경우 위협하기로 태깅합니다.

— 찌르기(piercing) : 도구(흉기)를 사용하여 상대방의 신체를 찌르는 행위

— 쓰러짐(falldown) : 강하게 밀쳐져 바닥으로 넘어지거나, 쓰러지는 행위

→ 바닥에 완전히 쓰러지지 않더라도, 신체가 수직 상태에서 바닥을 향해 낙하하게 되면 그 과정을 태깅합니다. (단, 쓰러짐을 유지하는 상태는 태깅하지 않습니다)

→ 자발적인 쓰러짐(Ex. 지쳐쓰러짐 등)의 경우 태깅하지 않습니다.

3. 절도 : 타인의 재물을 절취하거나, 권리자의 동의 없이 일시적으로 사용하는 것

1) 이상행동(event별) 세부동작(action)

— 줍기(pick up) : 바닥에 떨어진 물건(가방, 지갑, 핸드폰, 현금 등)을 줍는 행위

→ 사람의 몸과 떨어져있는 물건을 가져가는 경우, 물건(오토바이, 자전거 등)에 걸려있는 물건을 가져가는 경우 줍기로 태깅합니다.

→ 예 : 차량 내부에 있는 물건을 훔치는 경우 줍기에 해당

— 운반하기(carring) : 물건(공사장 자재, 택배 등 물건)을 옮기는 행위

— 당기기(pulling) : 손, 팔을 사용하여 상대방이 소지하고 있는 물건을 잡아당기는 행위

→ 사람의 몸에 닿아있는 물건을 가져가는 모든 행위는 당기기로 태깅합니다.

→ *예 : 쓰러져있는 사람의 몸을 뒤져서 물건을 훔치는 경우 당기기에 해당.

→ 손, 팔을 사용하여 물건을 흔드는 행위

→ 차 문을 여는 행위도 당기기에 해당합니다.

→ 바닥에 있는 물건을 주운 후 걸어가거나 뛰어가는 경우, 줍기 → 운반하기 로 태깅합니다.

— 절도 대상 물건을 전달하는 행위에 대해서, 절도 행위를 한 사람과 운반하는 사람이 같은 경우 전달에 대한 태깅은 하지않습니다.

— 절도 행위를 한 사람과 운반하는 사람이 다른 경우 운반하기 직전 전달 행위에 운반하는 사람에 대해 '당기기' , '운반하기' 태깅을 합니다.

4. 기물파손 : 유기물 또는 무기물을 무단으로 사용하지 못하게 파손하는 것

1) 이상행동(event별) 세부동작(action)

— 휘두르기(punching) : 손, 팔이나 도구를 사용하여 물건(문, 유리창 등)에 휘두르며 훼손하거나 파손하는 행위

— 차기(kicking) : 발을 사용하여 물건(문, 유리창 등)을 파손하는 행위

— 밀치기(pushing) : 신체를 사용하여 물건(문 등)을 밀어서 파손하는 행위

— 당기기(pulling) : 신체를 사용하여 물건을 당겨 파손하는 행위

→ 문, 현수막 등의 물건이나 건물의 일부를 마구 흔들어 파손하는 행위

— 던지기(throwing) : 도구(벽돌, 각목, 쇠파이프 등)를 던져 물건을 파손하는 행위

→ 물건을 바닥에 던져 파손시키는 행위

5. 실신 : 병이나 충격으로 의식을 잃고 쓰러져서 한동안 일어나지 못하는 것

1) 이상행동(event별) 세부동작(action)

— 쓰러짐(falldown) : 걷다가 혹은 서 있다가 갑자기 바닥으로 쓰러져 미동없이 누워있는 행위

→ 서 있다가 쓰러질 때를 태깅합니다.

→ 바닥에 완전히 누워있는 경우에만 태깅합니다 (ex. 웅크리기의 경우 태깅하지 않습니다.)

→ 쓰러져있는 행위가 10초 이내로 유지되는 경우 10초 이내로 쓰러짐을 태깅합니다. (쓰러지는 행위가 6초에서 멈추면 6초까지 태깅)
(이때 서 있다가 쓰러질 때의 끝 프레임과 새로 쓰러짐을 태깅하는 시작 프레임의 수가 같아야 합니다.)

— 비틀거림(totter) : 바로 걷지 못하고 몸을 앞뒤양옆으로 흔들거리며 걸어가는 행위, 몸을 흔들거리면서 서 있는 행위

6. 배회 : 목적이 없이 비슷한 자리를 맴도는 것

1) 이상행동(event별) 세부동작(action)

— 맴돌기(around) : 원의 크기에 상관없이 걸음을 멈추지 않고 원을 그리며 도는 행위

— 가다 서다 반복(stop and go) : 가다 멈추다를 반복하는 행위

→ [가다 - 서다]가 두 번 반복할 경우 [가다 - 서다 - 가다 - 서다] 하나의 set로 보고 태깅합니다.
(*가다 멈추는 시간이 길어져도 상관없이 태깅합니다.)

→ 제자리에서 방향을 바꾸는 경우는 '가다'라고 판단하기 어렵습니다. 방향성을 지닌 상태로 2~3 걸음 이상 가는 경우를 [가다 - 서다] 에서 '가다'로 합니다.

7. 침입

1) 이상행동(event별) 세부동작(action)

— 월담(climbwall) : 벽을 타고 올라가는 행위

— 당기기(pulling) : 신체를 사용하여 문을 세게 당기는 행위

— 걷기(walking) : 오브젝트가 문을 열고 들어가는 경우, '걷기(walking)'로 태깅.

8. 투기 : 유기물 또는 무기물을 무단으로 버리는 것

1) 이상행동(event별) 세부동작(action)

— 던지기(throwing) : 손, 팔을 사용하여 쓰레기(쓰레기 봉투 등)를 쓰레기 더미에 던지는 행위

→ 손을 기준으로 물건이 위로 올라가면 던지기로 태깅합니다.

— 떨어뜨림(drop) : 물건을 의도적으로 떨어뜨리거나 몰래 두고 가는 행위

→ 손을 기준으로 물건이 아래로 내려가면 떨어뜨림으로 태깅합니다.

— 운반하기(carring) : 큰 물체(가구, 의자 등)를 운반하여 버리는 행위

→ 물건을 리어카 등 다른 물체에 싣고 움직이는 행위도 운반하기로 태깅합니다.

9. 강도 : 폭행이나 협박 따위로 남의 재물을 빼앗는 것

1) 이상행동(event별) 세부동작(action)

— 휘두르기(punching) : 손이나 도구를 사용하여 상대방의 신체에 휘두르는 행위

— 차기(kicking) : 발을 사용하여 상대방의 신체를 차는 행위

— 밀치기(pushing) : 손, 어깨를 사용하여 상대방의 몸을 밀치는 행위

— 당기기(pulling) : 손, 팔을 사용하여 상대방의 몸을 잡고 세게 당기거나 흔드는 행위

— 던지기(throwing) : 도구(각목, 쇠파이프, 야구방망이, 벽돌)를 상대방에게 세게 던지는 행위

— 위협하기(threaten) : 직접적으로 상대방의 신체에 폭력을 가하지는 않으나 손을 머리위로 들어올려 상대방을 때리는 제스쳐를 취하는 행위

→ 칼을 쥔 상태에서 휘두르지 않고 상대방을 향해 겨누는 액션은 위협하기로 태깅하지 않습니다.

— 찌르기(piercing) : 도구(흉기)를 사용하여 상대방의 신체를 찌르는 행위

— 쓰러짐(falldown) : 강하게 밀쳐져 바닥으로 넘어지거나, 쓰러지는 행위

→ 바닥에 완전히 쓰러지지 않더라도, 신체가 수직 상태에서 바닥을 향해 낙하하게 되면 그 과정을 태깅합니다.

— 빼앗기(steal) : 상대방이 소지하고 있던 물건을 잡아 당기거나, 가져가는 행위

(최종적으로 물건을 가져가는 경우에 태깅합니다.)

→ 바닥에 떨어진 물건을 줍는 행위 (최종적으로 물건을 가져가는 경우에 태깅합니다.)

10. 데이트폭력 및 추행 : 남녀 두 사람 중 한 사람(나자)이 일방적으로 다른 한 사람을(여자)폭행 하는 것

1) 이상행동(event별) 세부동작(action)

— 휘두르기(punching) : 남성이 손이나 도구를 사용하여 여성의 신체에 휘두르는 행위

→ 손/도구를 상하좌우로 (상하좌우) 움직이는 모든 행위는 휘두르기로 태깅합니다.

→ 손/도구가 상대방의 신체에 닿았을 때 휘두르기로 태깅합니다.

(※ 단, 주체의 주먹/도구 등이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 휘두르기로 태깅합니다.)

— 차기(kicking) : : 발이 상대방의 신체에 닿았을 때 차기로 태깅합니다.

(※ 단, 주체의 발이 상대방의 신체에 닿지 않았더라도, 폭행을 하고자 하는 의도가 명확히 보이는 경우 차기로 태깅합니다.)

— 밀치기(pushing) : 남성이 손, 어깨를 사용하여 여성의 몸을 밀치는 행위 : 상대방의 몸에 손/어깨 등을 댄 후 앞으로 밀어버리는 경우 밀치기에 해당합니다.

— 당기기(pulling) : 남성이 손, 팔을 사용하여 여성의 몸을 잡고 세게 당기거나 흔드는 행위

→ 상대방의 몸을 잡는 행위(멱살잡기), 잡고 흔드는 행위 모두 당기기에 해당합니다.

— 던지기(throwing) : 남성이 도구(각목, 쇠파이프, 야구방망이, 벽돌)를 여성에게 세게 던지는 행위

→ 도구를 바닥에 던지거나 허공에 던지는 행위는 던지기로 태깅하지 않습니다.

— 안기(hugging) : 남성이 여성의 신체에 가까이 밀착하거나, 억지로 끌어안는 행위 (어깨동무도 안기에 해당합니다.)

— 만지기 (touching) : 남성이 여성의 신체를 의도적으로 만지는 행위, 더듬는 행위를 일컫습니다.

— 위협하기(threaten) : 직접적으로 여성의 신체에 폭력을 가하지는 않으나 손을 머리 위로 들어올려 상대방을 때리는 듯한 제스쳐를 취하는 행위
→ 이상행동을 하는 주체의 주먹/발/도구 등이 상대방의 신체에 닿지 않았을 경우 위협하기로 태깅합니다.

— 찌르기(piercing) : 도구(흉기)를 사용하여 상대방의 신체를 찌르는 행위

— 쓰러짐(falldown) : 강하게 밀쳐서 바닥으로 넘어지거나, 쓰러지는 행위 (쓰러짐(falldown)은 여성에게만 태깅합니다.)
→ 바닥에 완전히 쓰러지지 않더라도, 신체가 수직 상태에서 바닥을 향해 낙하하게 되면 그 과정을 태깅합니다. (단, 쓰러짐을 유지하는 상태는 태깅하지 않습니다.)

11. 납치 : 강제 수단을 써서 억지로 데리고 가는 것

1) 이상행동(event별) 세부동작(action)

— 휘두르기(punching) : 손이나 도구를 사용하여 상대방의 신체에(얼굴, 상반신, 하반신 등 신체의 모든 부위) 휘두르는 행위

— 차기(kicking) : 발을 사용하여 상대방의 신체를(얼굴, 상반신, 하반신 등 신체의 모든 부위) 차는 행위

— 밀치기(pushing) : 손, 어깨를 사용하여 상대방의 몸을 밀치는 행위

— 당기기(pulling) : 손, 팔을 사용하여 상대방의 몸을 잡고 세게 당기는 행위, 당기며 끌고 가는 행위도 포함됩니다. (팔짱끼기는 당기기에 해당합니다.)
→ 당기기와 안기가 동시에 진행되는 경우, 안기로 태깅합니다.

— 안기(hugging) : 상대방의 몸을 강하게 끌어안고 움직이지 못하게 압박하며 끌고 가는 행위 (어깨동무도 안기에 해당합니다.)
→ 당기기와 안기가 동시에 진행되는 경우, 안기로 태깅합니다.

— 위협하기(threaten) : 직접적으로 상대방의 신체에 폭력을 가하지는 않으나 손을 머리위로 들어올려 상대방을 때리는 제스쳐를 취하는 행위

— 찌르기(piercing) : 도구(흉기)를 사용하여 상대방의 신체를 찌르는 행위

— 쓰러짐(falldown) : 강하게 밀쳐져 바닥으로 넘어지거나, 쓰러지는 행위
→ 바닥에 완전히 쓰러지지 않더라도, 신체가 수직 상태에서 바닥을 향해 낙하하게 되면 그 과정을 태깅합니다. (단, 쓰러짐을 유지하는 상태는 태깅하지 않음)

12. 주취행동 : 술에 취해 자신의 의도와는 무관하게 움직이는 것

1) 이상행동(event별) 세부동작(action)

— 비틀거림(totter) : 바로 걷지 못하고 몸을 양 옆으로 흔들거리며 서 있거나 걸어가는 행위, 스스로 걷지 못하고 몸을 부축 받으며 걸어가는 행위
→ 비틀거리며 서 있는 행위도 비틀거림으로 태깅합니다.

— 주저앉음(sitdown) : : 바닥에 몸을 흔들거리며 앉아있거나, 쪼그려 앉아있는 행위
→ 서 있다가 바닥으로 주저 앉았을 때까지 태깅합니다.
→ 주저 앉아 있는 행위가 유지되는 경우 최대 10초간 주저앉아있음을 태깅합니다.

→ 주저앉음 태깅 후 주체가 쓰러졌다가 다시 주저앉는 경우 주저앉음을 새로 태깅합니다.

→ 의자에 앉아있는 액션은 주저앉음으로 태깅하지 않습니다.

4) 데이터셋 구조

- 이상행동 검출 데이터셋은 다음과 같이 두 가지 형태의 데이터로 구성됨
  - CCTV 동영상 mp4 파일
  - 정보 xml 파일 : 각 동영상 내의 action 및 anomaly를 포함하는 event 등의 정보를 가지고 있는 xml 형식의 텍스트 파일 (정보 xml 파일 : 정보 xml 파일은 레이블링 툴의 출력임)

[그림 5-4] xml format 예시


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<annotation>
    <folder>swoon</folder>
    <filename>100-1_cam01_swoon01_place02_day_summer.mp4</filename>
    <source>
        <database>NIA2019 Database v1</database>
        <annotation>NIA2019</annotation>
    </source>
    <size>
        <width>3840</width>
        <height>2160</height>
        <depth>3</depth>
    </size>
    <header>
        <duration>00:05:14.9</duration>
        <fps>30</fps>
        <frames>9448</frames>
        <inout>OUT</inout>
        <location>PLACE02</location>
        <season>SUMMER</season>
        <weather>SUNNY</weather>
        <time>DAY</time>
        <population>1</population>
        <character>M20</character>
    </header>
    <event>
        <eventname>swoon</eventname>
        <starttime>00:02:30.6</starttime>
        <duration>00:01:26.4</duration>
    </event>
    <object>
        <objectname>person_1</objectname>
        <position>
            <keyframe>4518</keyframe>
            <keypoint>
                <x>1975</x>
                <y>1030</y>
            </keypoint>
        </position>
        <action>
            <actionname>falldown</actionname>
            <frame>
                <start>4518</start>
                <end>4675</end>
            </frame>
            <frame>
                <start>6705</start>
                <end>6790</end>
            </frame>
            <frame>
                <start>6790</start>
                <end>7107</end>
            </frame>
        </action>
    </object>
</annotation>
```

— xml 포맷 항목 설명

- folder : 상위 folder 명 (event끼리 folder를 묶으므로 event명이 folder명이 됨)
- filename : xml 파일과 매칭되는 원본 영상 파일 이름
- source
  - database : Database 버전 명
  - annotation : annotation 구조 명
- size
  - width : xml 파일에 대응되는 영상의 width값
  - height : xml 파일에 대응되는 영상의 height값
  - depth : xml 파일에 대응되는 영상의 channels값
- header
  - duration : 영상 길이
  - fps : 초당 프레임 수
  - frames : 영상의 총 프레임 수
  - inout : 실내, 실외 촬영
  - location : 영상 속 장소
  - season : 계절
  - weather : 영상 속 날씨
  - time : 시간대
  - population : 인원 수
  - character : 성별, 연령대
- event
  - eventname : 이상행동 명
  - starttime : 이상행동 발생 시각
  - duration : 이상행동 지속 시간
- object : xml 파일과 대응되는 이미지 속의 Object 정보에 대한 tag
  - objectname : object명
  - position : object의 위치 정보. object를 명시하기 위함.
    - keyframe : object가 등장한 프레임(위치를 알려주기 위함)
    - keypoint : 위 프레임에서 object의 위치(가슴(명치)부근)
  - action
    - actionname: action명
    - frame : action이 발생하는 시작 프레임과 끝 프레임을 리스트로 나타냄

5.1.4 데이터셋 검증

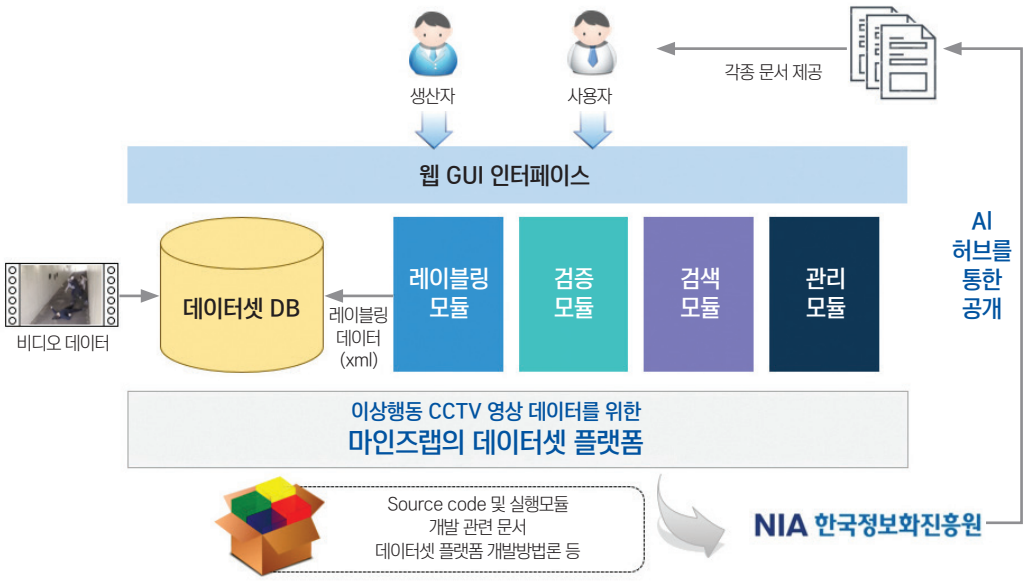
- 구축한 데이터셋 즉, 레이블링 달린 동영상들을 딥러닝 기반의 인공 지능 모델을 사용하여 학습하고 그 성능을 평가하여 데이터셋과 레이블의 유용성을 검증
- 딥러닝 기반의 인공 지능 모델은 지도 학습 방식과 비지도 학습 방식 두 가지를 선택

- 지도 학습 방식으로는 궤적 특징 방식을, 비지도 학습 방식으로는 다중 인스턴스 학습 방식으로 검증
- 촬영한 동영상을 이용하여 (4K 또는 FHD 1920*1080p, 확장자 MP4) 수원시에서 테스트베드로 이용

6. 인공지능 학습용 데이터 활용 지원

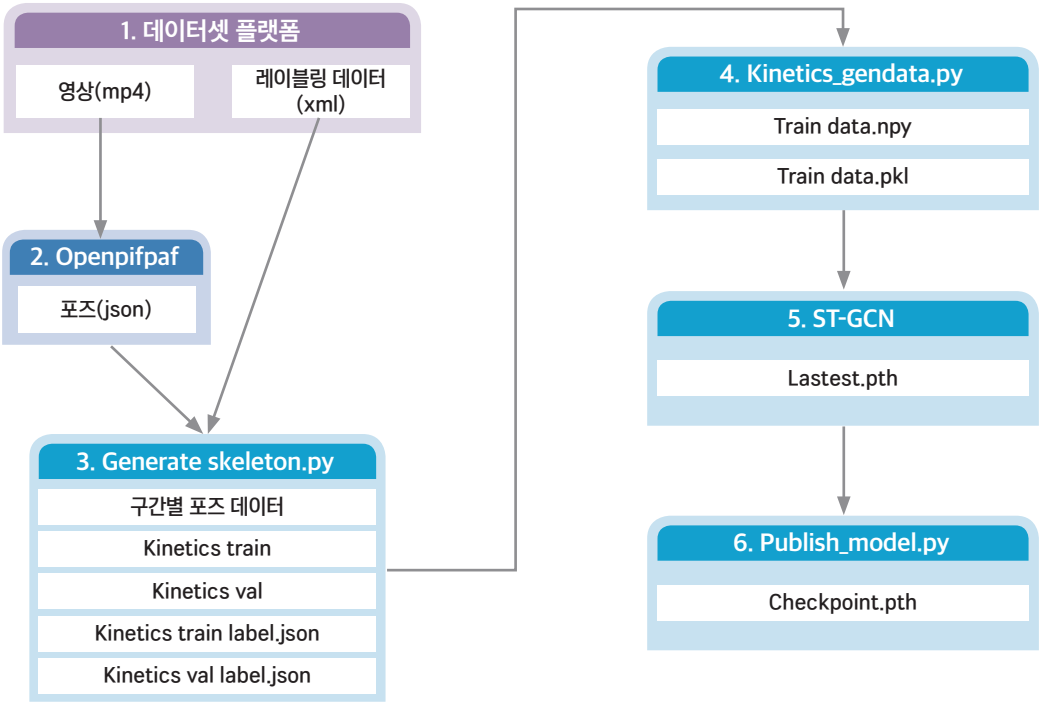
6.1 인공지능 학습용 데이터 공개·개방 방안

[그림 6-1] 인공지능 학습용 데이터 활용 시스템



- 구축한 이상행동 CCTV 영상 데이터셋은 데이터셋 플랫폼을 기반으로 배포
- 데이터셋 플랫폼은 데이터셋 DB, 레이블링 모듈, 검증 모듈, 검색 모듈, 관리 모듈 및 웹 인터페이스로 구성
- 데이터셋 플랫폼은 독립적 동작이 가능한 형태로, AI Hub 내 구축 혹은 연계할 수 있도록 충분한 기술지원을 제공
- 웹 인터페이스를 기반으로 데이터셋 플랫폼을 구현하여 데이터셋 생산자와 사용자 모두 웹 GUI를 통해 데이터셋을 추가하거나, 이용할 수 있음
- 데이터셋 플랫폼에 대한 사용 설명서 및 사용 예 등을 종합적으로 기술하여 배포
- 데이터셋 플랫폼의 소스 코드 및 모든 정보를 NIA 측에 제공하고, 이에 대한 대외 공개 여부는 NIA가 결정

[그림 6-2] 알고리즘 동작 프로세스



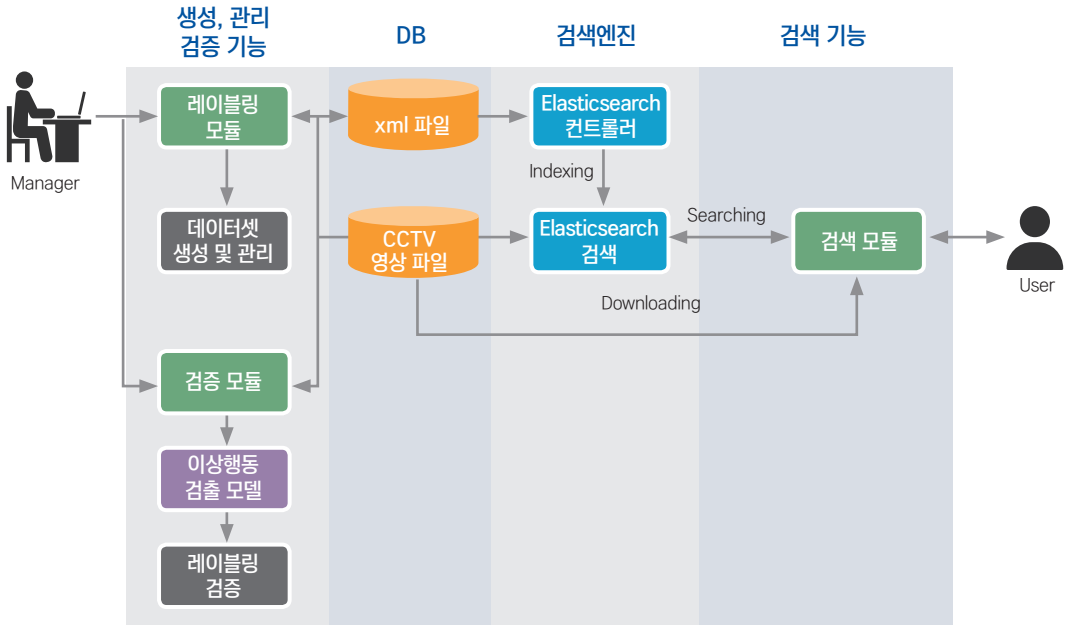
- 데이터셋 플랫폼은 작업 결과물을 학습데이터로 입력
- openpifpaf로 스켈레톤을 프레임단위로 추출하는 알고리즘을 통해 스켈레톤 데이터 생성
- generate skeleton.py 코드를 이용하여 스켈레톤 데이터와 레이블링 데이터를 매칭하여 특정 프레임 단위로 스켈레톤의 묶음 단위 추출
- 학습을 통한 traindata로 정제(이상동작 분석)
- ST-GCN을 이용한 이상행동 판독 확정
- publish_model.py 코드를 이용하여 출력

6.2 데이터셋 플랫폼

6.2.1 데이터셋 플랫폼 개요

- 데이터셋 플랫폼은 데이터셋의 생성, 검증, 관리 및 검색 등의 기능을 통합한 시스템이며 아래 (그림 6-2)와 같은 시스템을 목표로 구축됨
- 데이터셋 생성 및 관리를 위해 레이블링 모듈을 구현
- 검증 모듈을 포함하고, 두 가지의 이상행동 검출 모델을 포함하고 있으며, 이를 통해 데이터셋의 이상행동 검출의 가능성을 확인하고 레이블링의 유용성을 검증

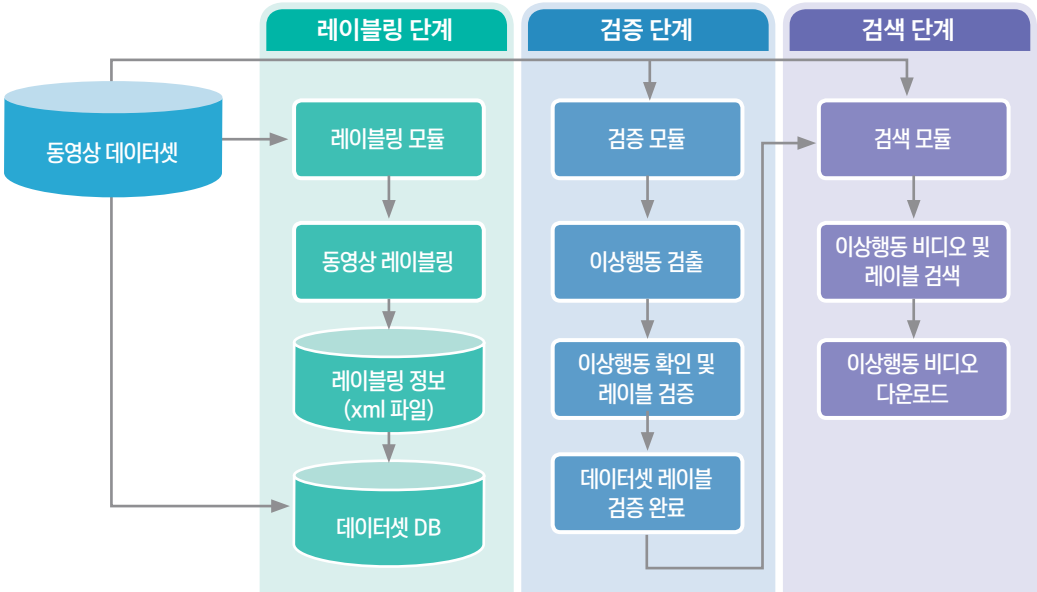
[그림 6-3] 데이터셋 플랫폼의 구성 및 동작 개념도



6.2.2 데이터셋 플랫폼 사용 순서

- 제작자는 비디오를 입력으로 레이블링 모듈을 이용하여 레이블을 추가
- 이때 자동으로 생성된 정보 xml 파일이 생성되며, 비디오와 xml 파일이 데이터셋 DB에 자동 추가
- 제작자는 검증 모듈을 이용하여 입력된 비디오에 대한 이상행동 검출 동작을 수행하고, 입력된 비디오에서 이상행동 검출 가능성을 확인 및 추가된 레이블에 대한 유용성을 검증
- 검증이 완료되면 사용자는 검색 모듈의 인덱싱 기능을 통해 입력된 비디오와 추가된 레이블이 검색이 가능하도록 구현
- 사용자는 검색 모듈을 이용하여 필요한 이상행동 비디오를 검색하고 이를 다운로드

[그림 6-4] 데이터셋 플랫폼 사용 순서



○ 6.2.3 데이터셋 플랫폼 구성 모듈

1) 레이블링 모듈

- 동영상을 재생하면서 프레임별로 동작 및 행동, 객체 정보를 입력하는 기능을 가진 모듈
- 동영상의 각 프레임, 또는 시간 간격 동안의 프레임에 미리 정의된 동작 및 객체 정보를 입력
- 관리자는 레이블링 모듈을 이용하여 데이터셋 DB에서 CCTV 동영상을 자유롭게 추가, 삭제 및 수정이 가능

2) 검증 모듈

- 지도 학습인 궤적 특징 방식과 비지도 학습인 다중 인스턴스 학습 방식 기반의 이상행동 검출 모델로 구성
- 이상행동 비디오와 정보 xml 파일로 두 개의 이상행동 검출 모델을 학습하고 입력 영상에 대해 이상행동을 검출

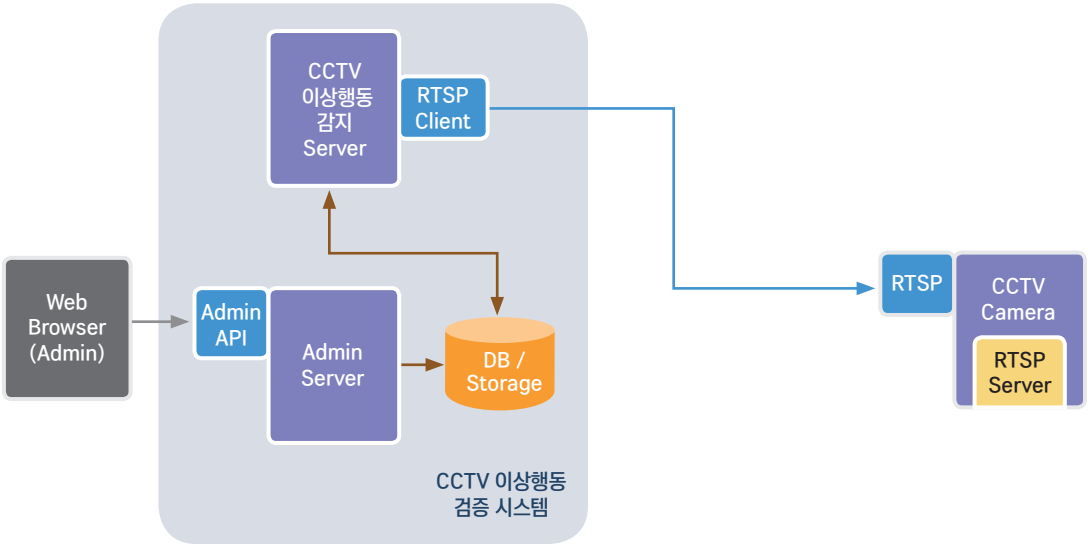
3) 검색 모듈

- 정보 xml 파일을 인덱싱 과정을 통해 검색 DB를 업데이트하며 사용자가 이상 행동에 대한 비디오를 검색하고 다운로드 받을 수 있는 기능을 제공

6.2.4 공공기관 CCTV이상행동 검증 모듈

- 수원시 도시안전통합센터에서 테스트베드 시스템 구축
- 테스트 베드 프로세스는 CCTV 에서 영상데이터를 받아 검증 모듈에서 결과값을 송출하여 관리자에게 알람을 주는 방식으로 테스트 진행
- 각각의 이상행동을 검출하는 것 보다 안전에 위협이 되는 상황 시 관리자에게 현 상황을 보고하는 것을 중심으로 테스트 진행

[그림 6-5] CCTV 이상행동 검증 모듈



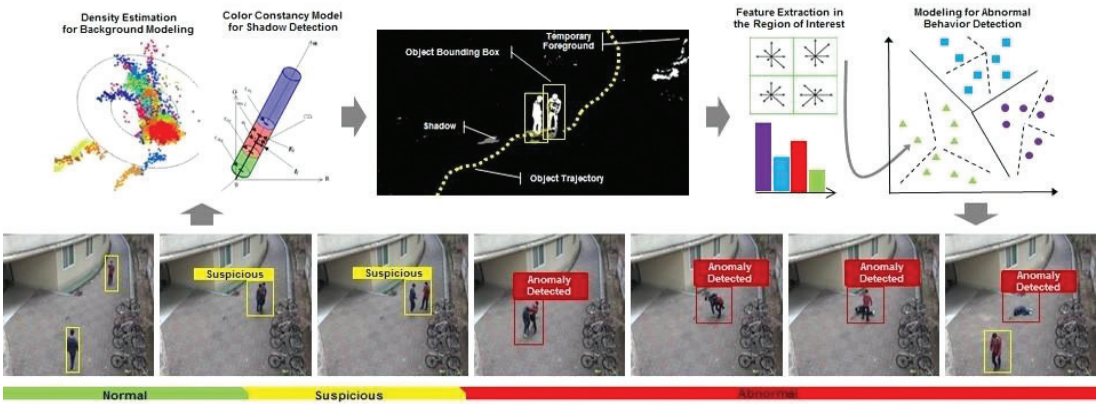
부록 I

딥러닝 기반의 인공지능 모델에 대한 개념

● I .1 궤적 특징 방식

○ Abstract

- Object Tracking 기술을 기반으로 사람이나 물체를 검출 및 추적하여 Object의 동선이나 포즈의 변화, Object간의 관계 등을 통해 이상행동을 정의하는 방법
- 이상행동을 감지하는 방법 중 가장 고전적이며 일반적인 방법이지만, 아직 성능을 높이고 해결해야 할 이슈들이 많이 존재함
  - CCTV와 같이 카메라가 고정되어있는 영상 환경에서 Object를 배경으로부터 분리하기 수월
  - Object를 검출하고 추적하는 것이 비교적 쉽기 때문에 많이 사용



○ Labeling

- Object 추적을 기반으로 하는 이상행동 검출 방식으로 Object가 추적되는 순간부터 동작
  - 검출된 Object에 대해 세부적인 포즈에 대한 레이블링 필요
  - 예를 들어, 다툼이나 폭행이라는 이상행동을 감지하기 위해서 먼저 사람을 검출 → 포즈에 대한 정의 (주먹을 휘두르고 있는지, 다른 사람을 미는 행동을 하는지 등) → 각 포즈에 대해 학습 → 행동 인식 가능
- 검출된 사람의 포즈들과 각 사람들의 관계를 통해 그 이상행동 판단
  - 포즈에 대한 학습은 필수적인 과정으로 기 정의된 각각의 포즈에 대해 매우 세부적이고 직접적인 레이블링 필요

○ Pros

- Object의 상태로 분명하게 정의할 수 있는 이상행동에 대해 정의 내리기 쉬움
  - Object의 상태에 따라 이상행동을 감지하는 직관적인 방식이기 때문
  - 예를 들어 월담, 침입, 배회 등의 이상행동은 사람의 동선을 추적하여 사람이 들어가지 않아야 하는 영역에 들어가거나, 일정 영역을 서성이는 것을 검출 및 추적하고 정의 내리기에 용이함

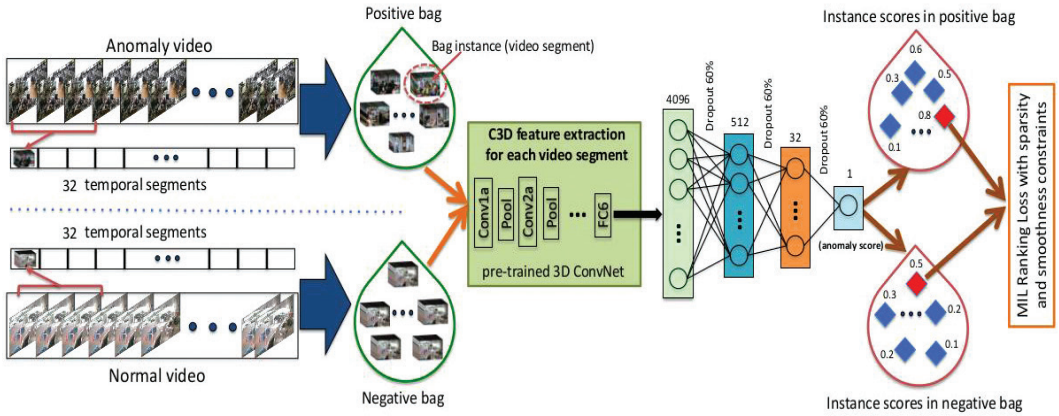
○ Cons

- Object 추적을 방해하는 요소가 들어간 혼잡한 장면에서 취약함
  - Object 추적을 기반으로 하는 방법이기 때문
  - 예를 들어 다수의 사람이 등장하여 섞이거나, 다른 Object에 의해 가려졌다가 나타나면 추적이 어렵고 다른 조명에 의해 그림자가 생기거나 급변하는 것에 대해서도 취약
  - 어떠한 상황에서도 강건하게 추적할 수 있는 추적 기술을 필요로 하지만 추적 기술은 이슈가 많은 어려운 분야임
- 정의하기 모호한 이상행동을 검출하기에는 적합하지 않음
  - 일반적으로 이상행동은 정의 내리기 모호한 경우가 많음
  - 예를 들어 다툼, 폭행, 납치 등의 이상행동은 상황을 일으키는 Object의 수도 매우 다양할 수 있고, 각 Object들의 포즈나 동선의 형태가 무제한적이므로 정의내리기 모호하기 때문

Ⅰ.2 다중 인스턴스 학습 방식

○ Abstract

- 다중 인스턴스 학습을 이용하여 비디오 내 이상행동 구간을 알아서 검출하고 학습하는 방법
  - 정상과 이상행동이 포함된 영상 데이터를 모두 활용하여 모델을 학습
  - C3D(3D ConvNet)를 이용하여 장면 내 모양과 모션에 대한 두가지 특징을 추출하고, 추출된 특징들이 이상행동을 판단하는 네트워크 구성

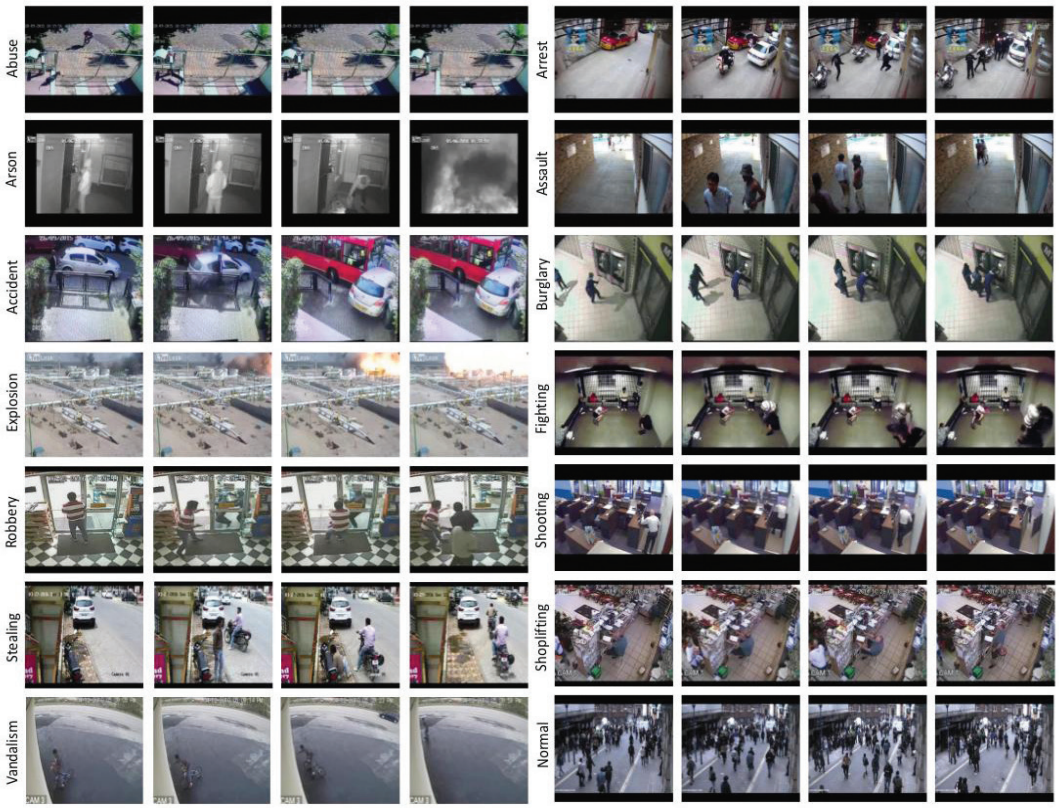


○ Labeling

- 비지도 학습 방식에 거의 가까운 약한 지도 학습 방식
  - 학습 데이터는 정상 상황만 포함되어있는 영상 부류와 이상행동이 조금이라도 포함된 영상 부류로 나눔
  - 정상 상황만이 포함되어있는 영상에 대해서는 정상으로 라벨링을 하고, 이상행동이 포함되어있는 영상에 대해서는 이상으로 레이블링
  - 알아서 이상행동에 대한 프레임 구간을 판단하므로, 성능 평가를 위해서 테스트 데이터셋에는 이상행동이 발생한 시점과 종료된 시점에 대한 라벨링이 필요함

○ Cons

- 이상행동을 판단하는 기준이 되는 학습 데이터에서 추출된 특징이 일반성이 보장되기 위해 한 상황에 대해 많은 데이터 필요
  - 이상행동은 무제한적인 특성을 가지고 있고, 같은 부류의 상황이라고 할지라도 무수히 많은 경우가 존재하기 때문
  - 이상행동에 대해 구분가능한 특징으로 표현하기 위해서는 무수히 많은 데이터를 통해 일반화해야 하기 때문



● Ⅰ.3 딥러닝 기반의 인공지능 모델에 대한 분석

- 두 가지 모델의 장단점 분석
  - 궤적 특징 방식: (단점) 무제한적이고 모호한 이상행동을 세부적으로 정의해야 하고, 일일이 레이블링 해야 함 / (강점) 검출하고자 하는 이상행동이 분명하게 정해져 있는 경우라면 집중해서 검출하기에 매우 유리한 모델임
  - 다중 인스턴스 학습 방식: (단점) 검출하고자 하는 이상행동에 대해 타겟팅을 하기 어렵고, 학습 데이터를 컨트롤 하는 것이 결과에 지대한 영향을 미치므로 학습 데이터에 대해 예민함 / (강점) 레이블링 작업이 거의 없으며, 무제한적이고 모호한 이상행동의 특징에 알맞은 모델임
- 이상행동 검출 모델 적용 방안
  - 각 모델의 장단점이 확실하므로, 12가지의 이상행동에 대해 각각 유리한 방식으로 이상행동 검출 및 레이블링의 유용성 검증한다

인공지능 학습용 데이터 구축 사업

과학기술정보통신부	김경만	과 장
	조찬영	
한국정보화진흥원	오성탁	본부장
	윤억수	팀 장
	김재욱	수 석
	유호진	수 석
	홍효진	수 석
	김민준	책 임
	이용진	책 임
	전진우	선 임
	최형인	선 임
	성연지	주 임
한국정보통신기술협회	박철순	소 장
	박정식	단 장
	신준호	팀 장
	배동석	책 임
	김송이	선 임

발 행 일 : 2020년 1월 10일(2쇄)
발 행 인 : 최 영 해
발 행 처 : 한국정보통신기술협회
경기도 성남시 분당구 분당로 47
(031·724·0114)
편집·제작 : (주)디자인여백플러스(02·2672·1535)

이 가이드북은 「과학기술정보통신부 정보통신진흥기금」을 지원받아 제작한 것으로 과학기술정보통신부의 공식의견과 다를 수 있습니다.