

'23년 인공지능 학습용 데이터 활용 가이드라인

< 한국인 피부상태 측정 데이터 >

작성 담당기관

사업 총괄

(주)카이로스랩

데이터 구축

아이이씨코리아(주)

AI모델 개발

단국대학교 산학협력단

데이터 활용

아이이씨코리아(주)

(주)카이로스랩

나준채

아이이씨코리아(주)

이은영

단국대학교 산학협력단

이정호

문서 버전

ver 3.0 ('24. 03. 11)

문서 변경 이력

버전 (개정이력)	변경일자 (개정일자)	변경내용 (개정내용)	작성자	검토자	승인자
ver 1.0	2023.12.29	한국인 피부상태 측정데이터 활용가이드 라인 초안 작성	이은영	나준채	나준채
ver 2.0	2024.02.22	문서 양식 변경	이은영	나준채	나준채
ver 3.0	2024.03.11	문서 양식 변경	이은영	나준채	나준채

목 차

5데이터 개요 및 활용 요약 1

제1장 데이터 개요 7

1. 데이터 설명서 7
2. 메타데이터 구조표 7
3. 데이터 통계 8
4. 데이터 포맷 10
5. 데이터 구축 절차 16
6. 데이터 저장 구조 및 파일 명명 규칙 25

제2장 데이터 구축 및 활용 28

1. 데이터 구축 28
2. 학습모델 29
3. 서비스 활용 34

- 별첨 : 1. 데이터 구축 현황표(데이터 구축 현황, S3 데이터 파일 현황)
2. 개인정보 수집·이용 동의서
 3. 서브라벨링 데이터 활용 가이드라인

데이터 개요 및 활용 요약

1. 공통 항목

데이터명	국문	한국인 피부상태 측정 데이터						
	영문	Korean skin condition measurement data set						
데이터 분야 ¹⁾		<input type="checkbox"/> 한국어	<input checked="" type="checkbox"/> 영상이미지	<input type="checkbox"/> 헬스케어	<input type="checkbox"/> 교통물류	<input type="checkbox"/> 농축수산	<input type="checkbox"/> 지식재산	
		<input type="checkbox"/> 문화관광	<input type="checkbox"/> 스포츠	<input type="checkbox"/> 제조	<input type="checkbox"/> 로보틱스	<input type="checkbox"/> 교육	<input type="checkbox"/> 법률	
		<input type="checkbox"/> 금융	<input type="checkbox"/> 재난안전환경					
데이터 유형 ²⁾		<input checked="" type="checkbox"/> 이미지	<input type="checkbox"/> 비디오	<input checked="" type="checkbox"/> 텍스트	<input type="checkbox"/> 오디오	<input type="checkbox"/> 센서	<input type="checkbox"/> 3D	<input type="checkbox"/> 기타
데이터 개요	소개	한국인 얼굴 이미지를 인식하여 피부상태를 예측하기 위한 데이터 안면 이미지 13,936장, 피부 상태 측정 데이터 84,688건, 라벨링 데이터 125,424건						
	구축 목적	한국인의 얼굴 부위별 다양한 각도로 촬영된 안면 이미지와 정밀한 피부측정 장비로부터 얻은 피부상태 측정 결과 데이터를 활용하여 얼굴 영역 검출, 개인별 피부 상태 예측, 피부 상태 분석 및 평가, 피부 건강 관련 연구 등의 분야 활용						
주요 키워드 ²⁾		<input type="checkbox"/> 컴퓨터비전	<input type="checkbox"/> 컴퓨터패턴인식	<input type="checkbox"/> 농업	<input type="checkbox"/> 축산업	<input type="checkbox"/> 수산업		
		<input type="checkbox"/> 자율주행	<input type="checkbox"/> 교통·모빌리티	<input type="checkbox"/> 안전	<input type="checkbox"/> 환경	<input type="checkbox"/> 제조		
		<input type="checkbox"/> 로보틱스	<input type="checkbox"/> 문화	<input type="checkbox"/> 관광	<input type="checkbox"/> 음성	<input type="checkbox"/> 자연어		
		<input type="checkbox"/> 헬스케어 의료	<input checked="" type="checkbox"/> 헬스케어 건강서비스			<input type="checkbox"/> 기타		

2. 메타데이터

데이터 형식	jpg, csv, json
데이터 출처	자체 수집
라벨링 유형	바운딩박스(이미지), 측정 데이터(텍스트)
라벨링 형식	json
데이터 활용서비스	맞춤형 화장품 추천 서비스, 피부 건강 모니터링 서비스, 피부 질환 진단 서비스, 피부 관리 교육 및 컨설팅 서비스, 피부 연구 및 제품 개발 지원 서비스
데이터 구축년도	2023년
데이터 구축량	125,424건
배포버전	ver 3.0

3. 세부데이터

데이터 설명 및 어노테이션 포맷 설명

데이터 구성																					
No	항목명	타입	구분	항목 설명	예시																
1	filename	string	필수	원천데이터 이름	"0001_01_F.jpg"																
2	id	number	필수	ID	0001																
3	date	string	선택	촬영일자	2023.05.13																
4	format	string	필수	포맷	jpg																
5	imgsize	string	필수	파일 크기(MB)	5.6																
6	width	number	필수	이미지 가로길이	1920																
7	height	number	필수	이미지 세로길이	1080																
8	device	string	필수	촬영 기기 정보	01																
9	angle	string	필수	촬영 각도	F																
10	gender	string	필수	성별	M																
11	birth	number	필수	생년월일	960129																
12	facepart	number	필수	<table><tr><td>1</td><td>이마</td></tr><tr><td>2</td><td>미간</td></tr><tr><td>3</td><td>눈가(R)</td></tr><tr><td>4</td><td>눈가(L)</td></tr><tr><td>5</td><td>볼(R)</td></tr><tr><td>6</td><td>볼(L)</td></tr><tr><td>7</td><td>입술</td></tr><tr><td>8</td><td>턱</td></tr></table>	1	이마	2	미간	3	눈가(R)	4	눈가(L)	5	볼(R)	6	볼(L)	7	입술	8	턱	1
1	이마																				
2	미간																				
3	눈가(R)																				
4	눈가(L)																				
5	볼(R)																				
6	볼(L)																				
7	입술																				
8	턱																				
13	face_box	number	필수	안면 바운딩 박스 x,y 좌표	"face_box":{ "h" : 698.14, "w" : 690.14, "x" : 189.11, "y" : 520.97, }																
14	moisture	number	선택	피부 측정 수분값																	
15	elasticity	number	선택	피부 측정 탄력값																	
16	pigmentation	number	선택	피부 측정 색소침착																	
17	pore	number	선택	피부 측정 모공값																	
18	expert	number	선택	전문가 진단	이마: 주름등급 입술: 건조도 볼: 색소 침착																

○ 공정별 데이터 포맷

구분	획득/수집 단계	정제 단계	가공 단계
데이터 구분	• 원시 데이터	• 원천 데이터	• 라벨링 데이터
데이터 형태	• 촬영-이미지(디지털 파일) • 측정-데이터(디지털 파일)	• 이미지(디지털 파일) • 텍스트(디지털 파일)	• JSON 파일
데이터 포맷	• 이미지규격: JPG(2136*3216 이상) • 텍스트	• 이미지규격: JPG(2136*3216 이상) • 텍스트	• 파일규격 : JSON

○ JSON 파일 예시

```
{
  "info": {
    "filename": "0001_01_R30.jpg",
    "id": "0001",
    "gender": "F",
    "age": 64,
    "date": "2023-09-11",
    "skin_type": 1,
    "sensitive": 0
  },
  "images": {
    "device": 1,
    "width": 2136,
    "height": 3216,
    "angle": 6,
    "facepart": 5,
    "bbox": [338, 1647, 1250, 2336]
  },
  "annotations": { "l_cheek_pore": 2, "l_cheek_pigmentation": 3 },
  "equipment": {
    "l_cheek_moisture": 71.66666666666667,
    "l_cheek_elasticity_R0": 0.209,
    "l_cheek_elasticity_R1": 0.076,
    "l_cheek_elasticity_R2": 0.6364,
    "l_cheek_elasticity_R3": 0.258,
    "l_cheek_elasticity_R4": 0.121,
    "l_cheek_elasticity_R5": 0.7091,
    "l_cheek_elasticity_R6": 0.9,
    "l_cheek_elasticity_R7": 0.3732,
    "l_cheek_elasticity_R8": 0.133,
    "l_cheek_elasticity_R9": 0.049,
    "l_cheek_elasticity_Q0": 41.8,
    "l_cheek_elasticity_Q1": 0.5618,
    "l_cheek_elasticity_Q2": 0.4466,
    "l_cheek_elasticity_Q3": 0.1152,
    "l_cheek_pore": 629
  }
}
```

4. 데이터 통계

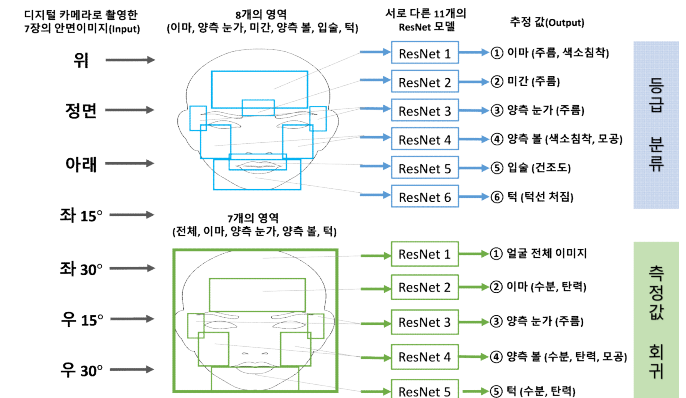
고해상도 다각도 안면이미지 13,936건, 피부상태 측정 데이터 84,688건, 메타데이터 6,432건

항목	구분	건수	비율(%)
성별 분포	남	62,478	49.81
	여	62,946	50.19
	합계	125,424	100.00
연령대 분포	10대	12,402	9.89
	20대	22,815	18.19
	30대	22,815	18.19
	40대	22,932	18.28
	50대	22,698	18.10
	60대 이상	21,762	17.35
	합계	125,424	100.00
촬영 각도 분포	디지털 카메라 정면	9,648	7.69
	디지털 카메라 좌15도	9,648	7.69
	디지털 카메라 좌30도	9,648	7.69
	디지털 카메라 우15도	9,648	7.69
	디지털 카메라 우30도	9,648	7.69
	디지털 카메라 상	9,648	7.69
	디지털 카메라 하	9,648	7.69
	스마트패드 정면	9,648	7.69
	스마트패드 좌	9,648	7.69
	스마트패드 우	9,648	7.69
	스마트폰 정면	9,648	7.69
	스마트폰 좌	9,648	7.69
	스마트폰 우	9,648	7.69
	합계	13,936	100.00
얼굴 영역 위치 분포	전체	13,936	11.11
	이마	13,936	11.11
	미간	13,936	11.11
	눈가(왼쪽)	13,936	11.11
	눈가(오른쪽)	13,936	11.11
	볼(왼쪽)	13,936	11.11
	볼(오른쪽)	13,936	11.11
	입술	13,936	11.11
	턱	13,936	11.11
	합계	125,424	100.00



5. 활용 AI모델 설명

활용모델



서비스 활용 시나리오

- 맞춤형 화장품 추천
 - 개인의 피부 상태에 맞는 화장품을 제공하여 고객들의 만족도와 구매 경험을 향상
 - 정확한 피부 상태 분석을 통한 한국인 맞춤형 미용 상품 추천 및 개발에 활용
- 피부 임상시험
 - 피부의 수분량, 모공, 주름, 탄력, 색소침착 정도 등 다양한 피부 특성을 정량화하고 평가
- 피부 질환 진단
 - 피부 측정 데이터를 활용하여 피부 질환을 정확하게 진단하고 적절한 치료 계획 수립
 - 피부 질환 예방 및 치료를 위한 정밀한 진단 및 치료에 활용
- 피부 AI 연구
 - 피부 상태 측정 AI 모델의 공개와 데이터 시각화를 통해 국내 인공지능 기술혁신을 지원
- 피부 제품 개발
 - 피부 측정 데이터를 활용하여 피부 연구를 수행하고 새로운 제품 개발에 활용

데이터 구축 규모 및
데이터 분포

6. 구축 업체정보

▶ 주관기관			
기관명	(주)카일로스랩		
담당자명	나준채		
전화번호 (유선번호) ¹⁾	070-4699-7340		
이메일 ¹⁾	david@kailoslab.com		
담당업무	사업 총괄		
▶ 참여기관 ²⁾			
기관명1	아이이씨코리아(주)		
담당업무1	데이터 수집, 정제		
기관명2	단국대학교 산학협력단		
담당업무2	전문가 피부진단, 저작도구 개발, AI 모델 개발, 학습		
문의처 ³⁾	책임자	전화번호 (유선번호) ¹⁾	이메일 ¹⁾
	담당자		
데이터문의처	나준채	070-4699-7340	david@kailoslab.com
	이은영	031-206-2336	eunyoung.lee@ieckorea.co.kr
AI모델 관련 문의처	이정호	031-8005-2646	72210297@dankook.ac.kr
	양성일	070-4699-7340	simon.yang@kailoslab.com
저작도구 관련 문의처	나준채	070-4699-7340	david@kailoslab.com
	이정호	031-8005-2646	72210297@dankook.ac.kr

제1장 데이터 개요

1. 데이터 설명서

데이터셋 명	국문	한국인 피부상태 측정 데이터
	영문	Korean skin condition measurement data set
구축 목적	한국인의 얼굴 부위별 다양한 각도로 촬영된 안면 이미지와 정밀한 피부측정 장비로부터 얻은 피부상태 측정 결과 데이터를 활용하여 얼굴 영역 검출, 개인별 피부 상태 예측, 피부 상태 분석 및 평가, 피부 건강 관련 연구 등의 분야 활용	
소개	한국인 얼굴 이미지를 인식하여 피부상태를 예측하기 위한 데이터	
활용 분야	맞춤형 화장품 추천 서비스 피부 건강 모니터링 서비스 피부 질환 진단 서비스 피부 관리 교육 및 컨설팅 서비스 피부 연구 및 제품 개발 지원 서비스	
데이터 구축 절차 요약	- 안면 이미지 데이터 수집: 고해상도 카메라를 사용하여 다각도로 촬영된 안면 이미지 데이터 수집 - 정밀 피부측정 장비 데이터 수집: 실제 환경에서 4종 이상의 정밀 피부측정 장비를 활용하여 피부 상태 측정 데이터 수집	

2. 메타데이터 구조표

데이터 영역	영상이미지	데이터 유형	이미지
데이터 형식	jpg	데이터 출처	자체 수집
라벨링 유형	바운딩박스(이미지), 측정데이터(텍스트)	라벨링 형식	json
데이터 활용 서비스	피부상태 예측 서비스 등	데이터 구축년도/ 데이터 구축량	2023년 / 125,424건

3 데이터 통계

3.1 데이터 구축 규모

- 학습용 데이터 구성(원천데이터 이미지 기준)

데이터 구축 총량(100%)	학습 데이터 (80%)	테스트 데이터 (10%)	검증 데이터 (10%)
13,936건	11,141건	1,391건	1,391건

- 원시데이터 획득

파일포맷	데이터 규모	해상도	데이터 설명
jpg	7,700 장	2136*3216	디지털카메라 이미지
jpg	3,300 장	2448*3264	스마트패드 이미지
jpg	3,300 장	2448*3264	스마트폰 이미지
xlsx, omc, jpg, dx2	86,900 건	-	측정 데이터
텍스트	6,600 건		메타 데이터

- 원천데이터

파일포맷	데이터 규모	해상도	파일 설명
jpg	7,504 장	2136*3216	디지털카메라 이미지
jpg	3,216 장	2448*3264	스마트패드 이미지
jpg	3,216 장	2448*3264	스마트폰 이미지
csv	84,688 건	-	측정 데이터
텍스트	6,432건	-	메타 데이터

- 라벨링데이터

파일포맷	데이터 규모	라벨링 유형	라벨링 대상
json	125,424건	바운딩박스(이미지), 측정데이터(텍스트), 전문가진단(텍스트)	이미지내 측정 데이터

3.2 데이터 분포

○ 원천데이터 대비 가공(라벨링) 데이터 분포

원천 데이터					가공(라벨링) 데이터				
대분류	소분류	포맷	수량	단위	대분류	소분류	포맷	수량	단위
디지털 카메라	대상자번호	jpg	7,504	장	디지털 카메라	대상자번호	json	67,536	건
스마트패드	대상자번호	jpg	3,216	장	스마트패드	대상자번호	json	28,944	건
스마트폰	대상자번호	jpg	3,216	장	스마트폰	대상자번호	json	28,944	건
합계			13,936	장	합계			125,424	건

○ 디렉토리 구성 및 수량

데이터	분류 (1단계)	분류 (2단계)	라벨링 유형	객체수	포맷 (형식)	수량	파일 비율	단위
한국인 피부상태 측정 데이터	디지털카메라	대상자번호	BBBox, 전문가진단	7,504	json	67,536	1:9	건
	스마트패드	대상자번호	BBBox, 전문가진단	3,216	json	28,944	1:9	건
	스마트폰	대상자번호	BBBox, 전문가진단	3,216	json	28,944	1:9	건
	총수량					125,424	1:9	건

4. 데이터 포맷

○ 공정별 데이터 구성

구분	획득/수집 단계	정제 단계	가공 단계
데이터 구분	• 원시 데이터	• 원천 데이터	• 라벨링 데이터
데이터 형태	• 촬영-이미지(디지털 파일) • 측정-데이터(디지털 파일)	• 이미지(디지털 파일) • 텍스트(디지털 파일)	• JSON 파일
데이터 포맷	• 이미지규격: JPG(2136*3216 이상) • 텍스트	• 이미지규격: JPG(2136*3216 이상) • 텍스트	• 파일규격 : JSON

○ 기타 데이터 구성

구분	메타 데이터
데이터 구분	• 원천 데이터의 메타데이터(피실험자 정보)
데이터 형태	• CSV 파일
데이터 포맷	• 파일규격 : CSV

○ 원천 데이터 : 라벨링 데이터 파일 관계

구분	원천 데이터 파일 : 라벨링 데이터 파일 매칭						
데이터 파일 관계	1 : 1		1 : 多	0	多 : 1		기타
기타 설명	피실험자 촬영 이미지 기준으로 원천 데이터(1)-라벨링 데이터(多) 간 매칭되는 관계						

○ 어노테이션 포맷

구분	속성명	타입	필수 여부	설명	범위	비고
1	info	Object		데이터셋정보		
1-1	info.filename	String	Y	원천데이터 이미지 파일명		
1-2	info.id	String	Y	대상자 번호		
1-3	info.gender	String	Y	대상자 성별	F, M	F: 여성 M: 남성
1-4	info.age	Number	Y	대상자 나이	0-99	
1-5	info.date	String	Y	촬영일자		
1-6	info.skin_type	Number	Y	피부타입	0~5	0: 중성 1: 건성 2: 심한건성 3: 복합건성 4: 복합지성 5: 지성
1-7	info.sensitive	Number	Y	자가 피부 민감도	0, 1	0: 아니오 1: 예
2	images	Object				
2-1	images[].device	Number	Y	촬영 기기 정보	0~2	0: 디지털카메라 1: 스마트폰 2: 스마트폰
2-2	images[].width	Number	Y	이미지너비	[0~15000]	
2-3	images[].height	Number	Y	이미지높이	[0~35000]	
2-4	images[].angle	Number	Y	이미지 촬영 각도	0~8	0: 정면 1: 상 2: 하 3: 좌15도 4: 좌30도 5: 우15도 6: 우30도 7: 좌 8: 우
2-5	images[].facepart	Number	Y	안면 얼굴 영역	0~8	0: 전체 1: 이마 2: 미간 3: 왼쪽 눈가 4: 오른쪽 눈가 5: 왼쪽 볼 6: 오른쪽 볼 7: 입술 8: 턱
2-6	images[].b-box	Array	Y	B-box 정보		좌표 List

3	annotations	Object	전문가 라벨링		
3-1	annotations[.forehead_pigmentation]	Number	이마 색소침착 등급	0~5	가공 기준 참조
3-2	annotations[.forehead_wrinkle]	Number	이마 주름 등급	0~8	가공 기준 참조
3-3	annotations[.glabellus_wrinkle]	Number	미간 주름 등급	0~6	가공 기준 참조
3-4	annotations[.l_periocular_wrinkle]	Number	왼쪽 눈가 주름 등급	0~6	가공 기준 참조
3-5	annotations[.r_periocular_wrinkle]	Number	오른쪽 눈가 주름 등급	0~6	가공 기준 참조
3-6	annotations[.l_cheek_pore]	Number	왼쪽 볼 모공 등급	0~5	가공 기준 참조
3-7	annotations[.l_cheek_pigmentation]	Number	왼쪽 볼 색소 침착 등급	0~5	가공 기준 참조
3-8	annotations[.r_cheek_pore]	Number	오른쪽 볼 모공 등급	0~5	가공 기준 참조
3-9	annotations[.r_cheek_pigmentation]	Number	오른쪽 볼 색소 침착 등급	0~5	가공 기준 참조
3-10	annotations[.lip_dryness]	Number	입술 건조도 등급	0~4	가공 기준 참조
3-11	annotations[.chin_sagging]	Number	턱선 처짐 등급	0~6	가공 기준 참조
3-12	annotations[.acne]	Array	여드름 타입 및 위치		가공 기준 참조

4	equipment	Object	정밀장비 측정 값		
4-1	equipment[.forehead_moisture]	Number	수분_이마	0~100	
4-2	equipment[.r_cheek_moisture]	Number	수분_오른쪽볼	0~100	
4-3	equipment[.l_cheek_moisture]	Number	수분_왼쪽볼	0~100	
4-4	equipment[.chin_moisture]	Number	수분_턱	0~100	
4-5	equipment[.chin_elasticity_R0]	Number	탄력_턱_R0	0~1	
4-6	equipment[.chin_elasticity_R1]	Number	탄력_턱_R1	0~1	
4-7	equipment[.chin_elasticity_R2]	Number	탄력_턱_R2	0~1	
4-8	equipment[.chin_elasticity_R3]	Number	탄력_턱_R3	0~1	
4-9	equipment[.chin_elasticity_R4]	Number	탄력_턱_R4	0~1	
4-10	equipment[.chin_elasticity_R5]	Number	탄력_턱_R5	0~1	
4-11	equipment[.chin_elasticity_R6]	Number	탄력_턱_R6	0~2	
4-12	equipment[.chin_elasticity_R7]	Number	탄력_턱_R7	0~1	
4-13	equipment[.chin_elasticity_R8]	Number	탄력_턱_R8	0~1	
4-14	equipment[.chin_elasticity_R9]	Number	탄력_턱_R9	0~1	
4-15	equipment[.chin_elasticity_Q0]	Number	탄력_턱_Q0	0~100	
4-16	equipment[.chin_elasticity_Q1]	Number	탄력_턱_Q1	0~1	
4-17	equipment[.chin_elasticity_Q2]	Number	탄력_턱_Q2	0~1	
4-18	equipment[.chin_elasticity_Q3]	Number	탄력_턱_Q3	0~1	
4-19	equipment[.l_cheek_elasticity_R0]	Number	탄력_왼쪽볼_R0	0~1	
4-20	equipment[.l_cheek_elasticity_R1]	Number	탄력_왼쪽볼_R1	0~1	
4-21	equipment[.l_cheek_elasticity_R2]	Number	탄력_왼쪽볼_R2	0~1	
4-22	equipment[.l_cheek_elasticity_R3]	Number	탄력_왼쪽볼_R3	0~1	
4-23	equipment[.l_cheek_elasticity_R4]	Number	탄력_왼쪽볼_R4	0~1	
4-24	equipment[.l_cheek_elasticity_R5]	Number	탄력_왼쪽볼_R5	0~1	
4-25	equipment[.l_cheek_elasticity_R6]	Number	탄력_왼쪽볼_R6	0~2	
4-26	equipment[.l_cheek_elasticity_R7]	Number	탄력_왼쪽볼_R7	0~1	
4-27	equipment[.l_cheek_elasticity_R8]	Number	탄력_왼쪽볼_R8	0~1	
4-28	equipment[.l_cheek_elasticity_R9]	Number	탄력_왼쪽볼_R9	0~1	
4-29	equipment[.l_cheek_elasticity_Q0]	Number	탄력_왼쪽볼_Q0	0~100	
4-30	equipment[.l_cheek_elasticity_Q1]	Number	탄력_왼쪽볼_Q1	0~1	
4-31	equipment[.l_cheek_elasticity_Q2]	Number	탄력_왼쪽볼_Q2	0~1	
4-32	equipment[.l_cheek_elasticity_Q3]	Number	탄력_왼쪽볼_Q3	0~1	
4-33	equipment[.r_cheek_elasticity_R0]	Number	탄력_오른쪽볼_R0	0~1	
4-34	equipment[.r_cheek_elasticity_R1]	Number	탄력_오른쪽볼_R1	0~1	
4-35	equipment[.r_cheek_elasticity_R2]	Number	탄력_오른쪽볼_R2	0~1	
4-36	equipment[.r_cheek_elasticity_R3]	Number	탄력_오른쪽볼_R3	0~1	
4-37	equipment[.r_cheek_elasticity_R4]	Number	탄력_오른쪽볼_R4	0~1	
4-38	equipment[.r_cheek_elasticity_R5]	Number	탄력_오른쪽볼_R5	0~1	
4-39	equipment[.r_cheek_elasticity_R6]	Number	탄력_오른쪽볼_R6	0~2	
4-40	equipment[.r_cheek_elasticity_R7]	Number	탄력_오른쪽볼_R7	0~1	
4-41	equipment[.r_cheek_elasticity_R8]	Number	탄력_오른쪽볼_R8	0~1	
4-42	equipment[.r_cheek_elasticity_R9]	Number	탄력_오른쪽볼_R9	0~1	
4-43	equipment[.r_cheek_elasticity_Q0]	Number	탄력_오른쪽볼_Q0	0~100	
4-44	equipment[.r_cheek_elasticity_Q1]	Number	탄력_오른쪽볼_Q1	0~1	
4-45	equipment[.r_cheek_elasticity_Q2]	Number	탄력_오른쪽볼_Q2	0~1	
4-46	equipment[.r_cheek_elasticity_Q3]	Number	탄력_오른쪽볼_Q3	0~1	

4-47	equipment[].forehead_elasticity_R0	Number	탄력_이마_R0	0-1	
4-48	equipment[].forehead_elasticity_R1	Number	탄력_이마_R1	0-1	
4-49	equipment[].forehead_elasticity_R2	Number	탄력_이마_R2	0-1	
4-50	equipment[].forehead_elasticity_R3	Number	탄력_이마_R3	0-1	
4-51	equipment[].forehead_elasticity_R4	Number	탄력_이마_R4	0-1	
4-52	equipment[].forehead_elasticity_R5	Number	탄력_이마_R5	0-1	
4-53	equipment[].forehead_elasticity_R6	Number	탄력_이마_R6	0-2	
4-54	equipment[].forehead_elasticity_R7	Number	탄력_이마_R7	0-1	
4-55	equipment[].forehead_elasticity_R8	Number	탄력_이마_R8	0-1	
4-56	equipment[].forehead_elasticity_R9	Number	탄력_이마_R9	0-1	
4-57	equipment[].forehead_elasticity_Q0	Number	탄력_이마_Q0	0-100	
4-58	equipment[].forehead_elasticity_Q1	Number	탄력_이마_Q1	0-1	
4-59	equipment[].forehead_elasticity_Q2	Number	탄력_이마_Q2	0-1	
4-60	equipment[].forehead_elasticity_Q3	Number	탄력_이마_Q3	0-1	
4-61	equipment[].l_perocular_wrinkle_Ra	Number	주름_왼쪽눈가_Ra	0-100	
4-62	equipment[].l_perocular_wrinkle_Rq	Number	주름_왼쪽눈가_Rq	0-100	
4-63	equipment[].l_perocular_wrinkle_Rmax	Number	주름_왼쪽눈가_Rmax	0-300	
4-64	equipment[].l_perocular_wrinkle_R3z	Number	주름_왼쪽눈가_R3z	0-100	
4-65	equipment[].l_perocular_wrinkle_Rt	Number	주름_왼쪽눈가_Rt	0-350	
4-66	equipment[].l_perocular_wrinkle_Rz=Rtm	Number	주름_왼쪽눈가_Rz=Rtm	0-200	
4-67	equipment[].l_perocular_wrinkle_Rp	Number	주름_왼쪽눈가_Rp	0-200	
4-68	equipment[].l_perocular_wrinkle_Rv	Number	주름_왼쪽눈가_Rv	0-200	
4-69	equipment[].r_perocular_wrinkle_Ra	Number	주름_오른쪽눈가_Ra	0-100	
4-70	equipment[].r_perocular_wrinkle_Rq	Number	주름_오른쪽눈가_Rq	0-100	
4-71	equipment[].r_perocular_wrinkle_Rmax	Number	주름_오른쪽눈가_Rmax	0-300	
4-72	equipment[].r_perocular_wrinkle_R3z	Number	주름_오른쪽눈가_R3z	0-100	
4-73	equipment[].r_perocular_wrinkle_Rt	Number	주름_오른쪽눈가_Rt	0-350	
4-74	equipment[].r_perocular_wrinkle_Rz=Rtm	Number	주름_오른쪽눈가_Rz=Rtm	0-200	
4-75	equipment[].r_perocular_wrinkle_Rp	Number	주름_오른쪽눈가_Rp	0-200	
4-76	equipment[].r_perocular_wrinkle_Rv	Number	주름_오른쪽눈가_Rv	0-200	
4-77	equipment[].pigmentation_count	Number	스팟개수_정면	0-300	
4-78	equipment[].r_cheek_pore	Number	모공개수_오른쪽볼	0-3000	
4-79	equipment[].l_cheek_pore	Number	모공개수_왼쪽볼	0-3000	

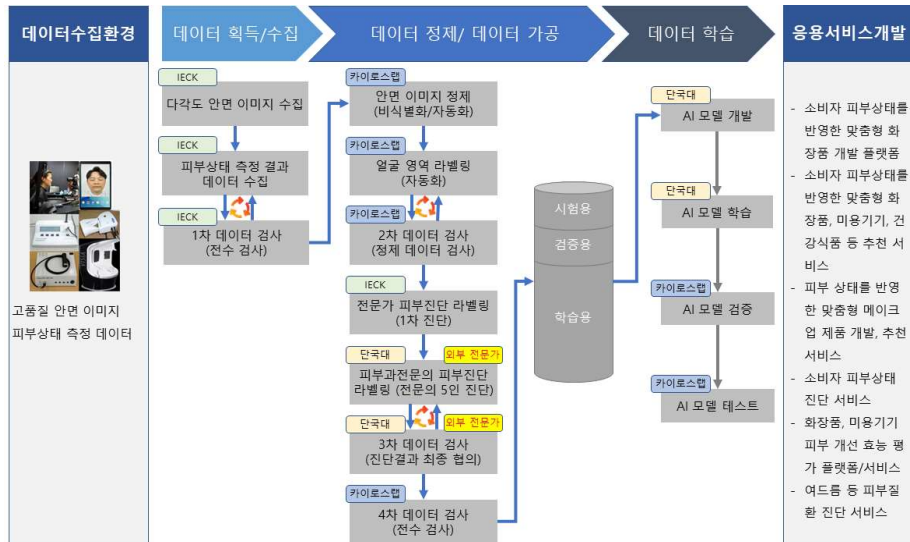
○ 어노테이션 예시

```
{
  "info": {
    "filename": "0001_01_R30.jpg",
    "id": "0001",
    "gender": "F",
    "age": 64,
    "date": "2023-09-11",
    "skin_type": 1,
    "sensitive": 0
  },
  "images": {
    "device": 1,
    "width": 2136,
    "height": 3216,
    "angle": 6,
    "facepart": 5,
    "bbox": [338, 1647, 1250, 2336]
  },
  "annotations": { "l_cheek_pore": 2, "l_cheek_pigmentation": 3 },
  "equipment": {
    "l_cheek_moisture": 71.66666666666667,
    "l_cheek_elasticity_R0": 0.209,
    "l_cheek_elasticity_R1": 0.076,
    "l_cheek_elasticity_R2": 0.6364,
    "l_cheek_elasticity_R3": 0.258,
    "l_cheek_elasticity_R4": 0.121,
    "l_cheek_elasticity_R5": 0.7091,
    "l_cheek_elasticity_R6": 0.9,
    "l_cheek_elasticity_R7": 0.3732,
    "l_cheek_elasticity_R8": 0.133,
    "l_cheek_elasticity_R9": 0.049,
    "l_cheek_elasticity_Q0": 41.8,
    "l_cheek_elasticity_Q1": 0.5618,
    "l_cheek_elasticity_Q2": 0.4466,
    "l_cheek_elasticity_Q3": 0.1152,
    "l_cheek_pore": 629
  }
}
```

5. 데이터 구축 절차

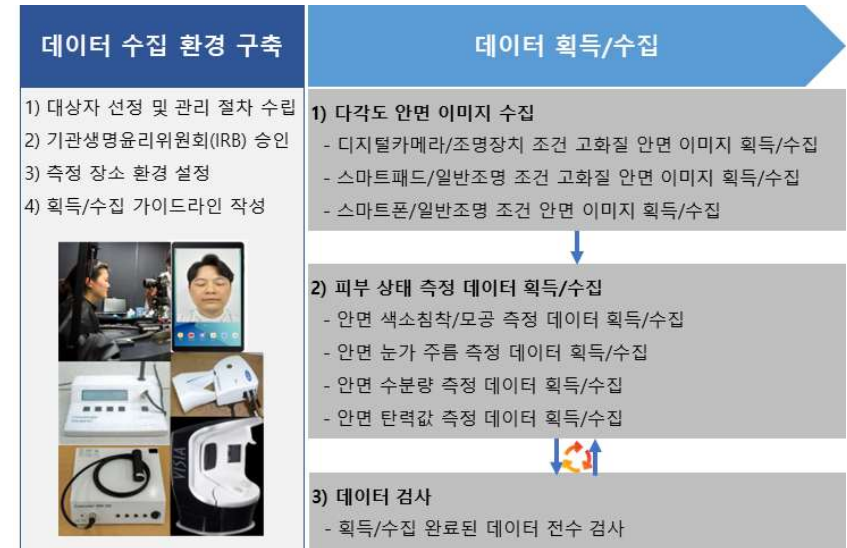
5.1 데이터 구축 절차

○ 데이터 구축 절차

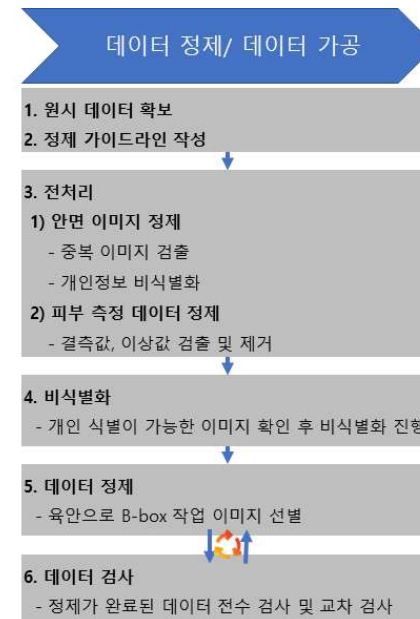


○ 공정별 데이터 구축 절차

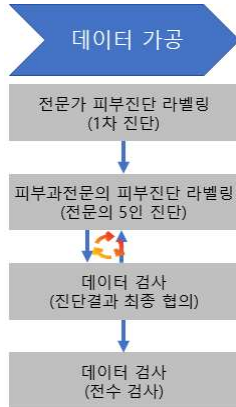
○ 획득/수집 데이터 구축 절차



○ 정제 데이터 구축 절차



○ 가공(라벨링) 데이터 구축 절차



5.2 데이터 구축 기준

○ 획득/수집 기준

구분	기준
법·제도·보안 준수	<ul style="list-style-type: none"> 안면 이미지 및 피부 상태 측정 데이터 획득/수집 활동(시험)은 인간대상 연구에 해당하므로 『생명윤리 및 안전에 관한 법률(생명윤리법)』에 근거하여 수행함 시험계획서 및 관련 문서에 대하여 시험 시작 전에 기관생명윤리위원회(IRB)의 승인을 받음. 최종 승인 이후에 시험을 개시할 수 있으며 최종 승인받은 시험계획서에 따라 시험을 진행할 수 있음 데이터 권리 획득 시 발생 가능한 이슈를 사전에 파악하고 대비 개인정보 보호와 저작권 관련 법률 및 규정을 엄격히 준수
데이터 다양성 확보	<ul style="list-style-type: none"> 데이터의 다양성과 편향 방지 데이터 획득 시 성별, 연령대의 분포를 고려하여 획득
데이터 획득/수집 시 품질고려	<ul style="list-style-type: none"> 품질관리 가이드라인 .0의 데이터 획득/수집 품질기준 및 개인정보 가이드라인 준수
데이터 획득/수집 환경 구성	<ul style="list-style-type: none"> 안면 이미지 획득/수집 <ul style="list-style-type: none"> 고해상도 카메라를 사용하여 다각도로 촬영된 안면 이미지 데이터 수집 조명장치가 설치된 측정실에서 안면 거치대에 얼굴을 고정하고 고성능 디지털카메라로 획득/수집한 다각도 고품질 안면 이미지 일반 천정조명 대기실에서 스마트패드와 스마트폰 카메라로 획득/수집한 다각도 안면 이미지 피부 상태 측정 데이터 획득/수집 <ul style="list-style-type: none"> 4종 이상의 정밀 피부측정 장비를 활용하여 피부 상태 측정 데이터 수집 항온항습장비가 설치된 측정실에서 정밀 피부측정 장비를 이용하여 획득/수집
학습모델 구축목적의 적합성	<ul style="list-style-type: none"> 가장 대중적인 CNN 모델인 ResNet을 사용하여, 학습 및 검증을 진행함으로써 사용자들에게 성능에 대한 Baseline을 제공함 작업 마다 서로 다른 영역의 이미지가 모델의 입력으로 들어가기 때문에 총 11개의 ResNet 모델을 입력 이미지의 종류에 따라 학습하여 모델을 구축함

○ 정제 기준


기준	정제 기준
촬영수단	<ul style="list-style-type: none"> 촬영 수단의 제한
객체의 크기비율	<ul style="list-style-type: none"> 촬영 대상이 이미지 내에서 차지하는 적정 크기 또는 크기 제한
촬영대상제한	<ul style="list-style-type: none"> 이미지 내에서 촬영 대상 객체 외 다른 것들이 포함 가능 여부 한 이미지 내에 포함될 수 있는 객체의 개수 제한(최소, 최대 등)
이미지 비율	<ul style="list-style-type: none"> 촬영 시 가로, 세로 방향 4:3, 16:9 등 가로세로 비율
화질 및 필터	<ul style="list-style-type: none"> 해상도 제한 여부(최소 해상도, 최대 해상도 등) 필터가 적용된 이미지 가능 여부
잘못된 촬영 허용수준	<ul style="list-style-type: none"> 촬영 대상의 초점 안맞음 허용 여부 아웃포커싱 사진 허용 여부 이미지 흔들림 허용 여부 촬영 대상의 잘림 또는 가려짐 허용 여부 그 밖의 사유로 촬영 대상이 잘 보이지 않거나 흐릿함 허용 여부 기울어진 사진 허용 여부
개인정보처리	<ul style="list-style-type: none"> 개인정보보호법 위배 여부 눈가림 처리된 얼굴 사진에 대한 제3자 제공 동의 획득
저작권	<ul style="list-style-type: none"> 저작권 침해 가능성 여부

○ 정제 예시



1. 이마

1-1. 양옆: 양옆 헤어라인 기준으로 안쪽으로 선택
1-2. 상단: 헤어라인 아래를 기준으로 선택
1-3. 하단: 양쪽 눈썹 기준 위를 기준으로 선택



2. 미간

2-1. 양옆: 양쪽 눈썹 안쪽을 기준으로 선택
2-2. 상단: 안쪽 눈썹 기준 눈썹 크기의 1/2 위쪽 기준으로 선택
2-3. 하단: 눈가림 비식별화 영역 위쪽 기준으로 선택



2. 미간

2-1. 양옆: 양쪽 눈썹 안쪽을 기준으로 선택
2-2. 상단: 안쪽 눈썹 기준 눈썹 크기의 1/2 위쪽 기준으로 선택
2-3. 하단: 눈가림 비식별화 영역 위쪽 기준으로 선택

3. 양쪽 눈가

- 3-1. 양옆: 눈가림 비식별화 영역과 얼굴 윤곽선 사이를 선택
- 3-2. 상단: 눈썹 아래를 기준으로 선택
- 3-3. 하단: 눈가림 비식별화 영역 기준으로 상단과 동일한 비율로 하단 위치 선택

4. 볼

- 4-1. 양옆: 코 라인과 얼굴 턱 라인 기준으로 선택
- 4-2. 상단: 눈가림 비식별화 영역 아래 기준으로 선택
- 4-3. 하단: 인중 위 부분 기준으로 선택

5. 입술

- 5-1. 양옆: 양쪽 입꼬리 기준으로 선택
- 5-2. 상단: 입술 윗부분 포함 기준으로 선택
- 5-3. 하단: 입술 아래 부분 포함 기준으로 선택

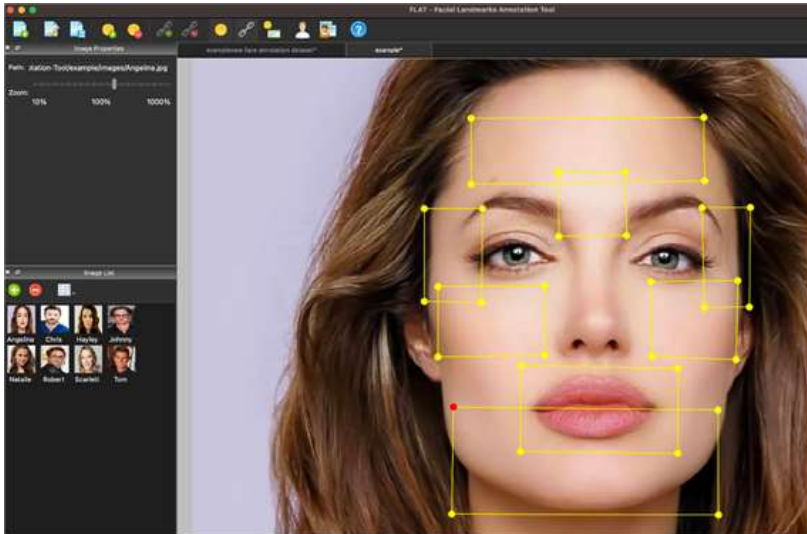
6. 턱 라인

- 6-1. 양옆: 상단과 하단 평행선 기준으로 얼굴 턱 선과 만나는 최대 영역까지 선택
- 6-2. 상단: 입술 가운데 라인을 기준으로 선택
- 6-3. 하단: 턱 하단을 기준으로 선택

○ 가공(라벨링) 기준

구분	가공(라벨링)방법	비고
공통	<ul style="list-style-type: none"> 안면 이미지 내 얼굴 영역 가공을 위한 라벨링(Landmark Object Detection) 가공된 얼굴 영역 데이터는 최소 8만 건 이상 수집 6인 이상의 전문가(피부과 전문의 5인 포함)가 여드름 병변(면포, 구진, 농포, 결절) 위치 및 종류, 주름, 색소침착, 모공, 건조도, 턱선 처짐을 진단한 결과로 라벨링 수행 피부상태 측정 결과 데이터 얼굴 영역별 라벨링 수행 자동 라벨링 저작 도구 사용 	
BBox	<ul style="list-style-type: none"> 고화질 안면 이미지 원천 데이터에서 각 영역별 Bounding Box 라벨링 작업 각 영역별 정밀 장비에서 측정된 피부 상태 측정 데이터(수분, 모공, 색소침착, 탄력)를 연결함 전문가가 이마의 주름 등급, 입술 건조도, 볼 색소 침착 정도, 메모 등을 작성하여 라벨링 값을 저장함 	

○ 가공(라벨링) 예시



6. 데이터 저장 구조 및 파일 명명 규칙

○ 디렉토리 구조

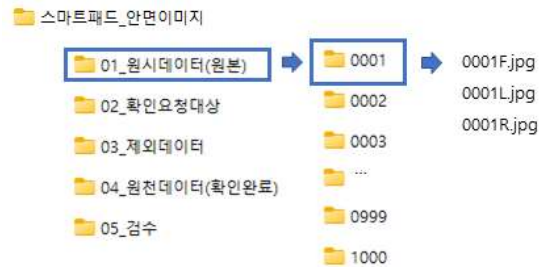


○ 파일 명명 규칙

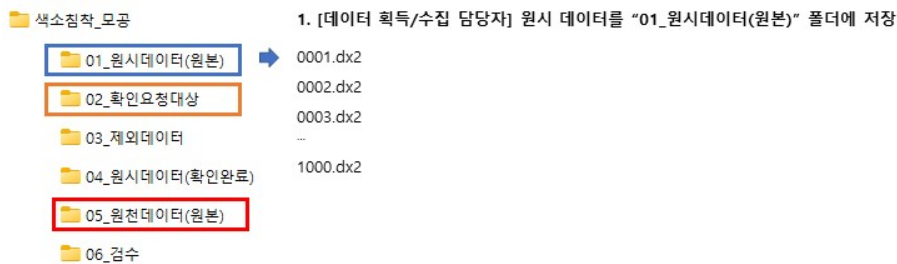
1) 디지털카메라 안면이미지 저장 방법



2) 스마트패드 안면이미지 저장 방법



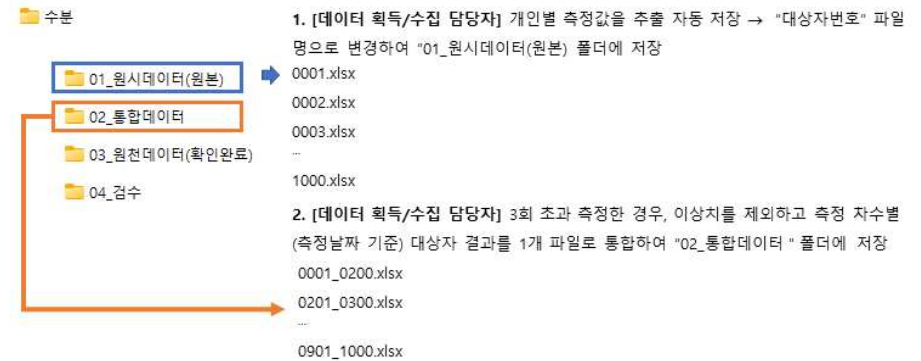
3) 색소침착, 모공 데이터 저장 방법



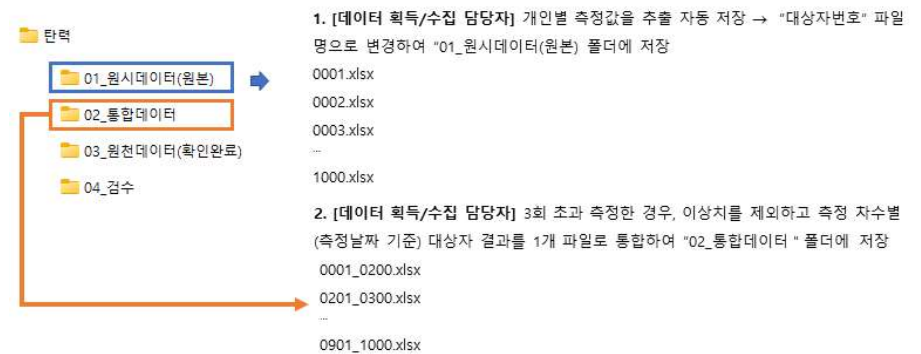
4) 눈가 주름 데이터 저장 방법



5) 수분 측정 데이터 저장 방법



6) 탄력 측정 데이터 저장 방법



제2장 데이터 구축 및 활용

1. 데이터 구축

1.1 데이터 구축과정 및 결과

구축과정	세부 프로세스	내용	결과
데이터 수집/획득	<ul style="list-style-type: none"> 다각도 안면 이미지 수집 ↓ 피부상태 측정 결과 데이터 수집 ↓ 데이터 검사 (전수 검사) 	<ul style="list-style-type: none"> 데이터 수집 환경 구축 획득/수집 가이드라인 작성 다각도 안면 이미지 획득/수집 피부 상태 측정 데이터 획득/수집 데이터 검사 	<ul style="list-style-type: none"> 최종 원시데이터 안면 이미지 13,936건 획득 최종 원시데이터 피부상태 측정데이터 84,688건 획득 목표대비 10% 초과 수집/획득을 통해 정제 시 탈락 데이터 대비
데이터 정제	<ul style="list-style-type: none"> 안면 이미지 정제 (비식별화/자동화) ↓ 얼굴 영역 라벨링 (자동화) ↓ 데이터 검사 (정제 데이터 검사) 	<ul style="list-style-type: none"> 원시데이터 확보 정제 가이드라인 작성 전처리 비식별화 데이터 정제 데이터 검사 	<ul style="list-style-type: none"> 원천데이터 안면 이미지 13,936건 구축 피부상태 측정데이터 84,688건 목표대비 5% 초과 구축으로 가공 시 탈락 데이터 대비
데이터 가공 (라벨링)	<ul style="list-style-type: none"> 전문가 피부진단 라벨링 (1차 진단) ↓ 피부과전문가 피부진단 라벨링 (전문가 5인 진단) ↓ 데이터 검사 (진단결과 최종 협의) ↓ 데이터 검사 (전수 검사) 	<ul style="list-style-type: none"> 가공 기준 작성 가공 도구 전처리 가공 작업 데이터 검사 	<ul style="list-style-type: none"> 라벨링데이터 125,424건 구축 검수 시 오류 수정으로 데이터 모두 활용
AI 모델 학습	<ul style="list-style-type: none"> AI 모델 개발 ↓ AI 모델 학습 ↓ AI 모델 검증 ↓ AI 모델 테스트 	<ul style="list-style-type: none"> 피부상태 등급 분류 성능: 평균 정확도 피부상태 측정 예측 성능: 평균 MAE 	<ul style="list-style-type: none"> 학습 11,148장 테스트 1,394장 검증 1,394장 유효성 지표/결과 <ul style="list-style-type: none"> Top-3 ACC./0.79 MAE./0.28

1.2 데이터 구축 유의사항

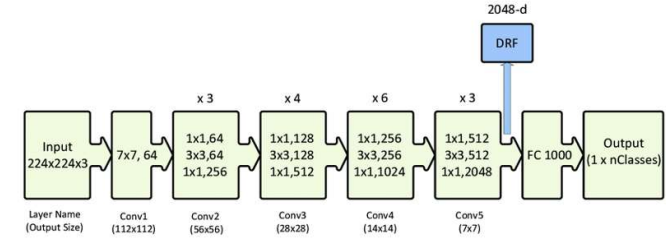
“해당사항 없음”

2. 학습모델

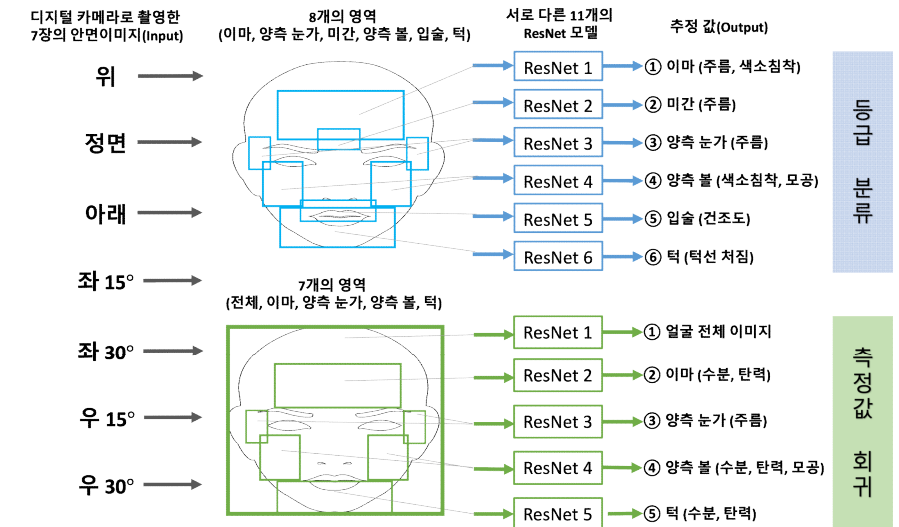
2.1 학습모델 및 알고리즘

○ 여러 개의 ResNet 모델을 이용한 피부상태 진단 및 측정값 예측 모델

<ResNet 50 - architecture>



<전체 프레임워크>



- 주어진 데이터셋에서 안면 전체이미지와 각 영역별 bounding box 정보가 들어있음
- 촬영 장비에 따라 안면 이미지가 여러 각도로 촬영 돼있음
- 각 ResNet 모델마다 영역별 이미지를 입력 이미지로 넣어주고 결과값을 얻어줌
- 등급 분류의 경우, 각 영역에 해당하는 전문가 피부상태 진단 등급을 얻어줌
- 측정값 회귀의 경우, 각 영역에 해당하는 피부 측정 장비값을 얻어주고 해당 값들은 정규화가 된 값들임
- CNN 모델을 활용함하여, 영역별 특징들을 잘 찾아내어 좋은 성능을 이끌어냄

<피부 상태 등급 분류>

유효성 검증 항목					
항목명	피부 상태 등급 분류 성능				
검증 방법	Docker Image 제출				
목적	피부 상태 등급 분류				
지표	정확도				
측정 산식	등급을 맞춘 개수 / 분류할 전체 이미지 수				
도커 이미지	NIA_2023_019.tar, 8.41GB				
실행 파일명	test.py				
유효성 검증 환경					
CPU	Intel(R) Core(TM) i9-10980XE CPU @ 3.00GHz				
Memory	128GB				
GPU	GeForce 3090 * 4				
Storage	1TB				
OS	Ubuntu (20.04)				
유효성 검증 모델 학습 및 검증 조건					
개발 언어	Python 3.8				
프레임워크	CUDA 11.7 Pytorch 2.0.1				
학습 알고리즘	ResNet-50				
학습 조건	batch_size = 1 optimizer = Adam max_epoch = 100 initial lr = 1e-3 power = 0.9 r_decay = lr_poly(initial_lr * ((1 - float(iter) / total_iter) ** (power))) weight_decay = 0.0005				
파일 형식	• 원천 데이터: JPG • 라벨링 데이터: JSON				
전체 구축 데이터 대비 모델에 적용되는 비율	구분	클래스명	비율	수량	
		1 이마 주름	100%	13,936	
		2 이마 색소침착	100%	13,936	
		3 미간 주름	100%	13,935	
		4 눈가 주름	100%	22,566	
		5 볼 색소침착	100%	27,802	
		6 볼 모공	100%	27,802	
		7 입술 건조도	100%	13,936	
		8 턱선 처짐	100%	13,936	
모델 학습 과정별 데이터 분류 및 비율 정보	구분	클래스명	training(80%)	validation(10%)	test(10%)
		1 이마 주름	11,149	1,394	1,393
		2 이마 색소침착	11,149	1,394	1,393
		3 미간 주름	11,149	1,394	1,393
		4 눈가 주름	18,049	2,233	2,284
		5 볼 색소침착	22,244	2,781	2,777
		6 볼 모공	22,244	2,781	2,777
		7 입술 건조도	11,149	1,394	1,393
		8 턱선 처짐	11,149	1,394	1,393
제한사항	• 측면 이미지의 경우, 눈가 영역이 매우 작아 리사이징 오류가 발생하는 등의 문제로 영역의 너비, 높이가 50 pixel 미만인 이미지 제외				

<피부 상태 측정 예측>

유효성 검증 항목					
항목명	피부 상태 측정 예측 성능				
검증 방법	Docker Image 제출				
목적	피부 상태 측정값 예측				
지표	MAE				
측정 산식	(예측한 측정값 - 실제 측정값)의 절대 오차				
도커 이미지	NIA_2023_019.tar, 8.41GB				
실행 파일명	test.py				
유효성 검증 환경					
CPU	Intel(R) Core(TM) i9-10980XE CPU @ 3.00GHz				
Memory	128GB				
GPU	GeForce 3090 * 4				
Storage	1TB				
OS	Ubuntu (20.04)				
유효성 검증 모델 학습 및 검증 조건					
개발 언어	Python 3.8				
프레임워크	CUDA 11.7 Pytorch 2.0.1				
학습 알고리즘	ResNet-50				
학습 조건	batch_size = 1 optimizer = Adam max_epoch = 100 initial lr = 1e-3 power = 0.9 r_decay = lr_poly(init_lr * ((1 - float(iter) / total_iter) ** (power))) weight_decay = 0.0005				
파일 형식	• 원천 데이터: JPG • 라벨링 데이터: JSON				
전체 구축 데이터 대비 모델에 적용되는 비율	구분	클래스명	비율	수량	
	1	이마 수분	100%	13,936	
	2	이마 탄력	100%	13,936	
	3	눈가 주름	100%	20,304	
	4	볼 수분	100%	23,607	
	5	볼 탄력	100%	23,607	
	6	볼 색소침착	100%	23,607	
	7	턱 수분	100%	13,936	
	8	턱 탄력	100%	13,936	
모델 학습 과정별 데이터 분류 및 비율 정보	구분	클래스명	training(80%)	validation(10%)	test(10%)
	1	이마 수분	11,149	1,394	1,393
	2	이마 탄력	11,149	1,394	1,393
	3	눈가 주름	15,787	2,233	2,284
	4	볼 수분	18,049	2,781	2,777
	5	볼 탄력	18,049	2,781	2,777
	6	볼 색소침착	18,049	2,781	2,777
	7	턱 수분	11,149	1,394	1,393
	8	턱 탄력	11,149	1,394	1,393
제한사항	9	정면 색소침착	11,149	1,394	1,393
	• 측면 이미지의 경우, 눈가 영역이 매우 작아 리사이징 오류가 발생하는 등의 문제로 영역의 너비, 높이가 50 pixel 미만인 이미지 제외				

2.2 모델 학습 준비 및 유효성 검증 결과

○ 도커 환경 준비 및 실행

- OS는 Ubuntu 20.04를 사용하였고 모델은 ResNet-50을 사용하며, 고성능의 사양이 요구되지는 않음
- GPU당 최소 1GB의 메모리가 필요함
- System RAM은 최소 6GB의 메모리가 필요함

○ 도커 이미지 불러오기

- 도커 파일 안에 Dataset과 모델 Checkpoint가 모두 포함되어있어 따로 다운 받을 필요 X

\$ docker load -i NIA_2023_019.tar

○ 도커 컨테이너 실행

- 피부 상태 등급 분류

\$ docker run -it -v ./NIA/test nia2 python /NIA/test.py --mode class

- 피부 상태 측정 예측

\$ docker run -it --gpus '"device=0"' -v ./NIA/test nia2 python /NIA/test.py --mode regression

- ✓ --mode를 통해 분류 및 예측 모드를 선택할 수 있음
- ✓ -v 뒤에 본인이 log 파일을 저장하고 싶은 위치를 선택함
- ✓ 위와 같이 할 시, 현재 위치에 class_test.txt 혹은 regression_test.txt 파일이 저장됨
- ✓ 또한, volume을 연결하여, 컨테이너 내에서 생성된 txt파일을 로컬로 가져올 수 있게 해줌
- ✓ --gpu를 통해 원하는 gpu를 선택하여 활용 가능함

○ 컨테이너 실행 모습

```
(nia) jseung@n1pk390kx:~$
$ docker run -it -v ./NIA/test nia2 python /NIA/test.py --mode regression
2023-12-02 00:10:12,284 Set Dec 2 00:18:12 KST 2023

2023-12-02 00:10:12,284 total 116
drwxr-xr-x 0 root root 4096 Dec 2 00:17 .
drwxr-xr-x 1 root root 4096 Dec 2 00:17 ..
drwxr-xr-x 1 root root 19 Dec 1 22:47 .env
drwxr-xr-x 0 root root 4096 Dec 2 00:17 .git
drwxr-xr-x 1 root root 137 Nov 10 10:40 .gitignore
drwxr-xr-x 2 root root 4096 Dec 2 00:17 .vscode
drwxr-xr-x 1 root root 338 Nov 15 15:30 Dockerfile
drwxr-xr-x 2 root root 4096 Dec 2 00:17 pycache
drwxr-xr-x 4 root root 4096 Dec 2 00:17 checkpoint
drwxr-xr-x 1 root root 3639 Nov 30 10:39 data_loader.py
drwxr-xr-x 4 root root 4096 Dec 2 00:17 dataset
drwxr-xr-x 1 root root 3582 Nov 30 17:30 logger.py
drwxr-xr-x 1 root root 6335 Nov 29 12:05 nia.py
drwxr-xr-x 1 root root 21631 Dec 1 17:58 model.py
drwxr-xr-x 1 root root 53 Dec 1 23:44 requirements.txt
drwxr-xr-x 21 3882 1802 4096 Dec 2 00:17 test
drwxr-xr-x 12 root root 4714 Nov 30 15:13 test.py
drwxr-xr-x 1 root root 11276 Dec 1 23:54 test_model.py

2023-12-02 00:10:12,284 Command Line: /NIA/test.py --mode regression
Downloading: "https://download.pytorch.org/models/resnet50-11ad3fa6.pt" to /root/.cache/torch/hub/checkpoints/resnet50-11ad3fa6.pt
100%
/usr/local/lib/python3.8/site-packages/torch/_utils.py:412: UserWarning: Initializing zero-element tensors is a no-op
warnings.warn("Initializing zero-element tensors is a no-op")
2023-12-02 00:10:12,295 Resuming.....checkpoint/regression/100/v1.3
60%
```

- 먼저, 데이터셋을 불러옴 (약 3분 소요)

```
2023-12-02 00:26:35,348 Sub 0786 Equ 03 Angle_F_Area_0(count)====> Pred: 0.530 / Gt: 0.214 ==> MAE: 0.325
2023-12-02 00:26:35,348 Sub 0786 Equ 03 Angle_F_Area_0(count)====> Pred: 0.530 Gt: 0.214
2023-12-02 00:26:35,418 Sub 0786 Equ 03 Angle_F_Area_1(moisture)====> Pred: 0.552 / Gt: 0.633 ==> MAE: 0.081
2023-12-02 00:26:35,418 Sub 0786 Equ 03 Angle_F_Area_1(moisture)====> Pred: 0.552 Gt: 0.633
2023-12-02 00:26:35,419 Sub 0786 Equ 03 Angle_F_Area_1(elasticity)====> Pred: 0.547 / Gt: 0.584 ==> MAE: 0.137
2023-12-02 00:26:35,419 Sub 0786 Equ 03 Angle_F_Area_1(elasticity)====> Pred: 0.547 Gt: 0.584
2023-12-02 00:26:35,475 Sub 0786 Equ 03 Angle_F_Area_3(wrinkle)====> Pred: 0.211 / Gt: 0.100 ==> MAE: 0.031
2023-12-02 00:26:35,475 Sub 0786 Equ 03 Angle_F_Area_3(wrinkle)====> Pred: 0.211 Gt: 0.100
2023-12-02 00:26:35,575 Sub 0786 Equ 03 Angle_F_Area_5(moisture)====> Pred: 0.406 / Gt: 0.567 ==> MAE: 0.151
2023-12-02 00:26:35,575 Sub 0786 Equ 03 Angle_F_Area_5(moisture)====> Pred: 0.406 Gt: 0.567
2023-12-02 00:26:35,576 Sub 0786 Equ 03 Angle_F_Area_5(elasticity)====> Pred: 0.568 / Gt: 0.745 ==> MAE: 0.175
2023-12-02 00:26:35,577 Sub 0786 Equ 03 Angle_F_Area_5(elasticity)====> Pred: 0.568 Gt: 0.745
2023-12-02 00:26:35,577 Sub 0786 Equ 03 Angle_F_Area_5(pore)====> Pred: 0.224 / Gt: 0.130 ==> MAE: 0.094
2023-12-02 00:26:35,577 Sub 0786 Equ 03 Angle_F_Area_5(pore)====> Pred: 0.224 Gt: 0.130
2023-12-02 00:26:35,626 Sub 0786 Equ 03 Angle_F_Area_6(moisture)====> Pred: 0.240 / Gt: 0.573 ==> MAE: 0.324
2023-12-02 00:26:35,626 Sub 0786 Equ 03 Angle_F_Area_6(moisture)====> Pred: 0.240 Gt: 0.573
2023-12-02 00:26:35,627 Sub 0786 Equ 03 Angle_F_Area_6(elasticity)====> Pred: 0.734 / Gt: 0.620 ==> MAE: 0.386
2023-12-02 00:26:35,627 Sub 0786 Equ 03 Angle_F_Area_6(elasticity)====> Pred: 0.734 Gt: 0.620
2023-12-02 00:26:35,627 Sub 0786 Equ 03 Angle_F_Area_6(pore)====> Pred: 0.121 / Gt: 0.137 ==> MAE: 0.017
2023-12-02 00:26:35,627 Sub 0786 Equ 03 Angle_F_Area_6(pore)====> Pred: 0.121 Gt: 0.137
2023-12-02 00:26:35,629 Sub 0786 Equ 03 Angle_F_Area_8(moisture)====> Pred: 0.533 / Gt: 0.703 ==> MAE: 0.170
2023-12-02 00:26:35,629 Sub 0786 Equ 03 Angle_F_Area_8(moisture)====> Pred: 0.533 Gt: 0.703
2023-12-02 00:26:35,669 Sub 0786 Equ 03 Angle_F_Area_8(elasticity)====> Pred: 0.848 / Gt: 0.515 ==> MAE: 0.133
2023-12-02 00:26:35,669 Sub 0786 Equ 03 Angle_F_Area_8(elasticity)====> Pred: 0.848 Gt: 0.515
2023-12-02 00:26:35,709 count: 0.2225 // moisture: 0.1776 // wrinkle: 0.0059 // elasticity: 0.2444 // pore: 0.2961
```

- 각 피실험자의 이미지에 대한 테스트가 이뤄지며, 피실험자의 번호, 장비, 영역, 진단 클래스의 예측 및 정답, 그리고 MAE가 출력 (약 20분 소요)

```
2023-12-06 16:59:34,674 Sub_0120 Equ_01_Angle_Ft_Area_6(moisture)====> Pred: 0.643 / Gt: 0.488 ==> MAE: 0.163
2023-12-06 16:59:34,675 Sub_0120 Equ_01_Angle_Ft_Area_6(elasticity)====> Pred: 0.662 / Gt: 0.760 ==> MAE: 0.187
2023-12-06 16:59:34,675 count: 0.2188 // moisture: 0.2416 // wrinkle: 0.183 // elasticity: 0.266 // pore: 0.2792
2023-12-06 16:59:34,675 [1392 / 1393] Total Average MAE ==> 0.222
2023-12-06 16:59:34,912 Finish!
```

- Finish! 라는 문구 출력과 함께 최종 결과값이 출력

3. 서비스 활용

3.1 서비스 활용 예시 및 시나리오

AI 데이터셋 활용 분야	활용 방법
맞춤형 화장품 추천	개인의 피부 상태에 맞는 화장품을 제공하여 고객들의 만족도와 구매 경험을 향상
피부 임상시험	피부의 수분량, 모공, 주름, 탄력, 색소침착 정도 등 다양한 피부 특성을 정량화하고 평가
피부 질환 진단	피부 측정 데이터를 활용하여 피부 질환을 정확하게 진단하고 적절한 치료 계획 수립
피부 AI 연구	피부 상태 측정 AI 모델의 공개와 데이터 시각화를 통해 국내 인공지능 기술혁신을 지원
피부 제품 개발	피부 측정 데이터를 활용하여 피부 연구를 수행하고 새로운 제품 개발에 활용

별첨 1. 데이터 구축 현황표

- 시트 1 : 공모서(RFP)

[019] 한국인 피부상태 측정 데이터		
공모서(RFP)		
1. 데이터 개요		<ul style="list-style-type: none"> ○ 한국인의 얼굴 부위별 미용 목적의 피부 상태 정밀 측정 결과에 대한 이미지 데이터 ○ 인공지능 학습용 데이터 구축량 <ul style="list-style-type: none"> - 고해상도 카메라로 촬영된 고품질 안면 이미지 1만 장 이상과 정밀 피부측정 장비로부터 획득된 피부 상태 측정 데이터 6만 건 이상 수집 - 1만 장의 안면 이미지에 얼굴 영역 5만 건 이상, 전문가가 진단한 피부상태가 라벨링 된 데이터 ○ AI 임무(task) <ul style="list-style-type: none"> - 다양한 각도의 안면 이미지를 활용하여 피부 상태 측정 시 활용되는 각 얼굴 부위 검출 및 개인별 정밀 피부 상태 예측 ○ 활용 분야(예시) <ul style="list-style-type: none"> - 고객의 피부 유형에 따른 맞춤형 화장품 추천 및 개발 등
2. 데이터 구성	원천 데이터	<ul style="list-style-type: none"> • 고해상도 카메라를 활용하여 다각도로 촬영된 고품질 안면 이미지 1만 장 이상 ※ 촬영 인원 : 1천 명 이상 • 3종 이상의 정밀 피부측정 장비로부터 획득된 피부 상태 측정 결과 데이터 6만 건 이상 * (예시) 수분량, 모공, 주름, 탄력, 색소침착 정도 등에 대한 측정 데이터
	라벨링 데이터	<ul style="list-style-type: none"> • 1만 장의 안면 이미지에 얼굴 영역 5만 건 이상이 라벨링된 데이터 * (예시) 눈가, 볼, 미간, 이마 등 • 1만 장의 안면 이미지에 전문가가 진단한 피부 상태가 라벨링된 데이터 • 기타 명시되지 않은 사항은 데이터 구축 목적을 달성할 수 있도록 라벨링 항목을 구체화하여 제시
	메타 데이터	<ul style="list-style-type: none"> • 다음의 필수 정보를 포함하여 데이터 명세에 필요한 메타데이터 구축방안 제시 - 피부측정 장비 정보 등 - 성별, 연령대 등 개인 특성 데이터 • 기타 명시되지 않은 사항은 데이터 구축 목적을 달성할 수 있도록 메타데이터를 구체화하여 제시
3. 데이터 구축방법	데이터 수집방법	<ul style="list-style-type: none"> • 다음 사항을 포함한 원천데이터 수집 방안 제시 - 고해상도 카메라로 촬영된 안면 이미지 데이터 수집 방안 - 실제 환경에서 3종 이상의 정밀 피부측정 장비를 활용한 데이터 수집 방안 • 고품질의 이미지를 얻기 위한 하드웨어를 이용하여 안면 이미지 데이터에 대한 적정 가이드라인 설계 및 근거 제시 • 그 이외의 데이터 수집을 위한 절차, 장소 및 도구(HW-SW), 조직, 기준 등이 포함된 수집 계획 제시
	데이터 가공 방법	<ul style="list-style-type: none"> • 다음 사항을 포함하여 데이터 구성에 적합한 라벨링 방법 제시 - 카메라로 촬영된 고품질 안면 이미지 내 얼굴 영역 가공을 위한 라벨링(Segmentation, Landmark, Bounding Box 등) 방법 - 전문가가 진단한 피부 미용 관련 상태에 대한 구체적인 라벨링 방법 • 그 이외의 고품질의 데이터 구축을 위한 최적의 라벨링 방법 제시

- 시트 3 : 데이터 구축 현황

구축 데이터 개요

구축 목표항
(AI역의 개방 기준)

o 다양한 자료의 한국인 얼굴 이미지와 피부 측정 데이터까지 구축 및 제공

1. 고해상도 다각도 인면이미지 13,936 장
2. 피부상태 측정 데이터 84,688 건
3. 제1데이터 6,432 건
4. 전문가 라벨링 데이터 125,424 건

원시 데이터

설명	유형	형식	수량	단위	획득방법	AI역로 제출 여부
교활질 인면 이미지	교활질 카메라	.jpg	7,700	장	직접촬영	X
	스마트패드	.jpg	3,300	장	직접촬영	X
	스마트폰	.jpg	3,300	장	직접촬영	X
침양알성 인면 데이터	측정 데이터 파일(수분)	* .xlsx	4,000	건	직접촬영	X
	측정 데이터 파일(탄력)	* .xlsx	76,000	건	직접촬영	X
	측정 데이터 파일(주름)	* .omc	4,000	건	직접촬영	X
	측정 데이터 파일(색소침착/홍운)	* .dkt	1,000	건	직접촬영	X

* 제출여부와 상관없이 원시데이터 보관은 5년 필수

Training - 최종 학습용 데이터 (원천 데이터 + 라벨링 데이터)

원천 데이터						라벨링 데이터				
설명	분류(1단계)	분류(2단계)	형식	수량	단위	라벨링 유형	객체수	형식	파일 수량 / 비율	
한국인 피부상태 측정 데이터	교활질 카메라	-	-	.jpg	6,006	장	전문화 라벨링	11	jeon	54,054 1 : 9
	스마트패드	-	-	.jpg	2,574	장	전문가 라벨링	11	jeon	23,166 1 : 9
	스마트폰	-	-	.jpg	2,574	장	전문가 라벨링	11	jeon	23,166 1 : 9

Validation - 최종 학습용 데이터 (원천 데이터 + 라벨링 데이터)

원천 데이터						라벨링 데이터				
설명	분류(1단계)	분류(2단계)	형식	수량	단위	라벨링 유형	객체수	형식	파일 수량 / 비율	
한국인 피부상태 측정 데이터	교활질 카메라	-	-	.jpg	749	장	전문화 라벨링	11	jeon	6,741 1 : 9
	스마트패드	-	-	.jpg	321	장	전문가 라벨링	11	jeon	2,889 1 : 9
	스마트폰	-	-	.jpg	321	장	전문가 라벨링	11	jeon	2,889 1 : 9

Test - 최종 학습용 데이터 (원천 데이터 + 라벨링 데이터)

원천 데이터						라벨링 데이터				
설명	분류(1단계)	분류(2단계)	분류(3단계)	형식	수량	단위	라벨링 유형	객체수	형식	파일 수량 / 비율
한국인 피부상태 측정 데이터	교활질 카메라	-	-	.jpg	749	장	전문화 라벨링	11	jeon	6,741 1 : 9
	스마트패드	-	-	.jpg	321	장	전문가 라벨링	11	jeon	2,889 1 : 9
	스마트폰	-	-	.jpg	321	장	전문가 라벨링	11	jeon	2,889 1 : 9

[illegible]

별첨 2. 개인정보 수집·이용 동의서

개인정보 수집 · 이용 동의서

한국지능정보사회진흥원에서는 개인정보를 중요시하며, 「개인정보 보호법」 제15조, 제22조에 따라 아래와 같이 동의를 얻고자 합니다.

아래 사항에 대해 충분히 읽어 보신 후, 동의 여부를 체크 · 서명하여 주시기 바랍니다.

개인정보 수집 및 이용에 대한 동의	
<input type="radio"/> 개인정보 수집 · 이용 목적	
<input checked="" type="checkbox"/> AI 허브에 공개된 인공지능 학습용 데이터에 대한 사용자 이용문의 대응	
<input type="radio"/> 수집하는 개인정보 항목	
<input checked="" type="checkbox"/> 필수적 정보	
<input type="checkbox"/> - 개인식별정보(휴대폰 번호, 이메일)	
<input type="radio"/> 개인 정보 보유 · 이용 기간 : 준영구	
<div>※ 유의 사항 : 귀하는 상기 동의를 거부할 수 있습니다. 그러나 이에 대한 동의를 하지 않을 경우, AI 데이터 운영관리 및 AI 허브 고객 문의 대응에 제한을 받으실 수 있습니다.</div> <div><input checked="" type="checkbox"/> 동의함 <input type="checkbox"/> 동의하지 않음</div>	

본인은 본 "개인정보 수집 및 이용" 내용을 읽고 명확히 이해하였으며 이에 동의합니다.

2024년 3월 26일

한국지능정보사회진흥원 귀증 본 인 : 나준채 (서명 또는 인)

별첨 2. 개인정보 수집·이용 동의서

개인정보 수집 · 이용 동의서

한국지능정보사회진흥원에서는 개인정보를 중요시하며, 「개인정보 보호법」 제15조, 제22조에 따라 아래와

같이 동의를 얻고자 합니다.

아래 사항에 대해 충분히 읽어 보신 후, 동의 여부를 체크 · 서명하여 주시기 바랍니다.

개인정보 수집 및 이용에 대한 동의	
<input type="radio"/> 개인정보 수집 · 이용 목적	
<input checked="" type="checkbox"/> AI 허브에 공개된 인공지능 학습용 데이터에 대한 사용자 이용문의 대응	
<input type="radio"/> 수집하는 개인정보 항목	
<input checked="" type="checkbox"/> 필수적 정보	
<input type="checkbox"/> - 개인식별정보(휴대폰 번호, 이메일)	
<input type="radio"/> 개인 정보 보유 · 이용 기간 : 준영구	
<div>※ 유의 사항 : 귀하는 상기 동의를 거부할 수 있습니다. 그러나 이에 대한 동의를 하지 않을 경우, AI 데이터 운영관리 및 AI 허브 고객 문의 대응에 제한을 받으실 수 있습니다.</div> <div><input checked="" type="checkbox"/> 동의함 <input type="checkbox"/> 동의하지 않음</div>	

본인은 본 "개인정보 수집 및 이용" 내용을 읽고 명확히 이해하였으며 이에 동의합니다.

2024년 3월 26일

한국지능정보사회진흥원 귀증 본 인 : 이은영 (서명 또는 인)

별첨 2. 개인정보 수집·이용 동의서

개인정보 수집 · 이용 동의서

한국지능정보사회진흥원에서는 개인정보를 중요시하며, 「개인정보 보호법」 제15조, 제22조에 따라 아래와 같이 동의를 얻고자 합니다.

아래 사항에 대해 충분히 읽어 보신 후, 동의 여부를 체크·서명하여 주시기 바랍니다.

개인정보 수집 및 이용에 대한 동의	
<input type="radio"/> 개인정보 수집 · 이용 목적	
✓ AI 허브에 공개된 인공지능 학습용 데이터에 대한 사용자 이용문의 대응	
<input type="radio"/> 수집하는 개인정보 항목	
✓ 필수적 정보	
- 개인식별정보(휴대폰 번호, 이메일)	
<input type="radio"/> 개인 정보 보유 · 이용 기간 : 준영구	
<div>※ 유의 사항 : 귀하는 상기 동의를 거부할 수 있습니다. 그러나 이에 대한 동의를 하지 않을 경우, AI 데이터 운영관리 및 AI 허브 고객 문의 대응에 제한을 받으실 수 있습니다.</div>	
<input checked="" type="checkbox"/> 동의함	<input type="checkbox"/> 동의하지 않음

본인은 본 "개인정보 수집 및 이용" 내용을 읽고 명확히 이해하였으며 이에 동의합니다.

2024년 3월 26일

한국지능정보사회진흥원 귀중 본 인 : 이정호 (서명인)

별첨 2. 개인정보 수집·이용 동의서

개인정보 수집 · 이용 동의서

한국지능정보사회진흥원에서는 개인정보를 중요시하며, 「개인정보 보호법」 제15조, 제22조에 따라 아래와 같이 동의를 얻고자 합니다.

아래 사항에 대해 충분히 읽어 보신 후, 동의 여부를 체크·서명하여 주시기 바랍니다.

개인정보 수집 및 이용에 대한 동의	
<input type="radio"/> 개인정보 수집 · 이용 목적	
✓ AI 허브에 공개된 인공지능 학습용 데이터에 대한 사용자 이용문의 대응	
<input type="radio"/> 수집하는 개인정보 항목	
✓ 필수적 정보	
- 개인식별정보(휴대폰 번호, 이메일)	
<input type="radio"/> 개인 정보 보유 · 이용 기간 : 준영구	
<div>※ 유의 사항 : 귀하는 상기 동의를 거부할 수 있습니다. 그러나 이에 대한 동의를 하지 않을 경우, AI 데이터 운영관리 및 AI 허브 고객 문의 대응에 제한을 받으실 수 있습니다.</div>	
<input checked="" type="checkbox"/> 동의함	<input type="checkbox"/> 동의하지 않음

본인은 본 "개인정보 수집 및 이용" 내용을 읽고 명확히 이해하였으며 이에 동의합니다.

2024년 3월 26일

한국지능정보사회진흥원 귀중 본 인 : 양성일 (서명인)