

개요: 한국인 안면이미지 데이터셋이란?

딥러닝 기반의 안면 인식 기술의 발전을 위해서는, 대규모의 정제되고 색인된 안면 이미지 데이터셋은 물론이고, 다양한 포즈, 조도, 표정, 액세서리와 같은 다양한 환경이 고려된 데이터셋이 필수적이다. 본 데이터셋은 이와 같은 필요에 의해 한국과학기술연구원(KIST) 인공지능연구단 주도로 2017년부터 2019년까지 구축된 데이터셋이다. 성별과 연령대가 각기 다른 한국인 1,000명*을 대상으로 얼굴 포즈 20종, 조명 30종, 액세서리 6종, 표정 3종, 해상도 3종을 고려하여 촬영을 통해 구축하였으며, 예시 영상은 그림 1과 같다.

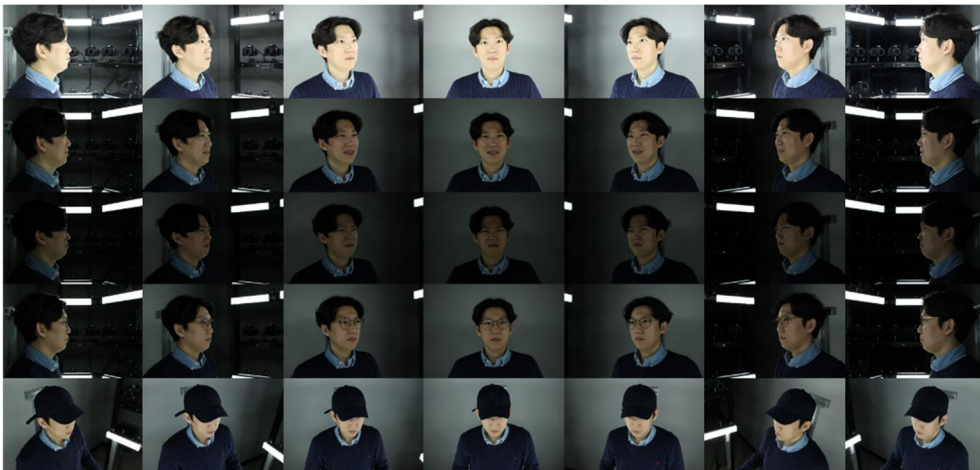


그림 1 한국인 안면 이미지 데이터셋 샘플 영상

*다년도에 걸쳐 구축한 데이터의 특성 상 안면 이미지 제공자에게 받은 데이터 활용 범위와 수량이 연차별로 상이함.

데이터셋의 구성

본 데이터셋은 안면 인식기가 다양한 각도와 조도에서 찍힌 얼굴 사진을 어느 수준까지 잘 식별할 수 있는지에 대한 물음에 답을 할 수 있도록 설계되었다. 데이터 구축 시 임의의 사진에서 정확한 각도와 조도를 알아내어 색인하는 일은 어려운 일이므로, 그림 2와 같이 촬영 조건을 알 수 있는 촬영 장비를 사용하였다.

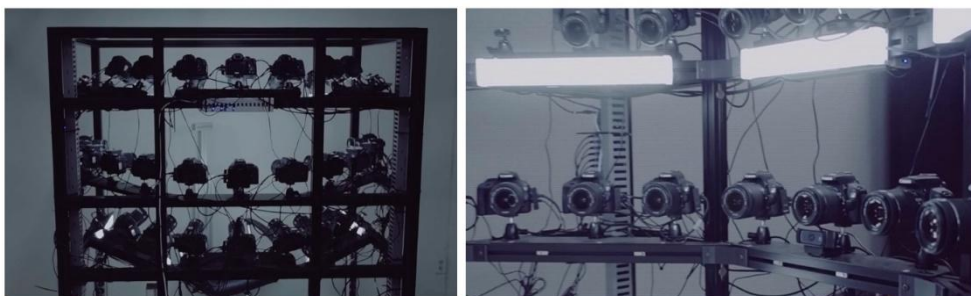


그림 2 한국인 안면 이미지 데이터셋 촬영 장비

촬영 장비에는 촬영 대상자를 중심으로 수직 방향 0도 기준에서 13개의 수평 방향 (± 90 , ± 75 , ± 60 , ± 45 , ± 30 , ± 15 , 0), 수직 상단 30도 기준에서 5개의 수평방향(± 45 , ± 15 , 0), 수직 하단 15도 기준에서 2개의 수평 방향 (± 40)을 고려하여 최소 20개의 각도에 카메라가 설치되어 있으며, 단 한 번의 촬영으로 여러 각도의 영상을 동시에 획득할 수 있도록 동기화 되어 있다.

또한 촬영 장비에는 조도를 제어할 수 있는 조명 제어 시스템이 설치되어 있으며, 조명을 0 Lux에서 1000 Lux까지, 방향을 바꿔가며 30가지 이상의 조합을 프로그래밍할 수 있도록 설계되었다.

최종적으로 표 1과 같이 얼굴 포즈 20종, 조명 30종, 액세서리 6종, 표정 3종, 해상도 3종을 고려하여 촬영 대상자 1명당 32,400 장의 안면 이미지 데이터셋이 구축되었다.

표 1 한국인 안면 이미지 데이터셋 구성 요소 및 상세 설명

구성 요소		상세 설명
안면 이미지 촬영 시스템	반구 형태의 프레임	좌우 90도, 상단 30도 / 하단 15도 커버
	최소 20대의 DSLR 및 Webcam	Canon EOS 100D
	Auto Illuminator	각도별, 조도별(Lux 단위) 조절 가능
안면 이미지 데이터	포즈	좌우 90도, 상단 30도 / 하단 15도 이내
	조명	30 가지 경우
	가림	일반 안경, 뿔테 안경, 선글라스, 모자
	표정	일반, 웃음, 찡그림
	해상도	고화질(864x576), 중화질(346x230), 저화질(173x115)
색인 정보	촬영 조건	폴더와 파일 이름으로 구분
	기타	특징점, 바운딩 박스 정보(특정 촬영 조건 내)

데이터셋의 설계 기준과 분포

본 데이터셋은 특정 성별과 연령대에 편향되지 않게 20대부터 50대의 다양한 사람을 섭외하여 촬영되었다. 다양한 사람을 고르게 섭외하여 촬영하려는 노력에도 불구하고, 2017년과 2018년에 촬영된 초기 데이터셋의 경우 그림 3과 같이 여성과 20대에 편향되어 있었다. 이를 해결하기 위해 2019년에는 남성과 30대 이상의 사람들을 주로 섭외하여, 최종적으로 그림 4와 같은 분포의 데이터셋이 되었다.

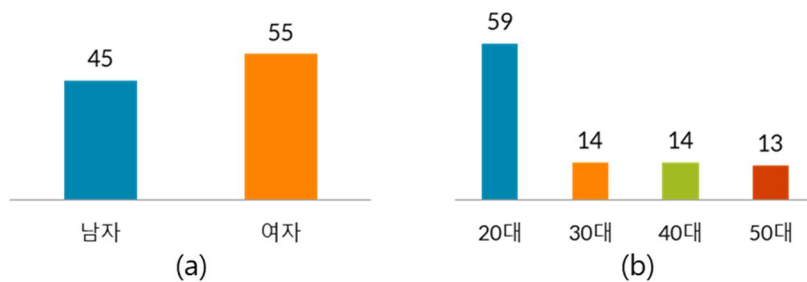


그림 3 한국인 안면 이미지 데이터셋 초기 400명 분포: (a) 성별 (b) 연령별

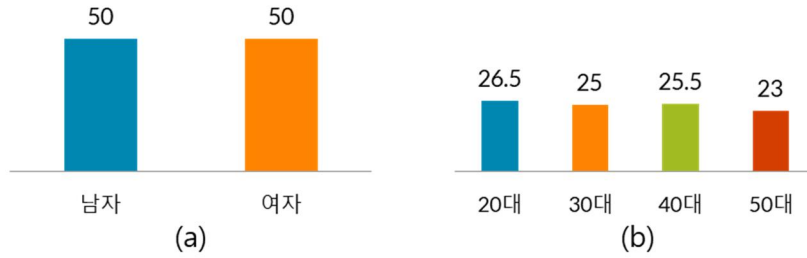


그림 4 한국인 안면 이미지 데이터셋 최종 1000명 분포: (a) 성별 (b) 연령별

본 데이터셋은 안면 인식기가 비제한적인 환경에서 찍힌 사진을 어느 수준까지 잘 식별할 수 있는지를 판단하기 위한 목적으로 처음 설계되었으며, 이를 역으로 활용하면 촬영 조건에 강인한 안면 인식기를 개발할 수 있다. 이것은 최근 활발하게 연구되고 있는 가짜 얼굴 생성 및 얼굴 정면화 기술들을 개발하는데도 똑같이 적용된다. 임의로 촬영된 사진들을 각도와 조도에 따라 구분하는 일은 어려운 일이므로, 촬영 조건을 정확히 알 수 있도록 제어 가능한 촬영 장치를 사용하였고, 상세 촬영 조건은 그림 5와 같다.

방향			Light				Camera (각도)			
		+			방향				방향	
수직	수평	상단	수직	수평	Lux	수직	수평	수직	수평	
		하단			L1	전체		C1	+90°	
		좌측			L2	전체		C2	+75°	
					L3	전체		C3	+60°	
					L4	전체		C4	+45°	
					L5	전체		C5	+30°	
					L6	전체		C6	+15°	
					L7	전체		C7	+0°	
					L8	+30°	전체	C8	-15°	
					L9	+30°	전체	C9	-30°	
					L10	+30°	전체	C10	-45°	
					L11	+30°	전체	C11	-60°	
					L12	-15°	전체	C12	-75°	
					L13	-15°	전체	C13	-90°	
					L14	-15°	전체	C14	+30°	
					L15	-15°	전체	C15	+15°	
					L16	전체	+90°	C16	0°	
					L17	전체	+90°	C17	+30°	
					L18	전체	+90°	C18	-45°	
					L19	전체	+45°	C19	+30°	
					L20	전체	+45°	C20	-15°	
					L21	전체	+45°			
					L22	전체	+0°			
					L23	전체	+0°			
					L24	전체	+0°			
					L25	전체	-45°			
					L26	전체	-45°			
					L27	전체	-45°			
					L28	전체	-90°			
					L29	전체	-90°			
					L30	전체	-90°			

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

Expression	
E01	무표정
E02	활짝 웃음
E03	찡그림

그림 5 한국인 안면 이미지 데이터의 촬영 조건

데이터 구조

본 데이터셋의 저장 구조는 그림 6과 같다. 식별 번호(ID)를 기준으로 본인 사진과 타인의 사진을 구분할 수 있다. 식별 번호에는 어떠한 개인정보도 포함되어 있지 않다.

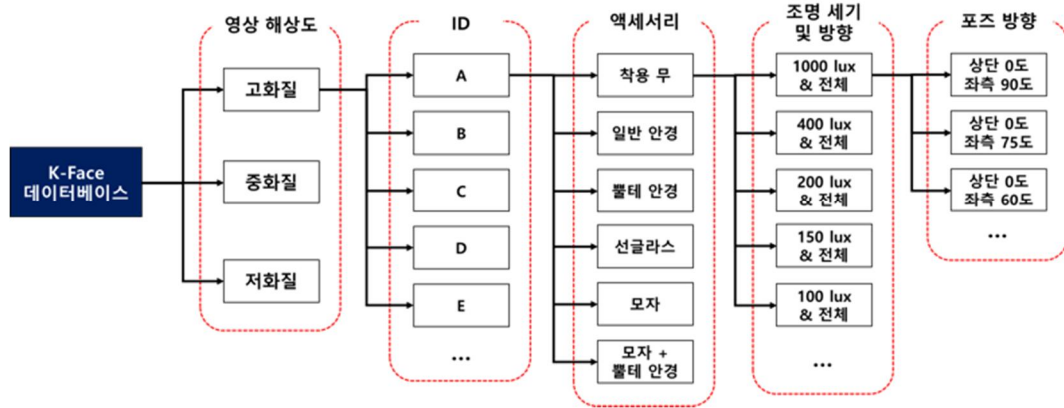


그림 6 한국인 안면 이미지 데이터셋 구조

데이터셋의 활용 범위를 높이고, 데이터셋을 보다 효율적으로 사용할 수 있도록 가림이 없는 특정 조명 조건(L1, L3, L6, L7)의 경우 특징점 정보와 바운딩 박스(Bounding Box) 정보를 색인하였다. 그림 7은 특징점 정보의 예시로 코 끝, 우측 눈 중심, 좌측 눈 중심, 우측 입꼬리, 좌측 입꼬리, 우측 귀, 좌측 귀의 순서로 총 7개의 위치 정보를 색인하였다. 그림 8은 바운딩 박스 정보의 예시로, 얼굴, 우측 눈, 좌측 눈, 코, 입, 우측 귀, 좌측 귀의 순서로 총 7개 영역 정보를 색인하였다. 바운딩 박스의 경우 보이지 않는 부분은 색인하지 않고 생략하였다.



그림 7 한국인 안면 이미지 특징점 예시



그림 8 한국인 안면 이미지 바운딩 박스 예시

데이터 예시

본 데이터셋의 예시 그림은 그림 9와 같다. 기본적으로 폴더 구조와 파일 이름으로 본인과 타인, 촬영 조건들을 구분할 수 있으며, 가림이 없는 특정 조명 조건(L1, L3, L6, L7)의 경우 이미지 파일과 함께 텍스트 파일도 함께 존재하는 것을 확인할 수 있다. 텍스트 파일에는 특징점의 위치 정보와 바운딩 박스의 코너 위치, 폭, 높이 정보가 기록되어 있다. 앞에서 설명한 바와 같이, 그림 9의 C1은 측면 90도에서 촬영된 영상으로 보이지 않는 얼굴 영역의 바운딩 박스 정보가 생략되어 있다.

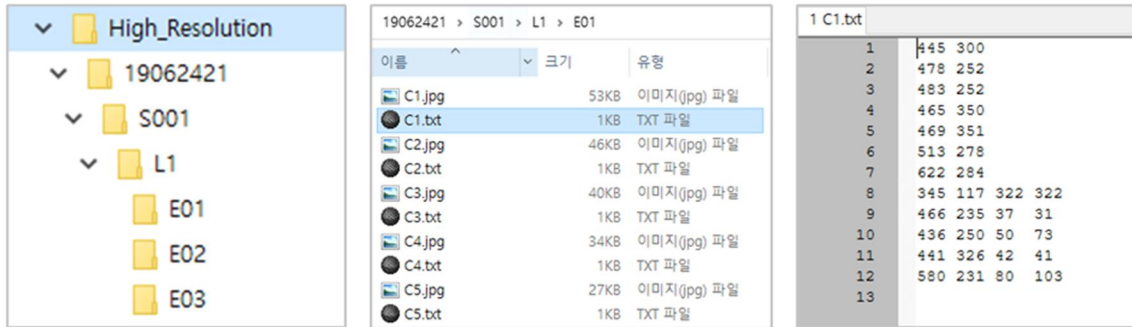


그림 9 한국인 안면 이미지 데이터셋 예시

데이터 구축 과정

본 데이터셋은 2017년에 200명, 2018년에 200명, 2019년에 600명을 섭외하여 구축하였고, 데이터 구축 과정은 그림 10과 같다. 연기자가 촬영장에 방문하면 우선 개인정보 활용에 대한 안내를 받고 동의서를 작성하고 촬영에 임하게 된다. 데이터를 구축한 기관만이 안면 이미지 데이터를 사용하는 것이 아니라, 제 3의 기관도 사용할 수 있도록 해야 하는 사업의 특수한 목적에 따라 개인정보 활용동의서는 해마다 갱신되었다.

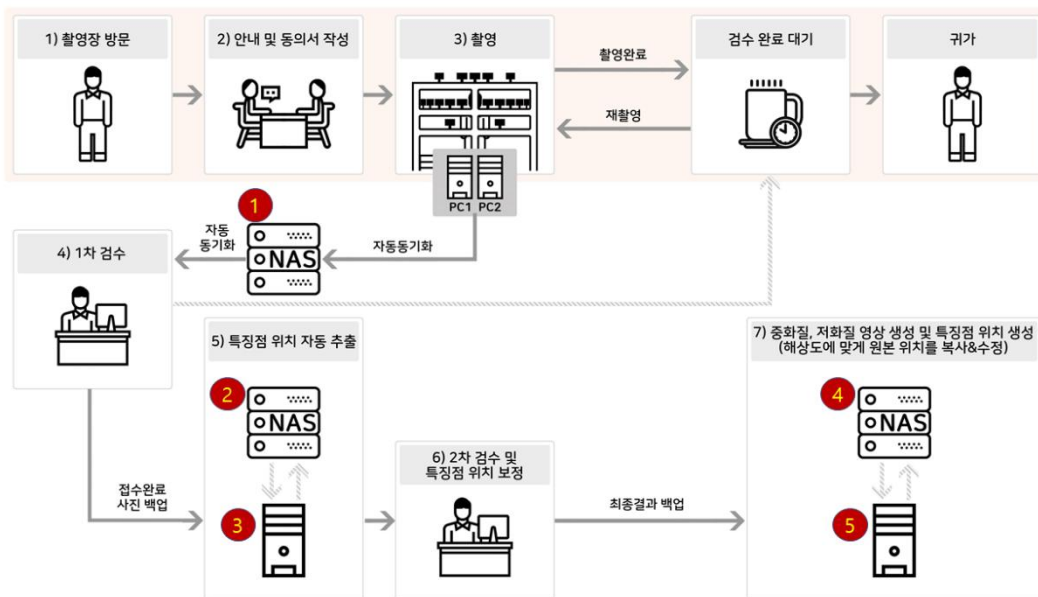


그림 10 데이터 구축 개요

촬영 시간은 짧게는 1시간 30분, 길게는 4시간이 소요되었으며, 2019년의 경우 단기간에 600명의 인원을 촬영해야 했기 때문에, 총 네 대의 촬영 장비를 제작하여 사용하였다.

자동화된 데이터 전송 시스템과 1차 검수 과정을 두어 데이터셋의 오류와 누락을 현장에서 발견하고 조치할 수 있도록 하였으며, 현장에서 발견하지 못한 오류가 2차 검수에서 발견된 경우 재촬영을 원칙으로 하였다.

검수와 품질 확보

본 데이터셋의 품질 확보를 위해 1차, 2차 검수 시스템을 운영하였다. 한 번의 촬영에서 획득하는 이미지 수량이 매우 많고, 촬영에 상당한 시간이 소요되기 때문에 되도록 현장에서 데이터의 오류와 누락을 발견하고 재촬영할 수 있도록 하였다.

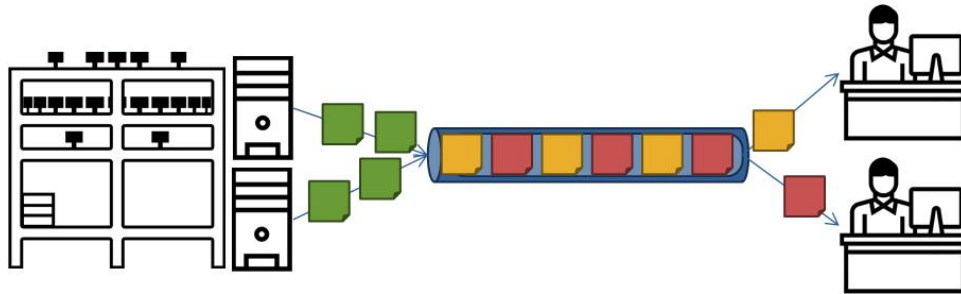


그림 11 데이터셋 현장 검수 시스템 예시

현장에서 발견하지 못한 데이터의 오류는 2차 검수 과정을 통해 발견하고 수정할 수 있도록 하였다. 복구할 수 없는 오류라고 판단되는 경우 재촬영을 원칙으로 하였고, 실제 재촬영 비율은 전체 데이터의 6.7% 수준이었다.

구축된 한국인 안면 이미지 데이터셋의 유효성을 확인하기 위해 교차 유효성 검증 (Cross-Validation)을 활용하였다. 그림 12와 같이 구축된 데이터를 랜덤하게 섞은 후 겹치지 않게 N개의 그룹으로 나누고 하나의 데이터셋(테스트용)을 제외한 나머지 데이터셋을 활용하여 모델을 학습하고 테스트하는 것을 N번 반복하였다.

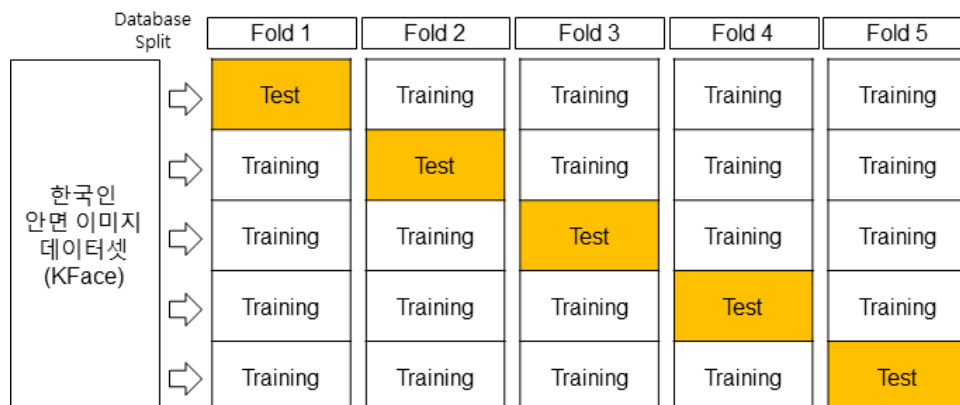


그림 12 데이터셋 유효성 검증 예시: N-Fold Cross Validation (N=5)

또 다른 유효성 검증 방법으로 그림 13과 같이 상용 데이터셋을 활용하는 검증 방법을 활용하였다. 유효성 검증에 활용한 안면 응용 인공 신경망 학습 모델은 안면 인식을 위한 LightCNN 모델과 가짜 얼굴 생성을 위한 PGGAN 모델이었다.

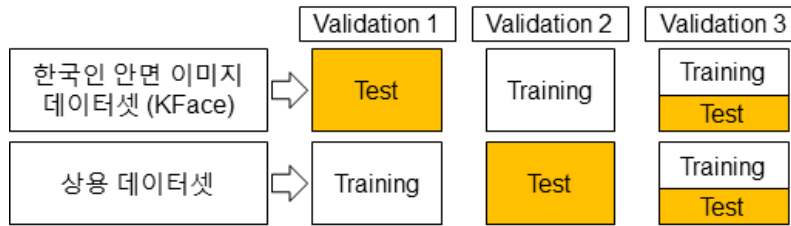


그림 13 데이터셋 유효성 검증 예시: Cross Validation

한국인 안면이미지 데이터셋은 안면 인식기 개발 뿐만 아니라 가짜 얼굴 생성, 얼굴 정면화, 가림 및 그림자 제거, 얼굴 영역 탐지 등 폭넓은 안면 응용 기술 개발에 적용 가능하며, 해당 기술에 꼭 필요한 인공 신경망 학습 및 검증에 훌륭한 재료가 될 것으로 기대한다.

데이터 구축 담당자

수행기관(주관) : 한국과학기술연구원 인공지능연구단, 이메일: jhcho@kist.re.kr