

●○ 감정인식 및 요약영상 과제

한국인 감정인식을 위한 복합 영상 데이터



●○ 개요: 감정인식 데이터셋이란

한국인을 대상으로 한 인공감성지능(Artificial Emotional Intelligence) 모델 개발을 위한 학습 데이터셋으로 KAIST에서 구축했으며, 총 50만장의 이미지로 구성되어 있다.

인공지능이 프론트 오피스 업무로 영역을 확대해나감에 따라 보다 자연스러운 인간-컴퓨터 상호 작용이 중요해지고 있다. 인공감성지능은 얼굴 표정, 심박수, 체온 등의 시각적, 신체적 요소에 감정을 추론하는 감정인식기술에 더불어, 감정 생성과 감정 증강까지를 아우르는 기술이다.

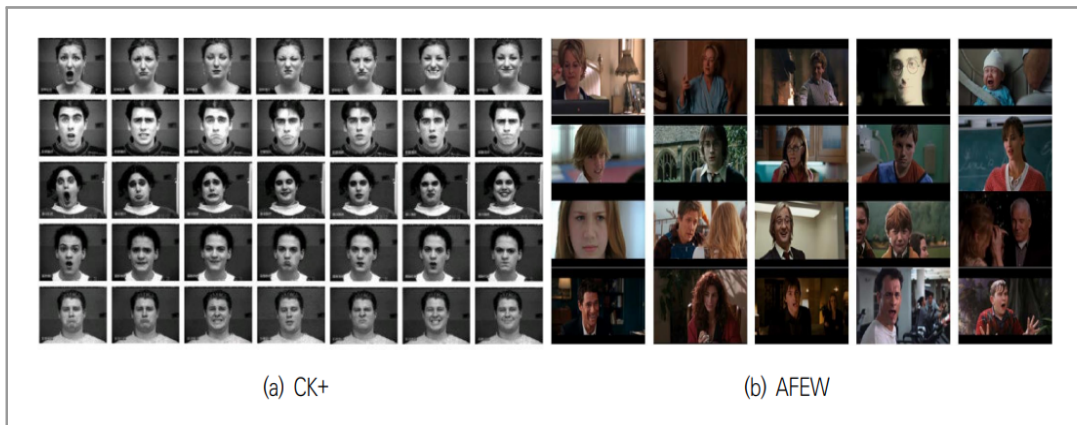


그림1 | 기존에 사용되던 감정인식 데이터셋

감정인식을 위해서는 CK+나 AFEW 데이터셋이 주로 사용되고 있다. 하지만 개인, 사회에 따라 감정의 표현 방식이 다르기에 표정만으로 감정을 추출하기 어렵다는 문제제기도 존재한다. 본 데이터셋에서는 한국인의 이미지를 수집하고, 각 이미지에 맞는 상황, 맥락 정보를 태깅하였다.

●○ 데이터셋의 구성

| 데이터 구축 규모와 특성 |

본 데이터셋은 한국인의 ‘안면이미지’를 통해 ‘감정인식’ AI 학습용 모델 구축의 목적으로 구성되며 특정 장소/배경, 특정 감정이 나타난 ‘1인 셀프카메라(Selfie)’ 50만장으로 구축되어 있다. AI 학습 모델의 정확도 증대를 위해, 특정 감정을 잘 표현할 수 있는 ‘전문인’을 별도 모집하여 25만장을 수집, 자연스러운 일반인의 표정을 묘사할 수 있는 ‘일반인’을 따로 모집하여 25만장을 수집하였으며, 이는 데이터의 포괄성과 품질에 대한 신뢰성을 확보할 수 있도록 구성되어 있다고 할 수 있다.

- 최종 구축된 안면 이미지 관련 인공지능 학습용 데이터는 사전에 정의된 총 7개 감정 Category와 총 10개 장소/배경 Category 정보가 종속된 총 50만장의 이미지, 약 3.8TB 분량
 - 안면 이미지는 1920*1080 이상의 해상도를 갖춘 이미지, 태깅 관련 CSV 파일 1건
 - 이미지는 1000*1000(px)로 크롭 및 리사이징하여 얼굴 크기가 100*100 ~ 400*400(px)
 - 이미지 내 얼굴 크기 10% 이상 40% 이하의 데이터

데이터 종류	포함 내용	제공 방식
일반인 대상 이미지	약 25만장	이미지 파일 (JPEG, JPG)포맷
전문인 대상 이미지	약 25만장	

●○ 데이터셋의 설계 기준과 분포

데이터셋 설계 시, 가장 고려했던 점이 ‘한국인’의 감정을 나타낼 수 있는 감정 분류체계를 정립하는 것이었으며, 특히 한정된 자원 하에 AI 학습용 데이터로써 향후 활용도 제고 차원에서 7개로 감정 체계를 압축하였다. 7가지 감정체계를 분류하는 기준은 Susan David의 감정분류체계에 따랐으며, 현재 세계적인 감정 이미지를 활용한 학습용 데이터 구축 흐름에 발맞추어 ‘중립(무표정)’ 감정을 별도 추가하였다.

또한 특정된 감정하에 이미지를 직접 촬영할 시 ‘일반인’ 대상으로는 이미지 데이터의 품질 확보가 우려되어 ‘전문인’을 일부 포함하여 진행하였으며, 한편 일반인 대상으로는 다양한 성/연령의 이미지 데이터 셋을 구축하기 위하여 노력하였다.

표 | A List of Emotions, Susan David, Harvard Business Review

Angry	Sad	Anxious	Hurt	Embarrassed	Happy
Grumpy	Disappointed	Afraid	Jealous	Isolated	Thankful
Frustrated	Mournful	Stressed	Betrayed	Self-conscious	Trusting
Annoyed	Regretful	Vulnerable	Isolated	Lonely	Comfortable
Defensive	Depressed	Confused	Shocked	Inferior	Content
Spiteful	Paralyzed	Bewildered	Deprived	Guilty	Excited
Impatient	Pessimistic	Skeptical	Victimized	Ashamed	Relaxed
Disgusted	Tearful	Worried	Aggrieved	Repugnant	Relieved
Offended	Dismayed	Cautious	Tormented	Pathetic	Elated
Irritated	Disillusioned	Nervous	Abandoned	Confused	Confident

표 | 감정 분류체계 7종

No.	대감정	세부감정
1	분노	툼툼대는, 좌절한, 짜증내는, 방어적인, 악의적인, 안달하는, 구역질 나는, 노여워하는, 성가신
2	슬픔	실망한, 비통한, 후회되는, 우울한, 마비된, 염세적인, 눈물나는, 낭패한, 환멸을 느끼는
3	불안	두려운, 스트레스 받는, 취약한, 헛갈리는, 당혹스러운, 회의적인, 걱정스러운, 조심스러운, 신경쓰이는
4	상처	질투하는, 배신당한, 격리된, 충격 받은, 궁핍한, 희생된, 억울한, 괴로워하는, 버려진
5	당황	격리된, 시선 의식하는, 외로운, 열등한, 죄책감의, 부끄러운, 혐오스러운, 한심한, 헛갈리는
6	기쁨	감사하는, 믿는, 편안한, 만족한, 흥분한, 느긋한, 안도하는, 신이 난, 자신하는
7	중립	무표정, 감정이 0인 상태

표 | 성/연령대별 목적적 임의할당 - 게시자 리크루팅 목표 기준

연령대	성 by 연령 데이터 구축 할당 내역(단위:명)			
	일반인		전문인	
	남	여	남	여
20대	1,000명 (성별 50% 비율 유지목표)		125	125
30대				
40대				
50대 이상				
총계	1,000명		250명	

한편, 장소/배경에 관련하여서는 현재 코로나19상황으로 다양한 장소/배경에 대한 자연스러운 방문이 어려운 상황이기 때문에, 국내 장소 카테고리를 10가지로 압축하여 학습용 데이터의 '장소/배경' 이미지를 구축하였다. 장소/배경 압축 기준은 'MIT Places CNN' 모델의 365가지 장소/분류 기준을 참고하여, 국내상황에 맞게 재가공 하였으며, 다음과 같이 구성된다.

표 | 장소 및 배경 관련 분류 기준

장소/배경	예시
공공시설/종교/의료	교회, 사찰, 관공서, 병원, 보건소 등
교통/이동수단(엘리베이터 포함)	택시, 기차, 자가용, 비행기, 육교, 다리 등
상업시설/점포/시장	재래시장, 대형마트, 종합 쇼핑몰, 식당, 빵집, 카페, 꽃집 등
스포츠 관람 및 레저시설	야구 경기장, 축구경기장, 실내 아이스링크, 볼링장, 체육관 등
오락/공연시설	콘서트장, 오락실, 놀이동산, PC방 등
행사/사무공간	사무실, 컨퍼런스 룸, 결혼식장, 구내식당 등
실외 자연환경	공원, 등산로, 호수, 바다 등
도심 환경	도심지 대로변, 빌딩 숲 등
문화재 및 유적지	문화재 및 유적지,
숙박 및 거주공간	집(실내, 실외 포함), 캠핑장, 호텔 등

●○ 데이터 구조<메트릭스>

분류		일반인 이미지	전문인 이미지
내용		감정 및 장소/배경	감정 및 장소/배경
수량		25만	25만
항목(포함하고 있는 데이터)			
감정	title	Y	Y
장소/배경	source	Y	Y

본 데이터 셋의 구조는 '원본 이미지(JPEG, JPG)' 파일을 제공하며, Ground Truth 값으로써의 '감정', '장소/배경'에 대한 태깅 정보를 제공한다. 다만 본 데이터 셋은 '개인정보보호법' 및 '생명 윤리법'에 적용됨에 따라, 안면 이미지와 부합되어 개인을 특정 할 수 있는 '연령/성별'에 대한 정보는 원천 데이터 상에서는 제공되지 않는다.

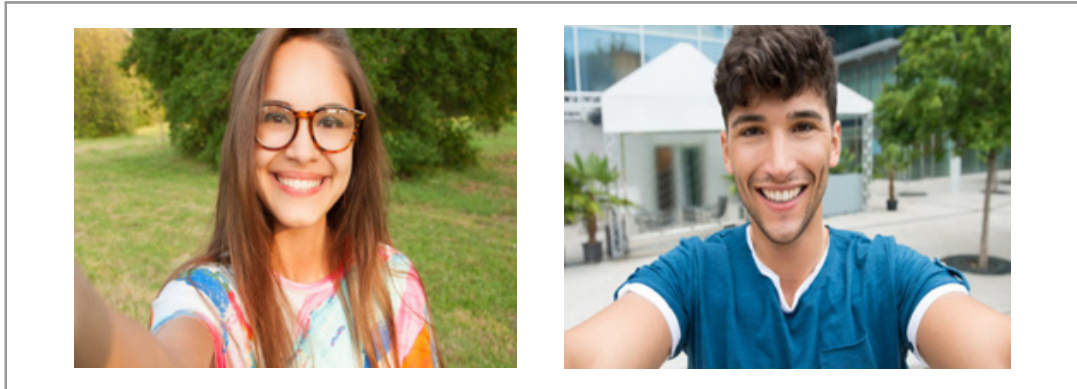
●○ 데이터 셋 예시

- 현재 예시와 관련해서는 '과업 수행중에 있으며' IRB 피험자 동의서를 최종 완료한 이미지 데이터를 보유하고 있지 않은 바 아래와 같은 예시만 제출합니다.

jpg, jpeg 형태 이미지 약 50만 장

(1) 인물의 감정 상태 포함

(2) 다양한 장소/배경 포함



●○ 데이터 구축 과정

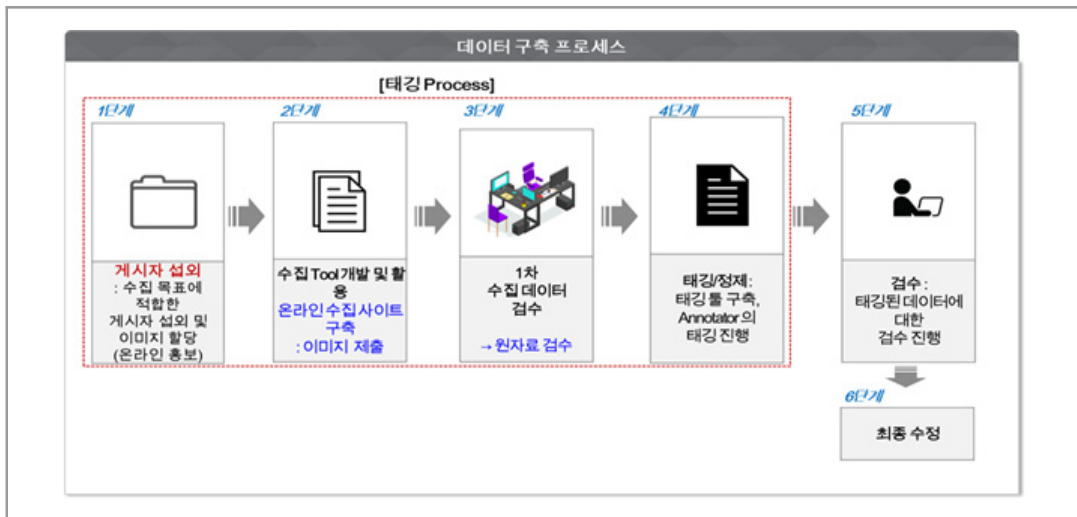


그림2 | 데이터 구축 프로세스

감정 이미지는 클라우드 소싱 방식을 이용해 일반인 집단이 자발적으로 게시하도록 하였다. 게시자 성비는 50:50으로, 게시 이미지는 게시자의 연령층에 따라 비례 할당했다. 또한, 편향성 방지와 감정의 다양성 확보를 위해 전문인(배우 등) 250명과 일반인 2500명을 대상으로 시나리오를 지정하여 50만 장의 이미지 데이터를 확보하였다.

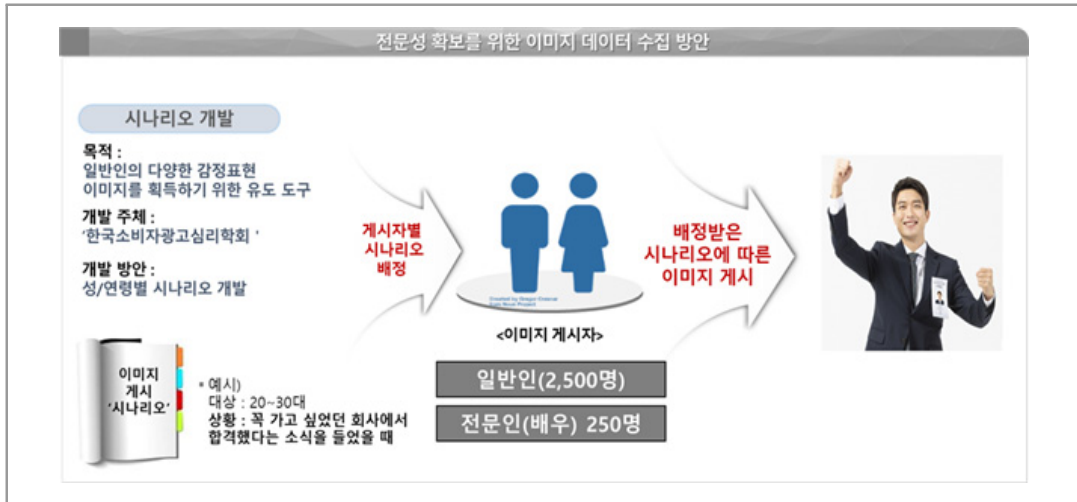


그림3 | 이미지 데이터 확보 방안 흐름도

이렇게 수집된 이미지들에 ‘사진 게시자’와 ‘Annotator’가 관계맥락 정보를 전용 툴을 이용해 태깅하도록 하였다. 게시자는 감정평가 대상자의 연령, 감정상태, 촬영 장소, 일시를 기재하였고, Annotator는 게시자가 등록한 정보가 blind 된 상태에서 감정평가 대상자의 심리 상태를 유추하여 기입하도록 하였다.

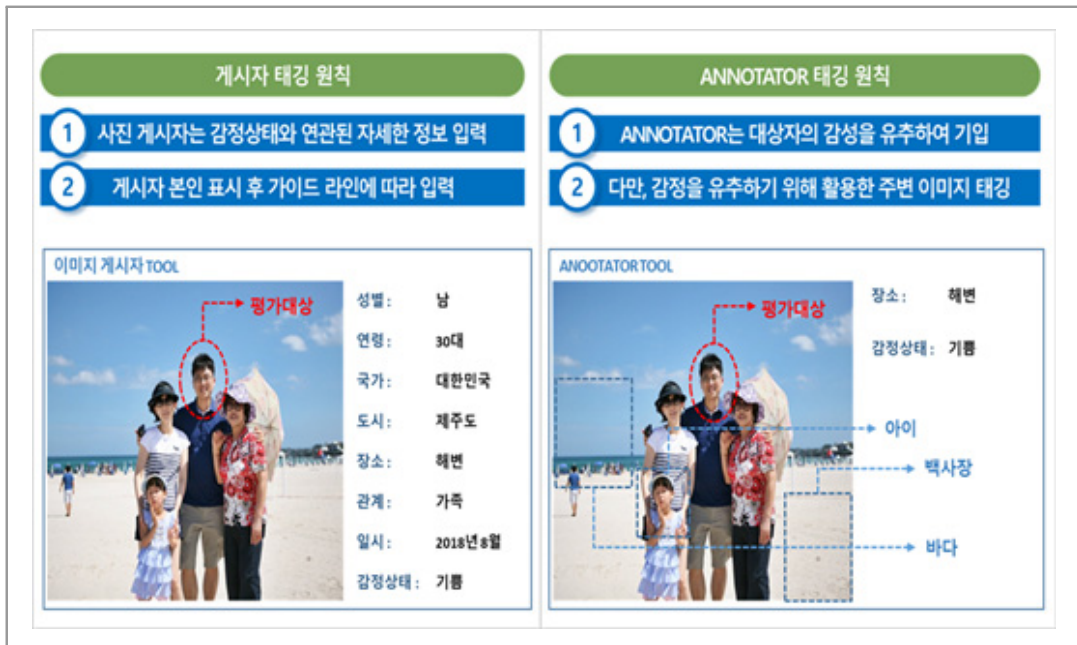


그림4 | 게시자, Annotator가 사용하는 태깅 툴

●○ 검수와 품질 확보

감정인식 과제는 50만장의 유저 수집 데이터로 이루어져 있습니다. 최종 레이블된 데이터의 결과물의 품질은 데이터의 수집단계에서부터 결정됩니다. 인공지능 데이터는 ‘garbage in, garbage out (질 낮은 데이터가 투입되면, 질 낮은 성능이 나온다)’ 규칙을 대표적으로 따르고 있습니다. 수집 단계에서부터 질 낮은 데이터가 수집될 경우 데이터 가공만으로는 양질의 학습 데이터를 만들 수 없습니다. 데이터헌트는 이번 과제의 대표 가공 업체로 수집단계에서부터 양질의 데이터가 수집될 수 있도록 관여해왔습니다. 일례로 데이터헌트는 아래와 같은 가공 전 조치를 취하였습니다.

1. 수집 가이드라인 수집기관과 함께 작성
2. 샘플 데이터 확보해 직접 사전 가공 → 문제점 수집기관에 전달
3. 샘플 데이터 결과물 활용기관(모델링)에게 데이터 퀄리티 확인
4. (가공 후) 수집된 데이터를 가공하면서 발견되는 불량 데이터를 수집기관에 반려해 재수집 요청 예정

위와 같은 선제적인 조치로 가공 전 불량 데이터를 막을 수 있습니다.

양질의 데이터가 수집된 전제라면 이제 이번 과제에 가장 적합한 가공 프로세스 수립이 데이터 품질을 결정하는 일만 남았습니다. 과제의 검수와 품질 확보를 위해 고려해야할 key factors들은 다음과 같았습니다.

1. 주관성: 얼굴의 감정 태깅은 어노테이터마다 기준이 다를 수 있음
2. 리소스: 방대한 데이터 양과 제한된 시간
3. 활용성: 최종 결과물로 어떤 모델을 만드는가?

● 주관성:

- 먼저, 주관성 문제를 해결해야 합니다. 이번 데이터 가공과 검수 프로세스는 한 이미지에 대해 3명의 각기 다른 어노테이터가 작업을 하고 3명의 어노테이터의 결과물을 모두 공개하는 것으로 주관성 문제를 해결할 수 있습니다. 사람의 감정을 어떻게 인식하느냐는 개개인의 편차가 있는 것이고 인공지능 모델도 여러 사람의 판단을 종합해 학습시키는 것이 더 정교한 감정인식 모델을 만들 수 있을 것이라 판단했습니다.

● 리소스:

- 가공 데이터의 결과물은 작업 시간(인력)에 비례할 수 밖에 없습니다. 현재 과제 연장 있었지만 수집단계에서 많은 시간이 소요되어 150만장(50만장 x3)을 한달안에 가공해야 하는 절대적인 시간 부족 현상이 있습니다. 또한, 과제비로 사전에 제한된 비용이 있기에 추가 비용을 발생

시켜 과제를 해결할 수 없습니다. 가공 기업에게 더 큰 리소스를 할하는 것이 가공 품질을 높이는 가장 확실한 방법임을 고려하여 추후 과제를 디자인하는 게 좋을 것이라 생각합니다. 이미 과제는 시작되었고 현재 주어진 리소스에서 최선의 방법을 컨소시엄내 모든 기업의 합의하에 도입하기로 결정하였습니다. 다음과 같은 조치들로 품질을 보장합니다.

1. 인공지능 전처리: 가공기업은 보유중인 인공지능 모델의 전처리를 이용해 이미 일정수준 이상의 전처리 결과물을 작업자에게 배포
2. 선택 제한: 작업시 불량데이터를 산출할 수 있는 여지를 애초에 차단. (예, 1개만 선택가능한 항목에 대해선 중복 선택 불가 등)
3. 결과물 비교: 3명의 작업자 결과물이 모두 다 다를 경우 모델링 기관에서 데이터 제외하거나 2명 이상의 작업자가 일치한 결과물만 모델에 사용하는 등 유동적으로 모델링에 적합한 데이터 도출

- 검수

- 위와 같은 조치들은 이번 과제의 데이터의 품질을 최대한으로 보장할 수 있습니다. 데이터 라벨링에 대한 검수조치는 단순히 필요한 데이터가 기입되었는지만 확인하는 수준이어야 합니다. 이는 주관성과 관련된 부분으로 검수자의 주관이 작업자의 주관보다 정확하다는 근거가 없기 때문입니다. 이번 과제의 특수성을 고려해 기존에 일반적인 전량 데이터값에 대한 '검수'는 없이 3명의 작업자가 작업을 하는 것으로 대체합니다. 이 또한 데이터 품질을 높이는데 최적의 검수 방안으로 생각합니다.

●○ 데이터 구축 담당자

수행기관(주관) : 카이스트

(전화: 010-4028-0212, 이메일: nerinary@kaist.ac.kr)