

AI 초개인화 맞춤형 피부진단을 위한 한국인 피부상태 측정 데이터 구축

이정호¹, 양주열², 최민서¹, 최상일¹

¹ 단국대학교 컴퓨터공학과

² 단국대학교 소프트웨어학과

72210297@dankook.ac.kr, 32192539@dankook.ac.kr, chlalstjek@dankook.ac.kr, choisi@dankook.ac.kr

Constructing a Dataset for Assessing Skin Condition in Koreans for AI-Personalized Customized Skin Diagnosis

Jeongho Lee¹, Juyeol Yang², Minseo Choi¹, Sang-Il Choi¹

¹Dept. of Computer Engineering, Dankook University

²Dept. of Software Engineering, Dankook University

요 약

최근 들어, 미용 상품을 선택하기 전에 자신의 피부 타입과 상태를 정확히 파악하고 맞춤형 상품을 선택하고자 하는 수요가 증가하고 있다. 이에 따라 피부 상태 측정을 위한 기술적 요소의 중요성이 더욱 두드러지고 있다. 그러나 현재까지 피부 상태 측정을 위한 데이터셋이 한국인을 대상으로 측정된 데이터셋이 없는 실정이다. 본 연구에서는 한국인의 피부 상태를 정밀하게 분석하기 위해 고해상도 디지털 카메라로 촬영된 이미지, 정밀 피부측정 장비를 활용하여 측정된 정밀 값, 그리고 피부과 전문의가 진단한 피부상태 진단 등급 데이터를 통합하여 제공을 한다. 추후 제작한 데이터셋을 활용하여 개인 맞춤형 미용상품 추천과 개발 등 다양한 분야에 활용하고자 한다.

1. 구축 배경

현대 사회에서 피부 건강과 외모에 대한 관심이 크게 증가하고 있으며, 많은 사람들이 자신의 피부 상태를 관리하고 개선하기 위해 다양한 화장품과 관리 방법을 찾고 있다. 각 개인의 피부는 고유한 특성을 가지고 있어 피부 타입, 민감도 등이 다양하기 때문에 정확한 피부 상태측정 기술이 요구된다. 비약적인 성장을 이룬 온 인공지능 기술과 빅데이터 분석을 통해 대규모의 피부 측정 데이터를 구축하고 인공지능 모델에 학습시켜 피부 상태를 평가하고 예측하는 것이 가능해져 최근 피부 관련 AI 모델에 대해서 관심이 부쩍 높아졌다.

최근까지 해외 기업들에서 여러가지 피부 상태 측정 어플들을 공개하였지만, 그에 비해 국내에서는 관련 서비스들을 찾아보기 힘든 실정이다. 본 연구에서는 고해상도 안면 이미지와 피부상태 라벨링 데이터를 통합하여 구축한 한국인 피부상태 데이터셋을 공개 할 계획이다.



(그림 1) 해외 기업들의 피부 상태 측정 서비스

2. 안면 이미지 촬영

사용자가 촬영하는 환경(ex. 조도, 습도, 온도 등)에 따라서 피부 상태가 크게 달라질 수 있다. 이러한 변형은 인공지능 모델이 이미지만을 가지고 정확한 피부 상태를 추정하는 것에 대해 악영향을 끼친다. 본 연구에서는 안면 이미지가 장소와 날씨에 영향을 받지 않고 항상 동일한 조건에서 AI 모델이 학습할 수 있게 통제된 환경을 만들어 촬영을 진행하며, 절차는 아래와 같다.

- 1) 모든 피실험자는 촬영하기 전에 세안을 하여, 얼굴에 이물질 제거한다.
- 2) 측정실의 온도, 습도가 일정하게 유지된 항온항

습실에서 15분간 대기를 하여, 외부 날씨 등 환경을 영향을 받지 않고 정확한 피부상태를 측정하게 해준다.

- 3) 조명 장치가 설치된 사진실 안에서 고해상도 디지털 카메라로 자체 제작한 귀 고정대를 사용하며, 레이저 포인트가 입가에 평행하게 얼굴 위치를 조절하여 최대한 모든 피실험자가 같은 위치를 맞춘다.
- 4) 카메라 각도를 자체 제작한 레일을 통해 정밀하게 조절하여 얼굴 정면, 위, 아래, 왼쪽 15도, 30도, 오른쪽 15도, 30도 순으로 촬영을 진행한다.
- 5) 스마트폰과 스마트 패드로 정면, 좌, 우 순으로 항온항습실에서 촬영을 한다.

3. 정밀 피부측정

정밀 피부측정 장비로 측정하는 종류는 크게 수분, 탄력, 주름, 색소침착, 모공으로 구분된다.



(그림 2) 사진실에서 촬영하는 모습

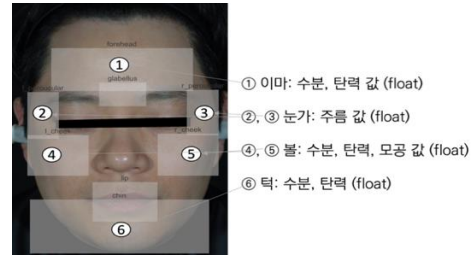
측정 장비에 해당하는 PRIMOS Lite[1](수분), Cutometer MPA580[2](주름)과 VISIA Complexion Analysis[3](색소침착/모공)는 모두 SCI 급 논문에 빈번하게 활용되며, 전세계 피부연구자에게서 표준장비로 인정받는 장비이다. 그리고, 측정하는 값들의 범위가 모두 상이하여, 정규화를 통해 0 과 1 사이의 실수 값으로 매핑해주었다.

4. 전문가 진단 라벨링

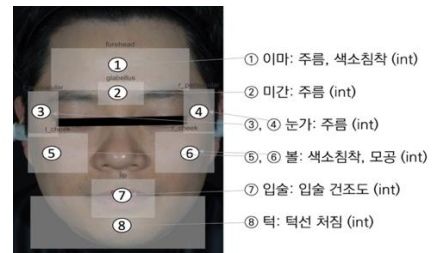
피부과 전문의 5인이 총 8개의 영역에 대한 각 부위별 피부상태 진단 등급을 매긴다. 영역은 이마, 미간, 양쪽 눈가, 양쪽 볼, 입술 그리고 턱으로 구분된다. 진단하는 종류는 주름(0-8), 색소침착(0-5), 모공(0-5), 입술 건조도(0-4), 턱선 처짐(0-6)이 있으며, 등급은 정수 값으로 매긴다. 진단 값의 정확도를 높이기 위해 총 2차 검증을 진행한다.

- 1) 1차 검증
 - A. 피부과 전문의 5인이 전문가 진단 라벨링을 진행한다.
- 2) 2차 검증
 - A. 5인의 전문가 진단 등급 값이 2인 이상의

진단 등급 값과 차이가 발생하게 되면, 내부 회의를 거쳐 최종 진단 등급 값을 결정한다.



(그림 5) 정밀 피부측정 장비 측정 라벨링



(그림 6) 전문가 진단 피부상태 라벨링

5. 데이터셋

1,000 명에 대한 안면 이미지 13,000 장을 수집하고 있는 중이며, 본 논문에서는 현재까지 측정한 100 명에 대한 안면 이미지 1,300 장을 AI 모델 학습에 활용하였다.

데이터셋에서 안면 이미지 데이터는 1 명에 대한 이미지가 디지털 카메라 7 장, 스마트폰 3 장, 스마트패드 3 장으로 총 13 장의 이미지로 구성 돼 있으며, 디지털 카메라의 경우에는 정면, 좌/우 15, 30도, 위, 아래로 구성 돼 있고 스마트폰과 스마트패드의 경우에는 정면, 좌, 우로 구성 돼 있다.

라벨링 데이터는 전문가 진단 라벨링과 정밀 측정 장비 라벨링으로 구분된다.

전문가 진단 라벨링은 주름, 색소침착, 모공, 입술, 턱으로 이뤄져 있으며 값의 형태는 정수형을 띄고 있다.

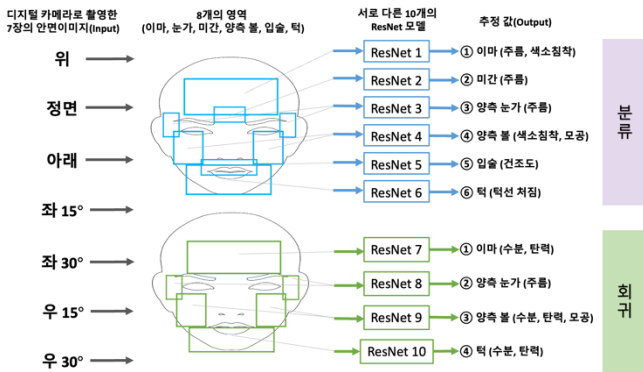
정밀 측정장비 라벨링은 수분, 탄력, 모공, 주름로 이뤄져 있으며, 값의 형태는 정규화를 진행하여 0 부터 1 사이의 실수 값을 띄고 있다.

6. AI 모델 학습

본 연구에서 획득한 데이터를 실제 AI 모델에 학습하여 피부상태 측정에 사용가능한 지에 대한 확인이 필요하여, 수년 간 널리 전 분야에 걸쳐 사용되어 충분히 검증이 된 CNN 모델인 ResNet-50[4] 모델을 사용하여 성능을 평가하였다.

모델의 작업은 크게 2 종류로 나뉘게 된다. 전문가 진단 등급을 분류하는 작업과 정밀 피부측정 장비의 값을 회귀하는 작업으로 나뉘며, 각 작업에 입력으로 들어가는 얼굴 영역에 맞춰 분류는 6 개의 ResNet 모

델이 사용되고 회귀에는 4 개의 ResNet 모델이 각각 사용된다.



(그림 8) 7 장의 멀티뷰 이미지를 입력으로 받는 AI 모델개요

7. 성능

현재까지 수집한 1,300 장의 안면 이미지 중 1,200 장을 모델 학습에 사용하였으며, 나머지 100 장은 모델 성능평가에 사용하였다. 성능평가의 결과는 전문가 진단 값의 평균 정확도가 주름, 색소침착, 모공, 턱선 처짐, 입술 건조도 순으로 75%, 74.07%, 72.22%, 77.78%, 88.89%이며, 정밀 측정장비 값의 MAE 오차는 주름, 탄력, 수분, 모공 순으로 0.0706, 0.053, 0.0617, 0.0791 이다.

8. 결론

지금까지 얼굴 피부상태 관련 데이터셋은 대부분 서양인에 대한 데이터이며, 동양인 데이터는 극히 드물고 특히 한국인 데이터는 전무하다. 그래서 한국인 얼굴 피부상태 데이터셋을 본 연구를 통해 구축하였고, 1 명에 대한 고정된 각도(정면, 상, 하, 좌/우 15, 30 도)에서 촬영한 7 장의 이미지를 획득하여 얼굴 전체적인 피부정보를 얻을 수 있으며, 통제된 환경에서 모든 이미지를 촬영하여 외부 방해요인을 최소화 해주었다. 또한, 안면 이미지에서 의미 있는 8 개의 영역으로 분할하여 영역별로 주요하게 보는 피부상태 종류를 구분 지었고 정밀 피부측정 장비와 피부과 전문의 5 인의 진단 라벨링이 제공된다. 검증된 AI 모델을 통해 본 데이터셋을 가지고서 안면 이미지만으로 피부상태 추정 가능성을 보여주었고 수치상으로 전문가 진단에서 평균 77.6%의 정확도와 정밀장비 측정값 평균 MAE 0.0661 오차를 보여주었다. 앞으로 1,000 명에 대한 13,000 장 안면 이미지와 라벨링 데이터를 통합하여 데이터셋 구축을 마무리 할 계획이다.

사사문구

“ 본 연구는 2023 년 과학기술정보통신부와 한국지능정보사회진흥원에서 지원하는 AI 데이터셋 프로젝트 (aihub.or.kr)으로 수행되었음”

참고문헌

- [1] Primos 매뉴얼, http://www.questar-china.com/download/GFM_Primos.pdf
- [2] Cutometer 홈페이지, <https://www.enviroderm.co.uk/products/cutometer-dual-mpa-580>
- [3] VISIA 홈페이지, <https://www.canfieldsci.com/imaging-systems/visia-complexion-analysis/>
- [4] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.