

AI-Powered Personal Beauty Advisor

2025 전기 졸업과제 최종보고서 <Underdog>



202155659 Han Nwae Nyein

202155658 Nyi Nyi Htun

202055594 전재원

지도교수 : 전상률

목 차

1. 서론.....	1
1.1. 연구 배경.....	1
1.2. 기존 문제점.....	1
1.3. 연구 목표.....	2
2. 배경 지식.....	3
2.1. 퍼스널 컬러 클러스터링.....	3
2.2. 클러스터 시각화 및 팔레트 설계.....	3
2.3. 얼굴 부위별 색상 오버레이.....	3
3. 연구 내용.....	4
3.1. 시스템 구현 및 개발 환경.....	4
3.2. 데이터 전처리.....	5
3.2.1. 적응형 조명 보정 (Adaptive Lighting Correction).....	5
3.2.2. 얼굴 감지 및 파싱 후 피부 영역 추출.....	7
3.3. 모델설계.....	8
3.3.1. 핵심 특징 벡터(Feature Vector) 설계.....	8
3.3.2. 최적 클러스터 개수(k) 탐색 설계.....	9
3.4. 모델학습.....	9
3.5. K 값 선정 및 클러스터링 평가.....	11
3.5.1. 주성분 분석(PCA)을 활용한 클러스터링 결과 시각화.....	11
3.6. 클러스터 분석 및 팔레트 설계.....	13
3.6.1. 클러스터별 대표 이미지 시각화.....	13

3.6.2.	HSV/HSB 색상 공간 분석	16
3.6.3.	클러스터 라벨링 및 초기 팔레트 설계	17
3.7.	컬러 오버레이를 위한 사전 학습 모델 설치	19
3.7.1.	메이크업 모듈 및 컬러 오버레이 개선	19
3.8.	웹 설계 및 구현	21
3.8.1.	기능 정의서	21
3.8.2.	결과물	23
4.	연구 결과 분석 및 평가	25
5.	결론 및 향후 연구 방향	27
5.1.	결론	27
5.2.	향후 연구 방향	27
6.	팀원별 역할 및 개발 일정	28
6.1.	팀원별 역할	28
6.2.	개발 일정	30
7.	참고 문헌	31

1. 서론

1.1. 연구 배경

현대 사회에서 패션과 뷰티는 단순한 외모 관리의 영역을 넘어, 개인의 정체성과 개성을 표현하는 중요한 수단으로 발전하였다. 머리 색, 콘택트 렌즈, 메이크업, 의류 등 다양한 요소를 통해 사람들은 자신이 원하는 이미지를 자유롭게 연출하고 변형할 수 있다.

이러한 맥락에서 퍼스널 컬러 진단은 자신에게 가장 잘 어울리는 색을 찾아, 자기 표현과 자기 개발을 돕는 중요한 도구로 주목받고 있다.

하지만 기존의 퍼스널 컬러 진단은 전통적인 4 계절, 12 유형 Armocromia 시스템에 의존하거나, 머리색과 눈동자 색 등 고정된 신체적 특징이나 전문가의 주관적 판단에 크게 의존하는 경우가 많았다. 특히 머리색과 눈동자 색은 염색이나 렌즈 착용으로 쉽게 변할 수 있어 진단 기준으로서의 안정성이 떨어진다. 이로 인해 제시되는 색 추천은 일반적이고 실제 스타일링 상황에서 적용 가능성이 낮으며, 사용자 기대와 결과 사이의 불일치가 발생하기 쉽다. 또한, 시뮬레이션 기능이 부족하여 사용자가 추천 색상을 직관적으로 경험하기 어렵다는 한계가 있다.

이에 본 연구에서는 피부 톤 분석에만 초점을 맞춘 AI 기반 퍼스널 컬러 시스템을 제안하고, 새로운 퍼스널 컬러 프레임워크를 구축하고자 한다. 또한, 색상 분석, 스타일 추천, 가상 스타일 체험을 통합한 AI-Powered Personal Beauty Advisor 를 구현함으로써, 누구나 쉽고 편리하게 자신에게 최적화된 스타일을 탐색할 수 있도록 하는 것을 목표로 한다.

1.2. 기존 문제점

기존 퍼스널 컬러 진단 서비스는 다음과 같은 한계를 지닌다.

1. **접근성 제약:** 전문가 대면 진단이 필요하거나 비용이 과도하게 소요되어 일반 사용자가 접근하기 어렵다.
2. **유연성 부족:** 사용자가 렌즈나 헤어 컬러와 같은 요소를 통해 스타일을 변경할 경우, 기존의 퍼스널 컬러 진단은 실제 활용도가 떨어질 수 있다.
3. **주관성 문제:** Armocromia와 같은 기존 체계는 고정적 분류에 머물러 있으며, 전문가의 경험과 주관에 크게 의존한다.
4. **정확도 및 성능 문제:** 조명 변화에 따른 피부톤 왜곡, 대규모 모델의 실시간 추론

속도 저하 등 기술적 제약이 존재한다.

5. **시각화 한계:** 추천 색상이 실제 적용된 모습을 제공하지 못해, 사용자에게 직관적 경험을 제공하기 어렵다.

이러한 한계는 사용자가 자신에게 최적화된 색상을 탐색하고 이를 실제 스타일링에 적용하는 데 제약을 초래한다.

1.3. 연구 목표

본 프로젝트의 목표는 기존의 제한적이고 주관적인 퍼스널 컬러 진단을 넘어, 데이터 기반의 새로운 체계와 AI 기술을 결합한 **맞춤형 뷰티 어드바이저 시스템**을 구축하는 것이다. 구체적으로 다음과 같은 목표를 설정하였다.

1. **정확한 피부톤 분석:** MediaPipe, Dlib, Facer 등을 활용한 얼굴 부위 정밀 추출과 Histogram Equalization, White Balance 보정 등을 통해 다양한 환경에서도 정확한 피부톤을 분석한다.
2. **새로운 퍼스널 컬러 체계 구축:** 5천 장의 얼굴 데이터셋과 K-Means 클러스터링 (Elbow Method, Silhouette Score 적용)을 기반으로 기존 12유형 체계가 아닌 데이터 기반의 새로운 컬러 그룹을 제시한다.
3. **컬러 가이드 제공:** 피부톤별로 최적화된 색상 팔레트를 제시하여 사용자가 참고할 수 있는 색상 가이드를 제공한다.
4. **스타일 제안 기능 구현:** 분석된 결과를 바탕으로 3가지 조합의 스타일 팔레트를 추천하여 사용자의 선택 폭을 넓힌다.
5. **스타일 체험 시각화:** 추천된 색상을 **렌즈, 립스틱, 헤어, 의상**에 가상 적용하여 실제감 있는 시뮬레이션 경험을 제공한다.
6. **확장성 있는 플랫폼:** 향후 사용자 DB 구축을 통해 개인 맞춤형 추천과 화장품·패션 브랜드 연계 서비스로 확장 가능성을 확보한다.

이를 통해 본 프로젝트는 단순한 색상 추천을 넘어, **데이터 기반 객관성 + 사용자 중심 유연성 + 실제감 있는 시각화**를 결합한 차세대 퍼스널 뷰티 플랫폼을 제안한다.

2. 배경 지식

2.1. 퍼스널 컬러 클러스터링

본 연구에서는 약 5 천 장 규모의 얼굴 이미지 데이터셋을 기반으로 K-Means 클러스터링을 수행하여 새로운 퍼스널 컬러 분류 체계를 구축하였다. 얼굴 감지 및 파싱을 통해 피부 톤 특징을 정량적으로 추출하고, Elbow Method 와 Silhouette Score 를 활용하여 최적의 군집 수를 산출함으로써 기존의 4 계절·12 유형 체계보다 더 객관적이고 데이터 기반적인 분류 방식을 제시하였다.

2.2. 클러스터 시각화 및 팔레트 설계

각 클러스터는 RGB, HSV, Lab 색 공간 분석을 통해 대표 색상을 정의하였으며, 이를 토대로 시각적 이름과 팔레트를 설계하였다. 이러한 팔레트는 사용자 맞춤형 색상 추천의 기반이 되며, 뷰티 및 패션 스타일링에 직접적으로 활용될 수 있다.

2.3. 얼굴 부위별 색상 오버레이

단순한 피부 톤 분류를 넘어, BiSeNet 기반 얼굴 분할 모델을 사용하여 머리카락, 입술, 눈, 의상 등 특정 부위를 정밀하게 분리하였다. 메이크업, 염색, 컬러 렌즈, 립 색상, 의상 스타일을 사실적으로 시뮬레이션하기 위해 본 연구에서는 소프트 라이트 블렌딩(soft light blending) 기반의 컬러 오버레이 기법을 구현하였다. 처음에는 일반적인 알파 블렌딩(alpha blending) 방식을 적용해 보았으나, 이 방법은 특히 검정색과 같이 어두운 영역의 색상을 다른 색으로 바꾸는 데 한계가 있어 원하는 결과를 얻기 어려웠다. 이러한 문제를 해결하기 위해 보다 효과적인 방법인 소프트 라이트 블렌딩 기법을 채택하였으며, 이를 통해 인위적인 색상을 얼굴이나 의상 영역에 자연스럽게 합성하면서도 원래 이미지의 질감, 명도, 음영을 그대로 유지할 수 있었다. 소프트 라이트 블렌딩은 기저 픽셀 BB 와 혼합 픽셀 SS 가 [0,1] 범위로 정규화되어 있을 때 다음과 같이 정의된다.

$$\text{Result} = \begin{cases} 2 \cdot B \cdot S + B^2 \cdot (1 - 2S), & \text{if } S < 0.5 \\ 2 \cdot B \cdot (1 - S) + \sqrt{B} \cdot (2S - 1), & \text{if } S \geq 0.5 \end{cases}$$

여기서,

- B 는 기저(base) 픽셀의 밝기 값이고,
- S 는 혼합(blend) 픽셀의 밝기 값을 의미한 다.

이 공식은 원본 이미지의 어두운 영역은 그대로 음영을 유지하고, 밝은 영역은 부드럽게 밝아지도록 하여, 단순한 알파 블렌딩에 비해 더욱 사실적이고 시각적으로 자연스러운 결과를 만들어낸다.

3. 연구 내용

3.1. 시스템 구현 및 개발 환경

AI 기반 퍼스널 뷰티 어드바이저는 백엔드, AI 엔진, 프론트엔드 인터페이스를 포함한 최신 기술 스택을 활용하여 풀스택 웹 애플리케이션으로 개발되었다.

- **백엔드 개발:** 서버 사이드 아키텍처는 Python 과 Flask 웹 프레임워크를 기반으로 구축되었다. Flask 는 가볍고 유연성이 뛰어나 사용자 요청 처리, 이미지 데이터 관리, 분석 결과 제공을 위한 강력한 RESTful API 를 구현하기에 적합했다.
- **AI 및 컴퓨터 비전 엔진:** 시스템의 핵심은 AI 엔진이며, 여러 핵심 라이브러리를 통합하였다.
 - **PyTorch & Facer:** 얼굴 분석의 핵심 연산을 위해 사용되었다. 특히 PyTorch 기반의 *facer* 라이브러리를 활용하여 얼굴 감지(retinaface) 및 얼굴 파싱(farl/celebm)을 수행했으며, 이는 피부 픽셀을 분리하는 데 필수적이었다.
 - **Scikit-learn:** 머신러닝 기법의 중심 라이브러리로, K-Means 알고리즘을 통해 피부 톤을 특징 벡터로 요약하고 최종 클러스터링을 수행하여 새로운 퍼스널 컬러 체계를 정의했다.
 - **OpenCV & Pillow:** 이미지 파일 입출력, 색공간 변환(RGB → Lab), 감마 보정(Gamma Correction)과 같은 조명 보정 파이프라인 구현 등 다양한 이미지 처리 작업에 활용되었다.
- **프론트엔드 개발:** 사용자 인터페이스는 HTML, CSS, JavaScript 를 사용해 개발되었다. 이를 통해 반응형이면서 직관적이고 사용하기 쉬운 웹 인터페이스를 구현하여, 사용자가 사진을 업로드하고, 분석 결과를 확인하며, 가상 메이크오버 기능을 활용할 수 있도록 했다.

3.2. 데이터 전처리

피부 톤만을 기반으로 한 분석의 정확성과 안정성을 보장하기 위해, 우리는 종합적인 데이터 전처리 파이프라인을 구축하였다. 기존의 퍼스널 컬러 체계가 머리카락이나 눈 색 상처럼 변할 수 있는 요소에 의존하는 반면, 본 연구의 방법론은 오직 개인의 피부 톤에만 집중한다. 이러한 접근은 두 가지 주요 과제를 해결하기 위한 견고한 전처리 단계가 필요하다. 첫째는 조명 조건의 가변성이고, 둘째는 피부 이외의 얼굴 요소로 인한 간섭이다. 우리의 파이프라인은 이 두 가지 문제를 순차적인 두 단계 과정을 통해 해결한다.

3.2.1. 적응형 조명 보정 (Adaptive Lighting Correction)

실제 촬영된 사진은 다양한 조명 조건에서 촬영되며, 종종 최적이지 아닌 환경에서 촬영되어 피부 톤이 심하게 왜곡될 수 있고, 이로 인해 분석 결과가 부정확해질 수 있다. 이를 완화하기 위해, 본 시스템에서는 **comprehensive_lighting_correction** 함수로 관리되는 적응형 조명 보정 파이프라인을 구현하였다.

이 파이프라인은 단일 고정 보정을 적용하지 않는다. 대신 먼저 **analyze_lighting_conditions** 함수를 사용하여 각 이미지의 조명 문제를 진단하며, 저조도(*underexposure*), 과다 노출(*overexposure*), 낮은 대비(*low contrast*) 등 문제를 확인한다. 진단 결과를 기반으로, 다음과 같은 맞춤형 보정 순서가 적용될 수 있다:

- **화이트 밸런스 보정 (White Balance Correction):** 서로 다른 광원(예: 실내 전구의 노란빛, 그들의 파란빛)으로 인해 발생하는 비현실적인 색상 편향을 중화하여 색상을 자연스럽게 조정한다.
- **적응형 히스토그램 균등화 (Adaptive Histogram Equalization, CLAHE):** 이미지 전반의 국부 대비를 향상시켜, 평평하거나 불균일한 조명에서 손실될 수 있는 피부의 미세한 질감과 디테일을 드러낸다.
- **그림자 및 하이라이트 보정 (Shadow & Highlight Correction):** 얼굴의 가장 어두운 영역과 가장 밝은 영역을 독립적으로 조정하여, 깊은 그림자나 밝게 날아간 부분에서 손실된 디테일을 회복한다.
- **감마 보정 (Gamma Correction):** 전체 밝기를 비선형적이고 인지적으로 조정하는 기법으로, 중간 톤을 지능적으로 조절하여 극단적인 어두움이나 밝음에서도 디테일을 유지하면서 노출 부족 또는 과다 노출을 보정한다.

- **언샤프 마스크 (Unsharp Masking):** 가장자리 대비를 높여 얼굴 특징의 선명도와 정의도를 향상시키는 최종 샤프닝 단계이다.

이와 같은 지능적이고 다단계의 보정은 퍼스널 컬러 분석 단계에서만 적용되어, 가장 정확한 특징 추출을 보장한다. 이후 가상 스타일링 단계에서는 원본 이미지를 그대로 사용하여, 최종 메이크오버 결과가 실제와 유사하고 사용자 사진에 충실하도록 설계하였다

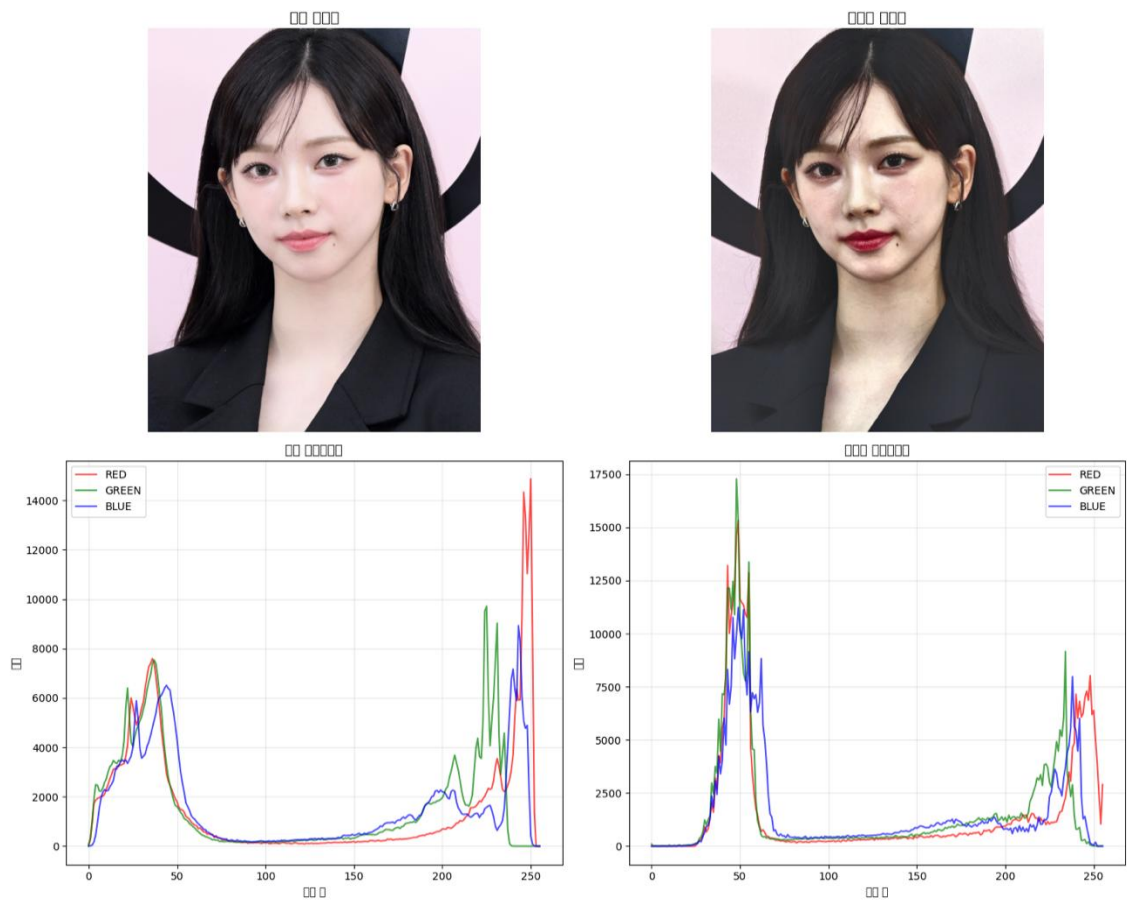


그림 3-1: 원본 이미지와 적응형 조명 보정 적용 후 이미지 비교

3.2.2. 얼굴 감지 및 파싱 후 피부 영역 추출

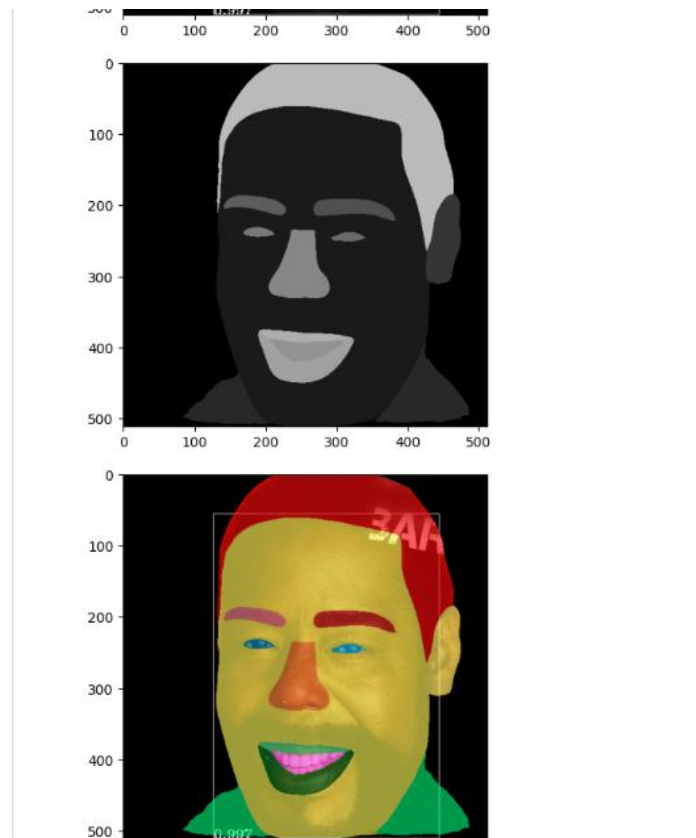


그림 3-2: 피부 영역 마스킹 및 얼굴 세분화 결과

조명 보정 이후, 다음 단계는 피부 픽셀을 얼굴의 다른 특징과 배경으로부터 분리하는 것이다. 이를 위해, 고성능 얼굴 분석 모델을 제공하는 facer 라이브러리를 활용하였다. 과정은 두 가지 모델로 이루어진다:

1. **얼굴 검출(Face Detection):** 먼저 얼굴 검출기가 얼굴의 위치를 식별한다.
2. **얼굴 파싱(Face Parsing):** 이후 의미론적 세분화(Semantic Segmentation) 모델이 얼굴 영역을 분석하여, 각 픽셀에 피부, 머리카락, 눈, 입술 등 레이블을 지정한 상세한 맵을 생성한다.

이 파싱 맵을 통해 정확한 피부 마스크를 만들 수 있으며, 피부가 아닌 요소를 효과적으로 제거한다. 그 결과, 오직 피부 픽셀만 포함된 이미지가 생성되며, 이는 이후 색상 특징 추출과 K-Means 클러스터링을 위한 깨끗하고 신뢰할 수 있는 기반을 제공한다.

3.3. 모델설계

본 연구는 기존의 사계절 퍼스널 컬러 분류 체계를 넘어, 데이터 기반의 새로운 컬러 그룹을 도출하는 것을 목표로 한다. 이를 위해 특정 레이블에 모델을 맞추는 지도학습(Supervised Learning)이 아닌, 데이터 자체의 숨겨진 구조를 발견하는 비지도학습(Unsupervised Learning) 방법론을 채택하였다. 핵심 알고리즘으로는 K-Means 클러스터링을 사용하였으며, 모델 설계는 '고유 특징 정의 및 추출', '최적의 군집 개수 탐색', '결과 해석을 위한 시각화'의 세 단계로 구성된다.

3.3.1. 핵심 특징 벡터(Feature Vector) 설계

정확한 클러스터링을 위해 이미지로부터 개인의 색상 특징을 정량화된 벡터로 변환하는 과정이 필수적이다. 본 모델은 퍼스널 컬러 진단에서 가장 중요한 요소가 '피부색'이라는 점에 착안하여, 다음과 같이 피부색 고유의 특징 벡터를 추출하도록 설계되었다:

1. 분석 영역 한정(Skin Segmentation)

- facer 라이브러리의 고성능 얼굴 인식(retinaface) 및 파싱(farl/celebm) 모델을 활용하여 입력된 이미지에서 얼굴 영역을 탐지하고, 이를 다시 피부, 눈썹, 눈, 코, 입 등의 세부 영역으로 분할한다.
- 설계의 핵심은 분할된 영역 중 오직 피부(skin)에 해당하는 픽셀만을 분석 대상으로 한정하는 것이다. 이를 통해 머리카락, 배경, 의상 등 퍼스널 컬러 분석의 노이즈가 될 수 있는 요소를 원천적으로 배제하여 모델의 정확도를 높였다.

2. 색상 공간 변환 (RGB to CIELAB)

- 추출된 피부 픽셀들을 컴퓨터 비전 표준인 RGB 색상 공간에서 CIELAB(Lab) 색상 공간으로 변환한다.
- Lab 색상 공간은 인간의 시각적 인지와 유사하게 설계되어, 색상 간의 유클리드 거리(Euclidean Distance)가 실제 사람이 느끼는 색상 차이와 비례하는 장점을 가진다. 이는 거리 기반 클러스터링 알고리즘인 KMeans의 성능을 극대화하는 중요한 설계 요소이다.

3. 특징 벡터 생성 (Color Quantization)

- 한 개인의 피부 영역에서 추출된 수많은 Lab 픽셀 값들을 대표하는 소수의 핵심 색상을 추출하기 위해, 픽셀 자체에 K-Means 알고리즘($k=7$)을 적용하는 색상 양자화(Color Quantization)를 수행한다.
- 이를 통해 한 사람의 피부톤을 가장 잘 나타내는 7개의 대표 색상(클러스터의 중심값)을 얻는다.

-
- 최종적으로, 이 10 개의 대표 색상(각 색상은 L^* , a^* , b^* 3 개 채널)을 하나의 벡터 로 결합하여, 총 30 차원(10 colors \times 3 channels)의 고유 특징 벡터(Unique Feature Vector)를 생성한다. 이 벡터는 한 사람의 피부색을 나타내는 고유한 '지문 (Fingerprint)' 역할을 수행한다.

3.3.2. 최적 클러스터 개수(k) 탐색 설계

K-Means 알고리즘은 사전에 군집의 개수(k)를 지정해야 한다. 본 연구의 목표는 '자 연스러운' 그룹을 찾는 것이므로, 최적의 k 를 데이터 기반으로 결정하는 과정이 중요하다.

- **평가 지표:** 실루엣 점수(Silhouette Score)를 핵심 평가지표로 사용한다. 실루엣 점수 는 각 데이터 포인트가 자신이 속한 클러스터와 얼마나 유사하고(cohesion), 다른 클 러스터와는 얼마나 다른지(separation)를 종합적으로 나타내는 지표이다. 1 에 가까울 수록 클러스터링이 명확하게 잘 되었다고 평가할 수 있다.
- **탐색 범위:** k 를 4 부터 20 까지 변화시키며 각 경우에 대한 클러스터링을 모두 수 행하고, 해당 k 값의 실루엣 점수를 기록한다.
- **최종 후보 선정:** 가장 높은 실루엣 점수를 기록한 상위 3 개의 k 값을 최종 후보 로 선정하여, 각 경우에 대한 심층적인 비교 분석을 수행한다. 이는 단 하나의 정 답을 강요하지 않고, 데이터가 제시하는 여러 유의미한 그룹화 가능성을 다각도로 검토하 기 위함이다.

3.4. 모델학습

본 모델에서의 '학습'은 전통적인 딥러닝 모델의 가중치 학습과는 다른 개념으로, KMeans 알고리즘이 최적의 클러스터 중심(Centroid)을 찾아가는 반복적인 프로세스를 의 미한다. 학습 과정은 다음과 같이 진행된다.

1. 데이터 준비 및 특징 추출 (Data Preparation & Feature Extraction)

- ArmocromiaDataset 클래스를 통해 전체 이미지 데이터셋을 로드한다.
- DataLoader 가 이미지를 배치(batch) 단위로 모델에 전달하면, 3.3.1 에서 설계된 특징 추출 파이프라인이 각 이미지에 대해 30 차원의 피부색 특징 벡터를 계산한 다.
- 데이터셋의 모든 이미지에 대한 특징 추출이 완료되면, (N_{images} , 30) 크 기의 최종 특징 행렬(Feature Matrix)이 메모리에 구축된다.

2. 특징 정규화 (Feature Scaling)

- 구축된 특징 행렬 전체에 sklearn 의 StandardScaler 를 적용한다.
- 이는 30 개 특징 각각의 값 범위를 표준 정규 분포(평균 0, 표준편차 1)에 가깝 게 조정하는 과정으로, 모든 특징이 K-Means 의 거리 계산에 동등한 중요도로 기 여하도록 보장하는 필수적인 전처리 단계이다.

3. K-Means 클러스터링 알고리즘 수행

- 초기화 (Initialization):** 지정된 k 개의 클러스터 중심점을 kmeans++ 방식을 사용하여 데이터 공간 내에 초기 배치한다. kmeans++ 방식은 초기 중심점을 서로 멀리 떨어뜨려 수렴 속도를 높이고 지역 최적해(local optima)에 빠질 위험을 줄여준다.
- 할당 (Assignment Step):** 정규화된 각 데이터 포인트(개인의 특징 벡터)에 대해, 모든 클러스터 중심점까지의 유클리드 거리를 계산하고 가장 가까운 중심점이 속한 클러스터에 해당 데이터를 할당한다.
- 업데이트 (Update Step):** 각 클러스터에 할당된 모든 데이터 포인트들의 평균값을 계산하여 새로운 클러스터 중심점으로 업데이트한다.
- 반복 및 수렴 (Iteration & Convergence):** 클러스터 중심점의 위치에 더 이상 큰 변화가 없을 때까지 할당(b)과 업데이트(c) 단계를 반복한다. 알고리즘이 수렴하면 최종 클러스터 구성이 확정되며, 이것이 하나의 '학습된 모델'이 된다.

최적 k 탐색을 위한 반복 학습

- 위의 3 단계 학습 과정은 설정된 k 범위(4~20) 내의 모든 k 값에 대해 독립적으로 반복 수행된다.
- 각 k 값에 대한 학습이 완료될 때마다 결과물(클러스터 할당 라벨)을 사용하여 실루엣 점수를 계산하고 기록한다.
- 모든 k 에 대한 학습 및 평가가 완료되면, 기록된 실루엣 점수를 기반으로 가장 성능이 우수한 상위 3 개의 k 값과 그에 해당하는 학습된 모델(클러스터 중심 점 및 할당 결과)을 최종 분석 대상으로 선정한다.

3.5. K 값 선정 및 클러스터링 평가

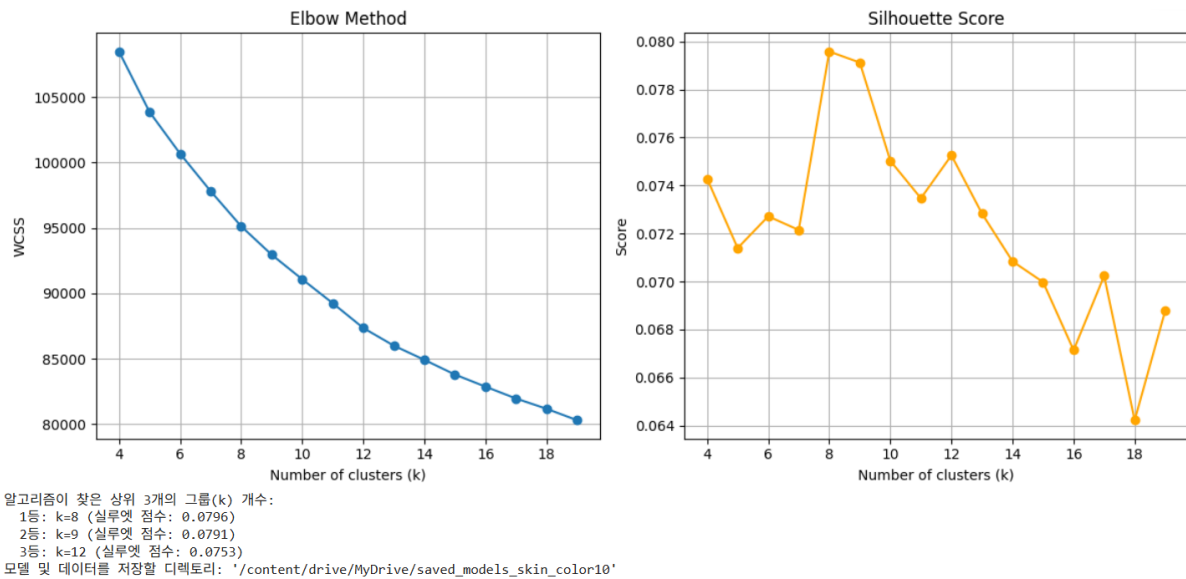


그림 3-3: Elbow Method 및 Silhouette Score의 시각화 결과

K-Means 클러스터링에서 최적의 클러스터 수(K)를 결정하기 위해, Elbow Method와 Silhouette Score를 활용하여 다양한 K 값에서의 클러스터링 성능을 평가하였다.

- **Elbow Method**는 클러스터 내 제곱합(WCSS)을 K 값에 따라 시각화하여, 클러스터 수를 늘려도 개선 효과가 급격히 감소하는 지점을 찾는다.
- **Silhouette Score**는 각 샘플이 속한 클러스터와 다른 클러스터 간의 유사도를 비교하여 군집의 명확성을 평가하며, 값이 클수록 군집이 잘 정의됨을 의미한다.

두 지표를 종합적으로 분석하여, 상위 3개의 K 값(실루엣 점수 기준)을 선정하고 최종적으로 **K=8**을 최적 값으로 결정하였다. 이 K 값은 이후 클러스터 시각화와 분석에 활용되었다.

3.5.1. 주성분 분석(PCA)을 활용한 클러스터링 결과 시각화

본 연구의 특징 추출 과정에서는 각 이미지에 대해 30 차원 특징 벡터(10 개의 대표 색상 × 3 개의 Lab 값)를 생성하였다. 이러한 고차원 표현은 30 차원 공간에서 유클리드 거리를 계산하는 K-평균 알고리즘에는 효과적이지만, 사람이 직접적으로 해석하거나 시각화하기에는 큰 한계를 가진다.

발견된 클러스터의 분리를 시각적으로 검증하기 위해 차원 축소 기법을 적용하였다. 본 연구에서는 고차원 데이터를 최대한 많은 정보(분산)를 유지한 채 저차원 공간으로 변환할 수 있는 강력한 통계 기법인 **주성분 분석(Principal Component Analysis, PCA)**을 활용하였다. PCA는 새로운 비상관 변수인 주성분(Principal Component)을 생성하여 데이터를 변환하는 방식으로, 제 1 주성분(PC1)은 데이터에서 가장 큰 분산을 설명하며, 제 2 주성분(PC2)은 그다음으로 큰 분산을 설명하는 방식으로 이어진다.

개념적으로 PCA는 30 차원 데이터를 2 차원 또는 3 차원 평면으로 투영했을 때 가장 유의미한 "그림자(shadow)" 혹은 투영(projection)을 찾아낸다. 본 분석에서는 30 차원을 가장 중요한 세 개의 주성분으로 축소하였으며, 이를 통해 각 이미지의 특징 벡터를 3 차원 공간 상의 점으로 표현할 수 있었다.

시각화 결과, 대부분의 클러스터는 비교적 잘 분리되어 독립적인 그룹을 형성하는 경향을 보였다. 그러나 **클러스터 2**는 명확한 중심점을 가지지 않았으며, 높은 분산을 나타내고 다른 클러스터와 중첩되는 양상을 보였다. 이러한 특성으로 인해 클러스터 2는 이상치(outlier)로 간주될 수 있으며, 후속 분석에서는 이 클러스터를 재검토하거나 제외하는 방안을 신중히 고려해야 한다.

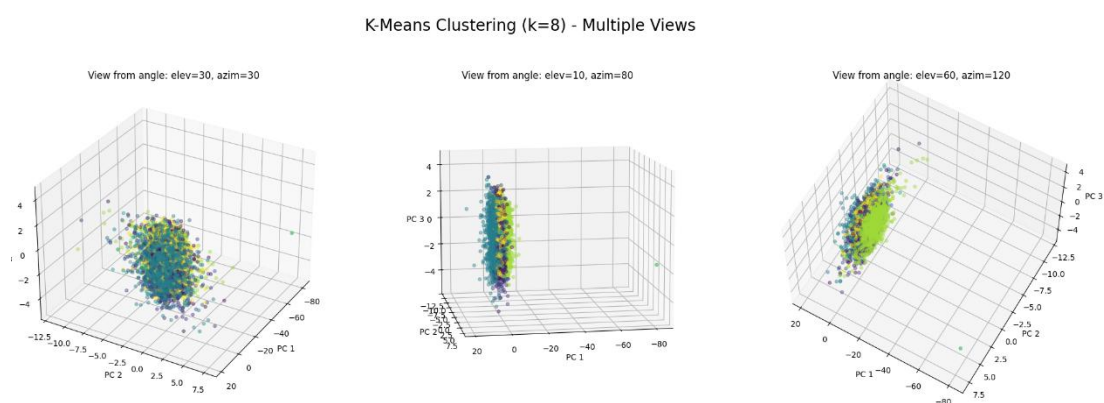


그림 3-4: 주성분 3개를 활용한 K-평균 클러스터(k=8)의 3차원 시각화

각 점은 이미지를 나타내며, 클러스터 할당에 따라 색으로 구분된다. 세 개의 플롯은 동일한 데이터를 서로 다른 관찰 각도((30, 30), (10, 80), (60, 120))에서 보여줌으로써 클러스터 분리를 종합적으로 확인할 수 있도록 한다.

3.6. 클러스터 분석 및 팔레트 설계

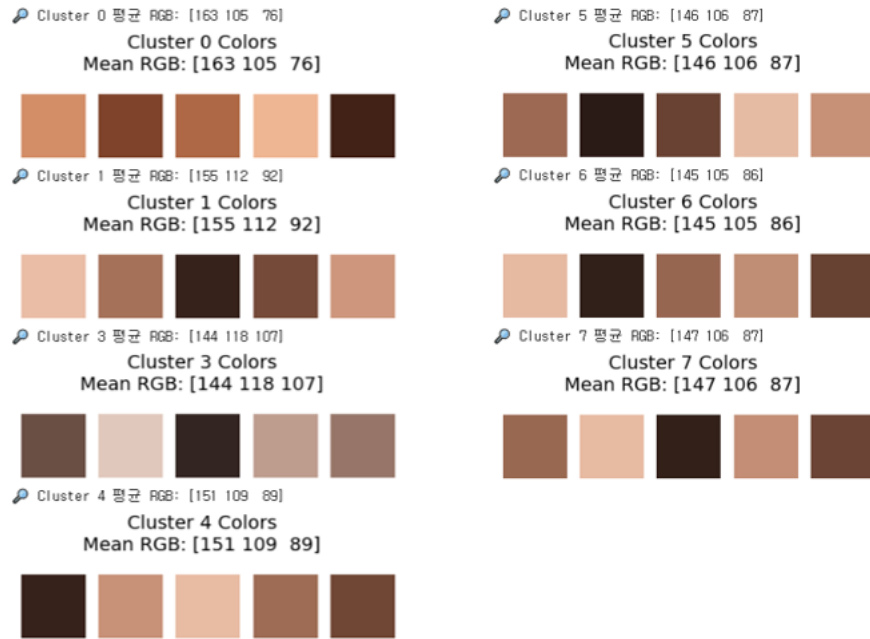
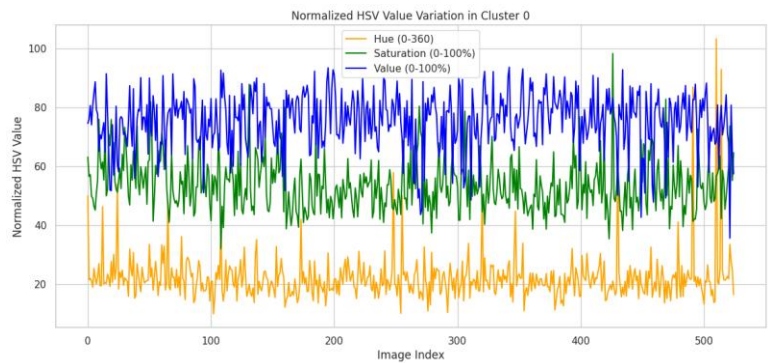


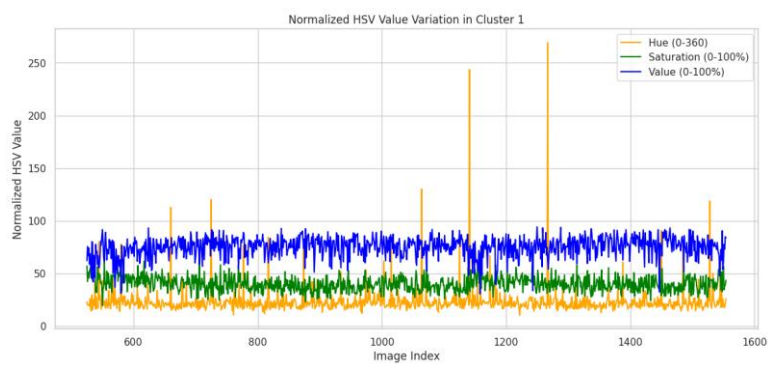
그림 3-5: 각 클러스터의 평균 RGB 값을 출력한 결과.

3.6.1. 클러스터별 대표 이미지 시각화

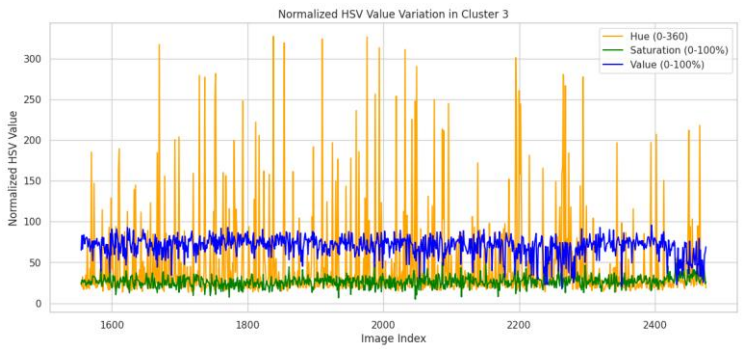
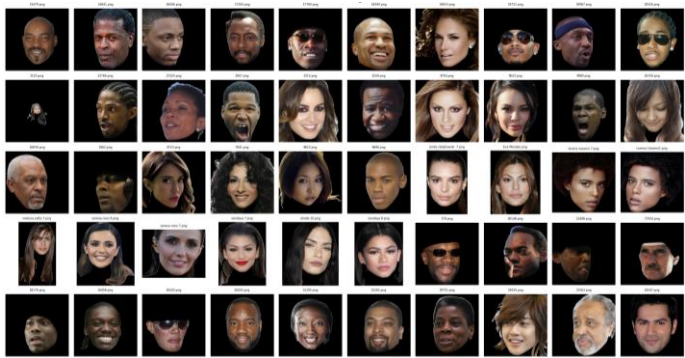
Cluster 0



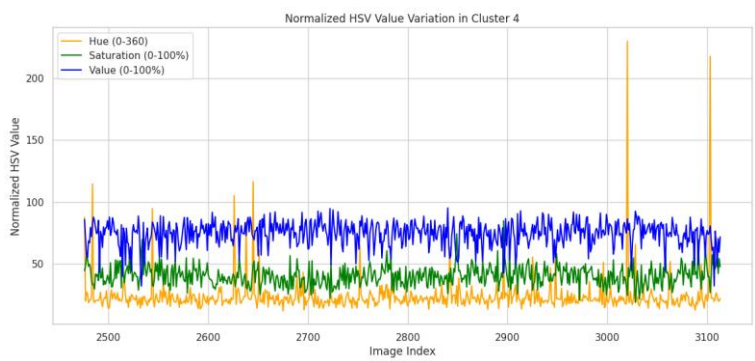
Cluster 1



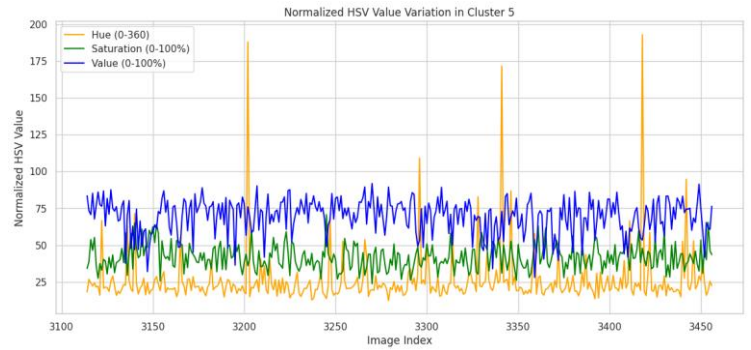
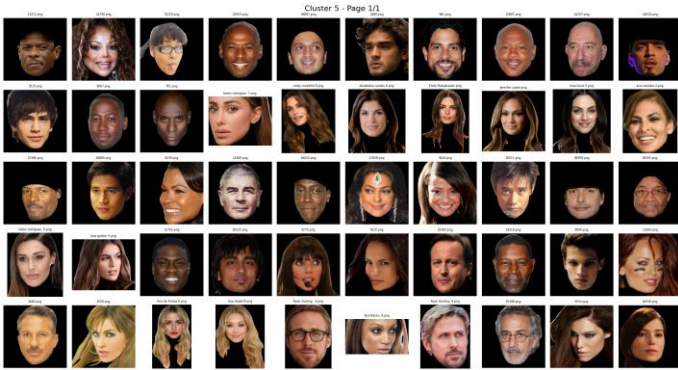
Cluster 3



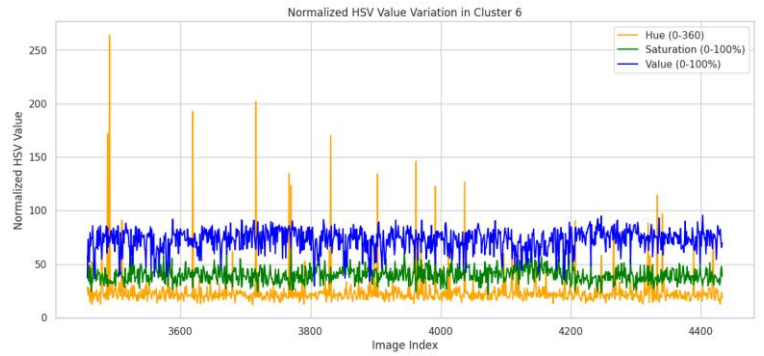
Cluster 4



Cluster 5



Cluster 6



Cluster 7

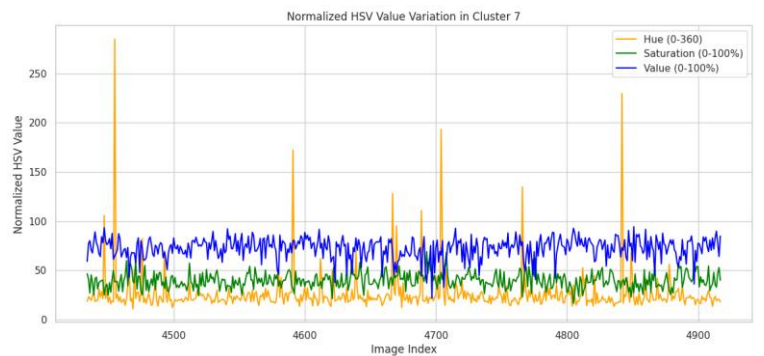


그림 3-6: 각 클러스터별로 이미지 그리드와 정규화된 HSV 값 분포 그래프를 나란히 (side by side) 배치하여 시각화. 이미지 그리드는 해당 클러스터에 속하는 대표 이미지를 4x5 형태로 보여주며, 파일 이름을 함께 표시. 그래프는 클러스터 내 이미지들의 정규화된 H, S, V 값을 각각 다른 색으로 나타내어 색상 특성을 직관적으로 확인할 수 있다.

도록 구성.

이 정보는 클러스터별 특성 분석 및 시각적 이름 지정에 위한 기초 자료로 활용됨.

3.6.2. HSV/HSB 색상 공간 분석

HSV는 색상(Hue), 채도(Saturation), 명도(Value)로 구성된 색상 표현 방식으로, 인간이 인식하는 색상 구조를 보다 직관적으로 표현할 수 있는 색상 공간이다.

- **H (Hue)** : 색상의 종류를 나타내며 0 도에서 360 도 범위로 표현된다. 0 도는 빨간색, 120 도는 초록색, 240 도는 파란색을 의미한다.
- **S (Saturation)** : 색의 선명함 또는 채도를 의미하며, 값이 높을수록 선명한 색을 나타낸다.
- **V/B (Value/Brightness)** : 색의 밝기를 의미하며, 값이 높을수록 밝은 색을 나타낸다.

본 보고서에서는 OpenCV의 HSV 범위를 고려하여, H(0~179), S(0~255), V(0~255)로 추출한 후, 시각화를 위해 H는 0~360, S와 V는 0~100%로 정규화하여 사용하였다.

HSV 기반 색상 분석을 통해 클러스터별 주요 색상 특성을 직관적으로 파악할 수 있도록 하였고, 클러스터 시각적 이름 지정에 활용하였다.

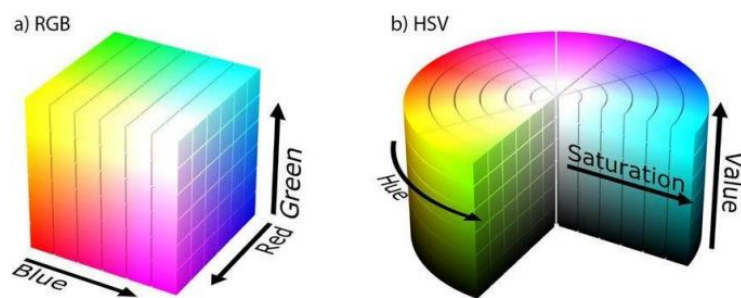


그림 3-7: RGB 색상 공간과 HSV 색상 공간 비교

3.6.3. 클러스터 라벨링 및 초기 팔레트 설계

본 프로젝트에서는 클러스터링된 피부 톤 데이터를 바탕으로 각 그룹의 평균 HSV, Lab, RGB 값을 종합 분석하여, 시각적 인상에 부합하는 팔레트 이름을 부여하였다. 또한, 색조 조화를 고려하여 각 클러스터에 어울리는 색상 조합을 함께 제안한다.

Cluster	시각적 이름 (Visual Name)	주요 특징 요약
0	Golden	가장 높은 평균 색상(Hue)과 채도(Saturation). 명확하게 따뜻한 톤.
1	Warm Beige	전체적으로 따뜻하며, 약간 올리브빛 언더톤 존재.
3	Muted Clay	가장 낮은 채도. 차분하고 은은한 색상.
4	Warm Apricot	전반적으로 오렌지 계열 색상. 안정적이고 따뜻한 이미지.
5	Peachy Pink	빨강-분홍 계열 변동 존재. 다소 사랑스럽고 생기 있는 톤.
6	Honey Buff	따뜻하고 달콤한 꿀 같은 골든 베이지 계열.
7	Beige Rose	클러스터 6과 유사하지만, 약간 더 부드러운 느낌.

Hue와 Value(밝기)를 조절하여 해당 클러스터와 조화를 이루는 팔레트를 제작후, 최종 팔레트를 완성하기 위해 적합한 색상을 수동으로 선별하였다

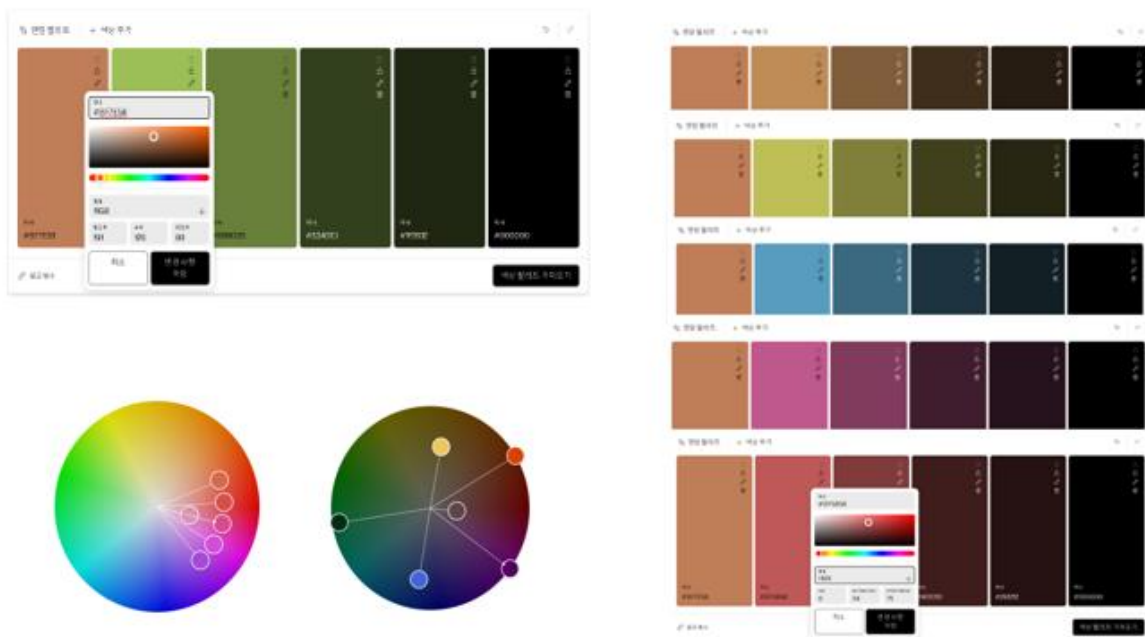


그림 3-8: 팔레트 생성 예시



그림 3-9: 피치 핑크 타입 팔레트

3.7. 컬러 오버레이를 위한 사전 학습 모델 설치

본 프로젝트에서는 BiSeNet 모델을 활용하였다. BiSeNet 은 경량화된 딥러닝 기반의 세분화(Segmentation) 모델로, 실시간 처리에 적합하다는 장점을 지닌다. 또한 CelebAMask-HQ 데이터셋으로 학습되어 얼굴을 19 개의 세부 영역으로 분할할 수 있으며, 이를 통해 정밀한 얼굴 파싱이 가능하다.

주요 사항

- 모델: BiSeNet
- 학습 데이터: CelebAMask-HQ
- 클래스 수: 19 개
- 사용한 사전 학습 모델: 79999_iter.pth

Label Name	Class Index
Hair	17
Upper Lip	12
Lower Lip	13
Left Eye	4
Right Eye	5
Clothes	16

3.7.1. 메이크업 모듈 및 컬러 오버레이 개선

가상 스타일링을 위해 머리카락, 입술 등 특정 얼굴 부위에 색상을 적용하는 메이크업 모듈을 구현하였다. 이 모듈은 얼굴 세분화 결과를 활용하여 각 부위를 인식하고, 해당 영역에만 선택적으로 색상을 입히는 방식으로 동작한다.

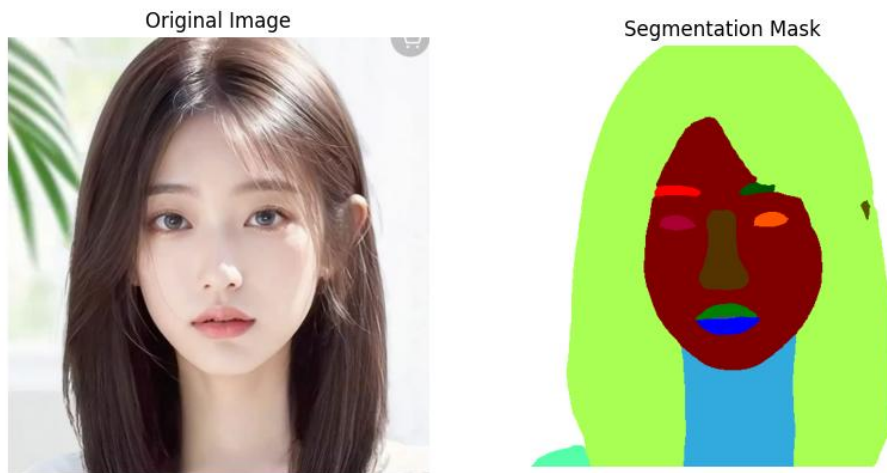


그림 3-10: 입력 이미지 및 입력 이미지에 대한 세분화 마스크 적용

기존의 사전 학습 모델이 제공하는 단순한 색상 오버레이 방식은 특히 검은 머리와 같이 어두운 영역을 변환할 때 색감이 부자연스럽고 평면적인 결과를 초래하는 한계가 있었다. 이를 보완하기 위해 본 연구에서는 소프트 라이트 블렌딩 기법을 추가로 적용하였다. 이 방법은 원본 이미지의 질감, 음영, 하이라이트를 유지하면서 색상을 자연스럽게 합성할 수 있어 보다 사실적인 결과를 제공한다. 또한 머리카락 영역에는 선명도를 강화하기 위한 샤프닝을 적용하여 디테일을 개선하였다.



그림 3-11: 입력 이미지 및 컬러 오버레이 적용 결과 이미지

이와 같은 개선을 통해 본 시스템은 머리 색상 변화나 립 컬러 적용과 같은 메이크업 시뮬레이션을 실제와 유사하게 구현할 수 있었으며, 퍼스널 컬러 진단 및 가상 스타일링 서비스의 시각적 품질과 사용자 만족도를 향상시킬 수 있었다.

3.8. 웹 설계 및 구현

웹 애플리케이션을 구현하기 위해 **Python, JavaScript, HTML, CSS**, 그리고 **Flask** 프레임워크를 활용하였다. 학습된 머신러닝 모델은 **.joblib** 형식으로 저장하여 웹 환경에서 불러올 수 있도록 하였으며, 이용자는 Flask 기반 인터페이스를 통해 이미지를 업로드할 수 있다. 업로드된 이미지는 전처리 과정을 거친 후 학습된 모델을 이용하여 사용자의 퍼스널 컬러를 예측하도록 설계하였다. 또한 사용자 관리와 개인화 기능을 지원하기 위해 **Firebase** 데이터베이스를 연동하였다. 이를 통해 사용자 이름, 이메일, 성별 정보와 같은 핵심 데이터를 안전하게 저장할 수 있으며, 사용자별 맞춤형 스타일 추천과 세션 연속성을 제공할 수 있게 하였다.

3.8.1. 기능 정의서

구분	요구사항명	기능명	상세 설명	데이터 (*는 필수)
사용자 인증	회원가입	사용자 정보 등록	사용자는 이름, 이메일, 비밀번호, 성별을 입력하여 회원가입을 할 수 있습니다.	*이름, *이메일, *비밀번호, *성별
		아이디/이메일 중복 확인	이미 가입된 아이디(이름)나 이메일이 있는지 확인합니다.	이름, 이메일
	로그인	자체 로그인	사용자는 가입 시 사용한 아이디(이름) 또는 이메일과 비밀번호로 로그인할 수 있습니다.	*아이디(이름) 또는 이메일, *비밀번호
	로그아웃	세션 종료	사용자의 로그인 상태를 종료하고 세션을 초기화합니다.	
	프로필 조회	내 정보 확인	로그인된 사용자는 자신의 프로필 정보(이름, 이메일, 성별 등)를 조회할 수 있습니다.	

퍼스널 컬러분석	이미지 업로드	파일에서 사진 선택	사용자는 로컬 컴퓨터에 저장된 이미지 파일을 업로드하여 분석을 요청할 수 있습니다.	*이미지 파일 (png, jpg, jpeg, webp)
		카메라로 사진 촬영	사용자는 웹캠을 이용하여 실시간으로 사진을 촬영하고 업로드할 수 있습니다.	*카메라 이미지 데이터
	AI 분석	얼굴 및 피부 영역 탐지	업로드된 이미지에서 AI 모델이 얼굴과 피부 영역을 자동으로 탐지합니다.	이미지 데이터
		조명 보정	이미지의 조명 상태(과다 노출, 노출 부족, 낮은 대비 등)를 분석하고 자동으로 보정하여 분석 정확도를 높입니다.	이미지 데이터
		퍼스널 컬러 진단	AI 모델(K-means)이 추출된 피부색 특징을 분석하여 8가지 퍼스널 컬러 타입 중 하나로 진단합니다.	피부 영역 색상 특징 (Lab)
	결과 확인	분석 결과 표시	진단된 퍼스널 컬러 타입(예: 골든 타입, 윈 베이지 타입 등), 상세 설명, 추천 컬러 팔레트를 화면에 표시합니다.	진단 결과, 추천 컬러 팔레트
		PDF 리포트 다운로드	분석 결과(원본 사진, 추천 스타일링 사진 포함)를 PDF 파일 형태로 다운로드할 수 있는 기능을 제공합니다.	원본/결과 이미지, 진단 결과
가상 스타일링 (메이크업)	가상 스타일링	추천 스타일링 적용	분석된 퍼스널 컬러에 맞는 여러 가지 추천 스타일(헤어, 렌즈, 립, 의상 컬러)을 원본 이미지에 적용하여 보여줍니다.	원본 이미지, 퍼스널 컬러 타입, 추천 팔레트 번호
		맞춤형 스타일링	사용자가 직접 헤어, 립스틱, 렌즈, 의상 색상을 선택하여 원본 이미지에 실시간으로 적용해볼 수 있습니다.	원본 이미지, 사용자가 선택한 색상 값

	개발자 도구	개발자용 테스트	특정 권한을 가진 개발자가 메이크업 기능을 테스트할 수 있는 별도의 페이지를 제공합니다.	개발자 계정
--	--------	----------	---	--------

3.8.2. 결과물

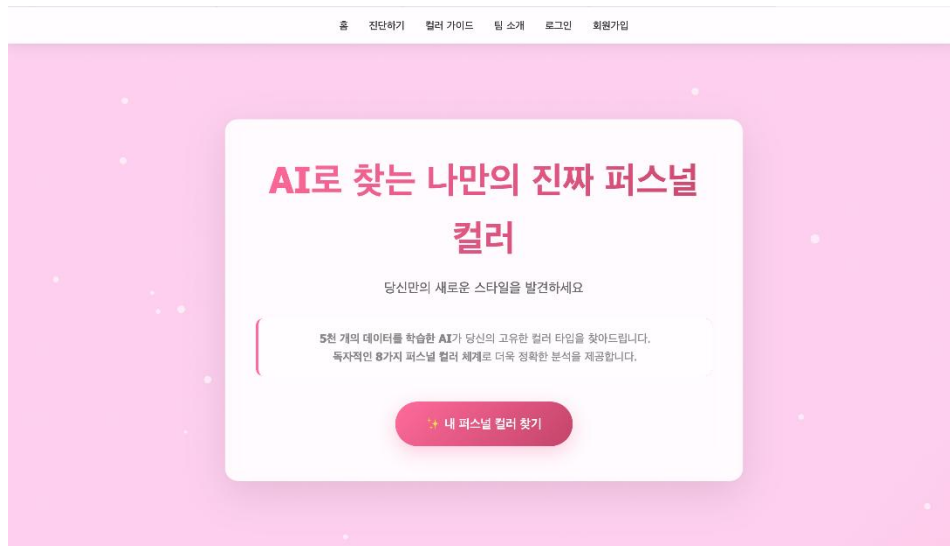


그림 3-12: 메인 페이지

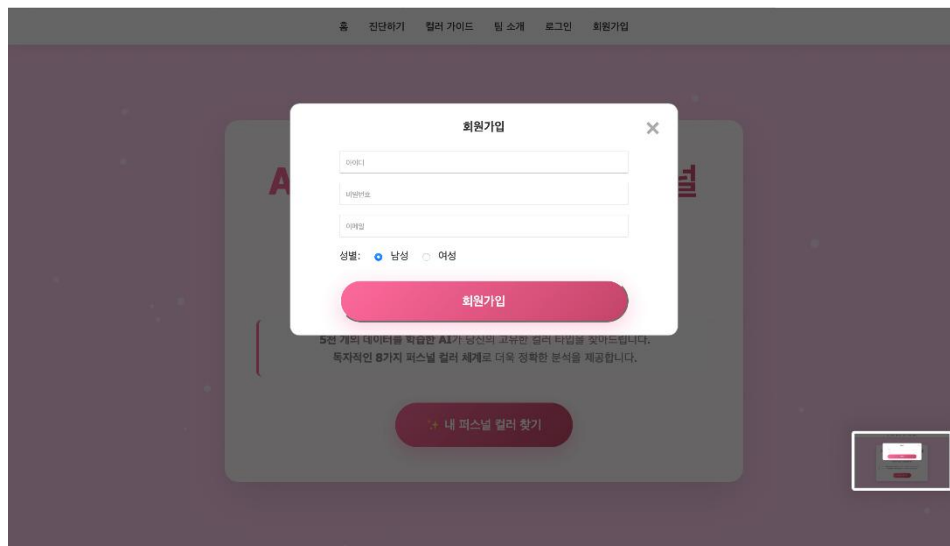


그림 3-13: 회원가입 화면

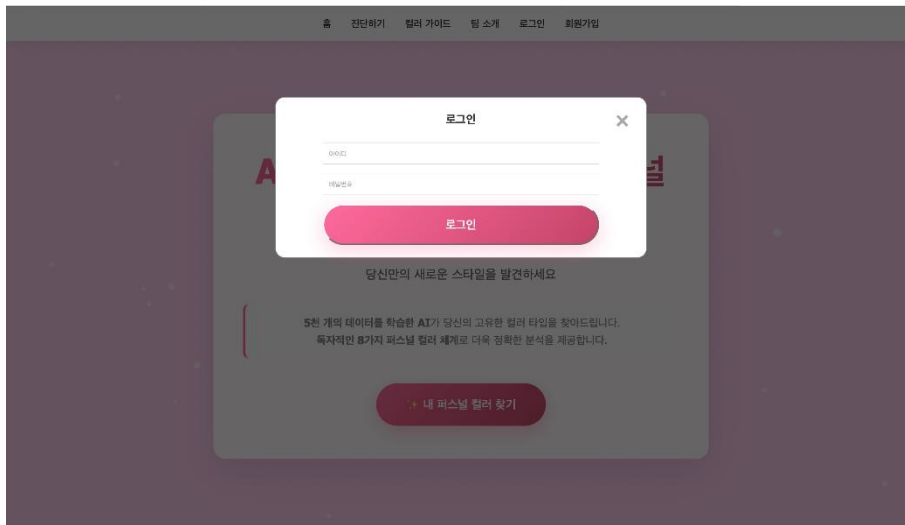


그림 3-14: 로그인 화면

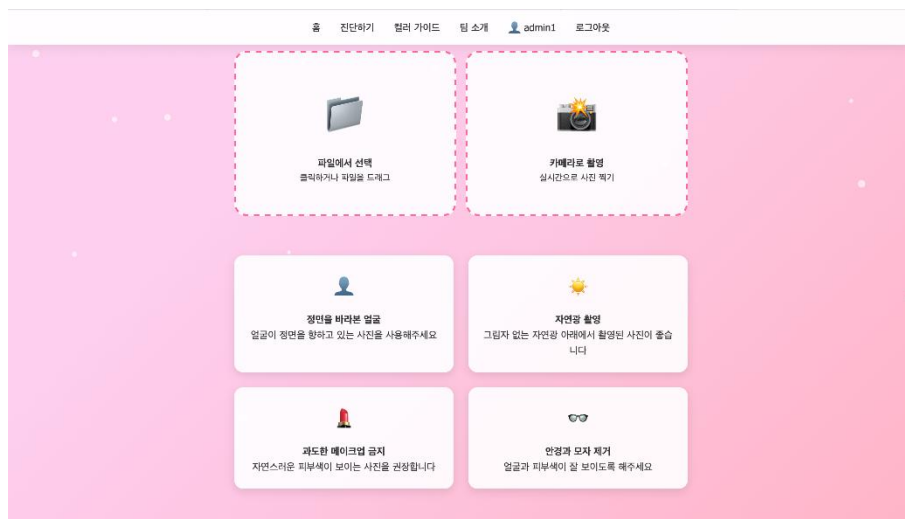


그림 3-15: 사진 업로드 및 촬영 안내 화면

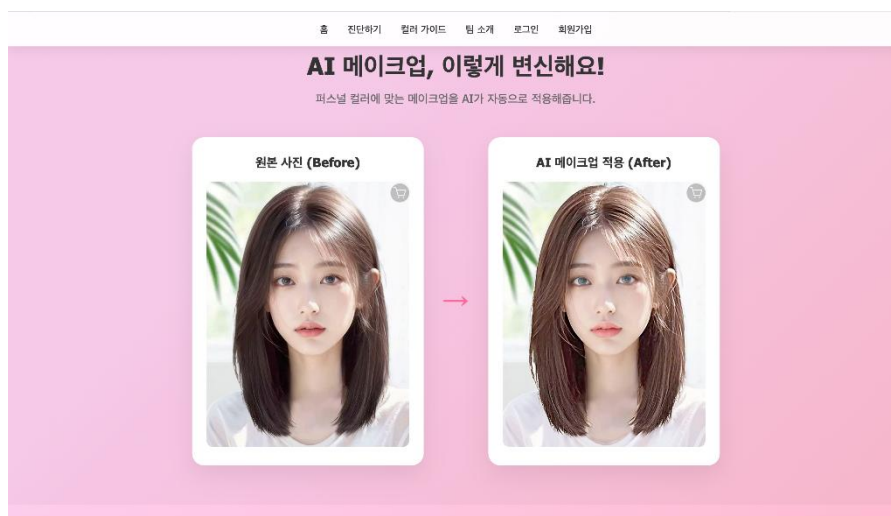


그림 3-16: AI 메이크업 적용 예시 화면

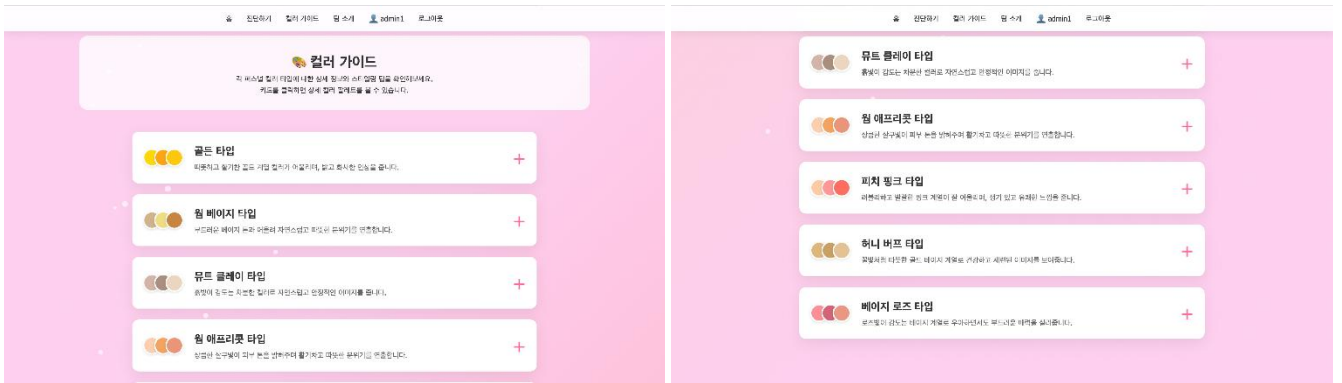


그림 3-17: 퍼스널 컬러 가이드 화면

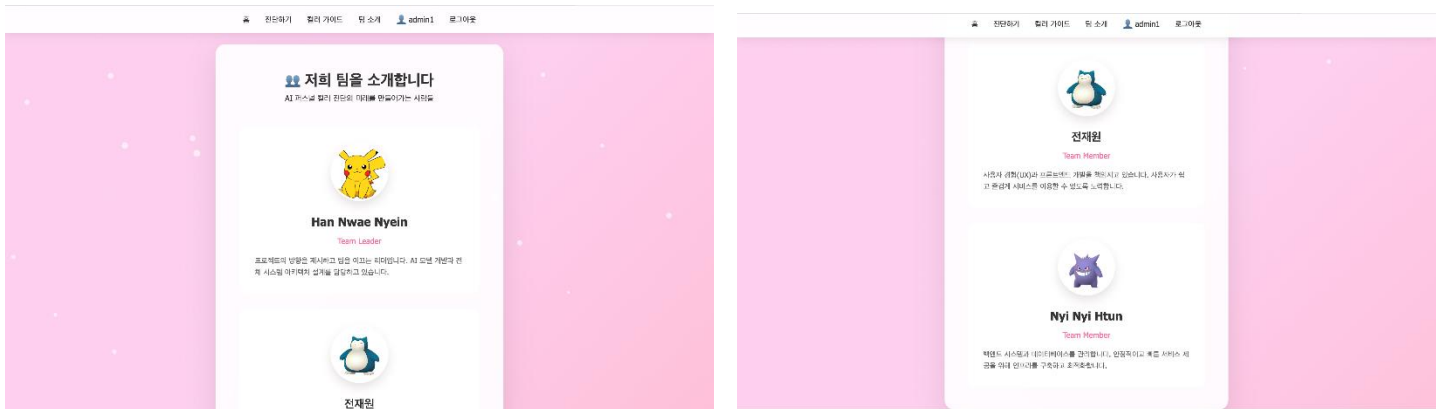


그림 3-18: 팀 소개 화면

4. 연구 결과 분석 및 평가

본 연구에서는 피부 톤 분석에 기반한 AI 퍼스널 뷰티 어드바이저를 개발하여 기존 아르모크로미아 체계의 한계를 극복하고자 하였다. 연구 결과는 다음과 같이 요약 및 평가할 수 있다.

1. 정확한 피부 톤 분석

*Facer*를 활용한 얼굴 검출 및 세분화를 통해 피부 영역을 정밀하게 추출할 수 있었다. 또한 Lab 색공간 변환과 클러스터링을 적용함으로써 대표적인 피부 톤 특징을 안정적으로 확보하였다. 다만, 조명 조건에 따른 변동성이 일부 존재하여 히스토그램 평활화와 화이트 밸런스 보정과 같은 전처리 기법의 추가적 보완이 필요함을 확인하였다.

2. 새로운 퍼스널 컬러 프레임워크 확립

기존의 12유형 체계 대신, 본 연구에서는 *K-Means 클러스터링*과 엘보우 기법 및 실루엣 점수를 적용하여 5,000장의 얼굴 이미지 데이터셋에서 최적의 그룹 수를 탐색하였다. 그 결과, $k=8$, $k=12$, $k=15$ 등 다양한 클러스터 수에서 의미 있는 분류가 도출되었으며, 이는 퍼스널 컬러가 보다 유연하고 데이터 기반으로 정의될 수 있음을 보여준다.

3. 컬러 가이드 제공

각 클러스터별 중심 색상 값을 기반으로 대표 팔레트를 도출하였으며, 이는 사용자에게 참고 가능한 컬러 가이드로 활용 가능성을 입증하였다. 나아가 이를 인터랙티브 컬러 카드 형태로 확장할 수 있는 가능성도 확인하였다.

4. 스타일 시뮬레이션 및 시각화

구현된 *메이크오버* 기능을 통해 추천된 색상을 립스틱, 헤어, 렌즈, 의상 등에 가상 적용할 수 있었다. 이를 통해 추상적인 추천이 아닌, 직관적으로 체험 가능한 결과를 제공하여 실용성과 사용자 만족도를 높였다.

5. 평가 및 한계점

본 시스템은 데이터 기반의 유연하고 시각적으로 확인 가능한 결과를 제시함으로써 기존 퍼스널 컬러 진단의 한계를 효과적으로 개선하였다. 그러나 다음과 같은 과제가 남아 있다:

- 실제 환경 적용을 위한 조명 보정 및 표준화 강화 필요
- 다양한 인종과 환경을 포괄할 수 있는 대규모 데이터셋 확보 필요
- 추천 엔진과 실시간 웹 서비스로의 확장 필요

종합적으로, 본 연구는 AI 기반 프레임워크를 통해 퍼스널 컬러 분석의 새로운 가능성을 제시하였으며, 향후 정확성·강건성·사용자 중심 기능 강화를 통해 더욱 발전할 수 있음을 확인하였다.

5. 결론 및 향후 연구 방향

5.1. 결론

본 프로젝트는 기존 퍼스널 컬러 진단의 한계를 넘어서는 **AI 기반 퍼스널 뷰티 어드바이저**를 제안하고 구현하였다. 기존 방식이 전문가의 주관적 판단이나 염색·렌즈 등으로 쉽게 변하는 머리색, 눈동자 색상에 의존한 것과 달리, 본 시스템은 **컴퓨터 비전 기술을 활용한 피부 톤 분석**에 집중하였다. 또한 히스토그램 평활화, 화이트 밸런스 보정 등 이미지 전처리 기법과 대규모 얼굴 데이터셋 기반의 클러스터링을 통해 **데이터 기반의 새로운 퍼스널 컬러 체계**를 구축하였다.

구현된 플랫폼은 립, 헤어, 렌즈, 의류에 최적화된 색상을 추천할 뿐만 아니라, **가상 시뮬레이션 기능**을 제공하여 사용자가 직관적으로 개인 맞춤형 스타일을 체험할 수 있도록 한다. 이를 통해 본 연구는 **퍼스널 컬러 진단의 정확성과 접근성을 동시에 향상시킬 수 있음을 입증**하였으며, 전문가 의존성을 낮추고 일반 사용자도 쉽게 활용할 수 있는 가능성을 제시하였다.

궁극적으로 본 연구는 **데이터 기반의 객관성 + 사용자 중심의 유연성 + 실제감 있는 시각화**를 결합한 차세대 퍼스널 뷰티 플랫폼의 기반을 마련하였다.

5.2. 향후 연구 방향

본 프로젝트의 기반을 토대로, 향후에는 보다 포괄적이고 지능적인 스타일 어드바이저 시스템으로 확장하고자 한다. 구체적인 연구 방향은 다음과 같다:

- **세분화된 피부 톤 클러스터링:**
기존 퍼스널 컬러 클러스터를 **피부 톤 기반으로 더 세분화**하여 파운데이션, 메이크업 등 화장품 색상까지 정밀하게 추천할 수 있도록 한다.
- **보색 기반 추천:**
피부 톤과 대비되는 색상을 분석하여 생기를 더해주는 포인트 컬러를 찾아내고, 의상이나 액세서리 추천에 활용한다.
- **조화색 팔레트 제공:**
피부 톤과 자연스럽게 어울리는 색상 조합을 제시하여 세련되고 조화로운 스타일 연출을 지원한다.
- **톤업 / 톤다운 컬러 추천:**

상황과 목적(예: 데일리, 포멀, 무드)에 맞게 명도와 채도를 조절한 컬러를 추천할 수 있도록 기능을 확장한다.

- **상황·스타일별 추천:**

계절, 장소, 이벤트(예: 데이트, 면접, 발표)에 적합한 색상 및 스타일 조합을 제안하여 실용성을 높인다.

- **컬러 조합 전략:**

단일 색상 추천을 넘어, 조화를 이루는 보조 색상을 함께 제시하여 전체적인 스타일링 완성도를 높인다.

- **의류 및 헤어스타일 통합:**

색상 추천에 머무르지 않고 의류 스타일과 헤어스타일까지 고려한 종합적인 스타일링 경험을 제공한다.

세분화된 피부 톤 분석을 강조함으로써, 본 연구는 **기존 퍼스널 컬러 시스템을 넘어서, 개인 맞춤형 데이터 기반 뷰티 추천**을 가능하게 한다. 상황별 스타일링 및 가상 시뮬레이션과 결합하면, 이 플랫폼은 **실용성과 개인 맞춤성을 갖춘 AI 통합 뷰티 어드바이저**로 발전할 수 있을 것이다.

6. 팀원별 역할 및 개발 일정

6.1. 팀원별 역할

학번	성명	팀원별 역할
202155659	Han Nwae Nyein	프로젝트 전체 시스템 아키텍처와 웹 애플리케이션 설계 방향을 정의함. 얼굴 이미지에서 피부색 특징을 추출하고, 소규모 K-Means 클러스터링을 적용하여 개인별 대표 피부색 벡터를 생성한 뒤, 메인 클러스터링 모델과 통합하여 사용자 맞춤형 색상 추천 시스템을 개발함. 색상 팔레트 설계 및 알고리즘 최적화 등을 담당하여 시스템의 핵심 기능을 완성함.
202155658	Nyi Nyi Htun	데이터베이스 및 서버 연동을 담당하여 사용자 데이터 관리와 시스템 안정성을 구현함. Flask와 연계된 데이터베이스 구조를 설계하고 사용자 입력, 추천 결과, PDF 생성 기록 등

		을 효율적으로 저장할 수 있도록 구축함. 또한, API 연동 및 서버 측 로직 최적화를 수행하여 웹 애플리케이션이 원활하게 동작하도록 기여함.
202055594	전재원	퍼스널 컬러 클러스터링 모델을 개발하고 이를 웹 애플리케이션과 통합함. K-평균 클러스터링(K-Means)과 엘보우 기법(Elbow Method)을 활용하여 최적의 클러스터 수를 도출하고, 각 클러스터의 대표 색상 팔레트를 정의함. 또한, 사용자 인터페이스 구현, 색상 시각화 및 시스템 통합 테스트를 수행하여 프로젝트 전반의 완성도를 높임.
	공동 역할	모든 팀원은 프로젝트 문서화(착수보고서, 중간보고서, 최종보고서) 작성과 시스템 통합 테스트에 함께 참여하여 프로젝트의 완성도와 안정성을 높이는 데 기여함.

6.2. 개발 일정

작업	5				6				7				8				9			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
자료 조사																				
착수보고서																				
색채 이론 연구 및 학습																				
색상 분석 모델 개발																				
이미지 세그멘테이션																				
다중 라이브러리 얼굴 검출 시스템																				
K-Means 클러스터링 최적화																				
중간보고서 작성																				
UI/UX 디자인																				
새로운 퍼스널 컬러 분류 모델																				
스타일 추천 시스템 구현																				
시스템 통합 및 최적화																				
성능 검증 및 테스트																				
통합 테스트 및 버그 수정																				
최종보고서 작성																				
발표 준비																				

7. 참고 문헌

Zhuang, L. (n.d.). BiSeNet: Real-time semantic segmentation for face parsing [Computer software]. GitHub. <https://github.com/zllrunning/face-parsing.PyTorch>

Stacchio, L. (n.d.). Deep-Armocromia: Face dataset for personal color analysis [Data set]. GitHub. <https://github.com/lorenzo-stacchio/Deep-Armocromia>

Figma. (2025). Explore 1000+ color palettes. <https://www.figma.com/color-palettes/>

Klook. (2024, May 24). A guide to personal color analysis in South Korea. <https://www.klook.com/blog/korea-color-analysis/>

YouTube, "I finally cleared doubts about my personal color ," YouTube, 2025. [Online]. Available: <https://www.youtube.com/shorts/zDwEil8QUU8>.