

UML 기반 딥러닝과 생성형 AI 융합 얼굴 특징 맞춤형 동물 타투 생성 시스템

정윤걸, 이정빈, 천효민, 김조은, 조희원, 이은서
국립경국대학교 컴퓨터공학과

ykjung0917@kakao.com, 2100238@naver.com, cjsgyals04@naver.com, joenkim200212@gmail.com,
gmldnjs1178@naver.com, eslee@gknu.ac.kr

A UML-Based Deep Learning and Generative AI System for Personalized Animal Tattoo Generation from Facial Features

Yun-Geol Jung, Jeong-Bin Lee, Hyo-Min Cheon, Jo-Eun Kim, Hui-Won Cho, Eun-Ser Lee
Dept. of Computer Science and Engineering Gyeongbuk National University

요 약

본 논문은 사용자의 얼굴 이미지를 입력 받아, 얼굴 특징 분석과 딥러닝 기반 얇은 동물 분류, 그리고 생성형 Artificial Intelligence(AI) 프롬프트 자동화를 결합한 맞춤형 타투 생성 웹 서비스를 구현하였다. MediaPipe 기반 얼굴 검출 및 특징 추출, Enhanced ResNet34 기반 동물 분류, 그리고 얼굴 특징을 반영한 자연어 프롬프트를 통해 자동 생성 파이프라인을 설계하여 시스템을 구현하였다.

1. 서론

생성형 인공지능의 발전은 텍스트 프롬프트 기반 이미지 생성, 개인화 콘텐츠 제작 등 다양한 분야에서 혁신을 이끌고 있다. 그러나 기존 생성형 AI 서비스는 사용자의 고유한 비정형적 시각적 특성을 효과적으로 반영하지 못하는 한계가 있다. 본 연구는 이러한 한계를 해결하기 위해 얼굴 이미지에서 추출한 정량적 특징과 딥러닝 기반 동물 분류 결과를 유기적으로 결합하여, 생성형 AI가 이해할 수 있는 자연어 프롬프트로 자동 변환하는 통합 파이프라인을 제안한다. 이를 통해 사용자의 얼굴과 얇은 동물 이미지를 기반으로 한 맞춤형 타투 디자인을 자동 생성하는 웹 서비스를 구현함으로써, AI 기반 개인화 콘텐츠의 실현 가능성을 제시한다.

2. 관련 연구

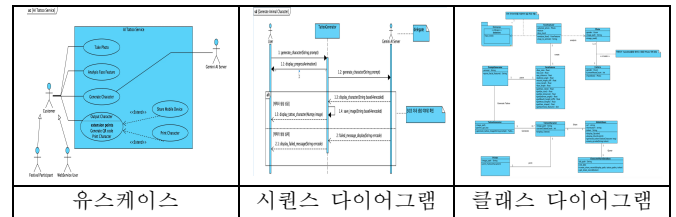
얼굴 특징 분석에는 MediaPipe, Dlib 등 다양한 라이브러리를 활용하여 얼굴 랜드마크 기반의 정량적 특징을 추출하는 연구가 활발하며, 이는 표정 인식 및 감정 분석 등 다양한 응용 분야의 기반이 된다. 한편, 동물상 분류는 주로 Convolutional Neural Network(CNN) 기반의 이미지 분류 모델이 사용되며, 최근에는 모델의 인식 정확도를 높이기 위해 Attention 메커니즘이나 Feature

Pyramid Network(FPN) 등 고도화된 구조가 도입되고 있다[1]. 생성형 AI 프롬프트 자동화 연구는 주로 텍스트-이미지 매핑, 자연어 처리 기반 프롬프트 엔지니어링에 집중되어 있으나, 사용자의 고유한 얼굴 특징을 정량적으로 분석하고 이를 생성형 AI 프롬프트를 직접적으로 연결하여 콘텐츠를 제작하는 연구는 드물다.

3. 실험 방법 및 실험 결과

3.1 전체 파이프라인 개요

본 시스템은 개발 초기 단계부터 Unified Modeling Language(UML)를 기반으로 체계적인 요구사항 분석 및 설계를 수행하였다. UML 다이어그램은 시스템의 구조적 명세와 동적 행위를 시각화 하여, 복잡한 딥러닝 및 생성형 AI 파이프라인의 통합 과정을 명확히 이해하고 구현하는 데 핵심적인 역할을 하였다.



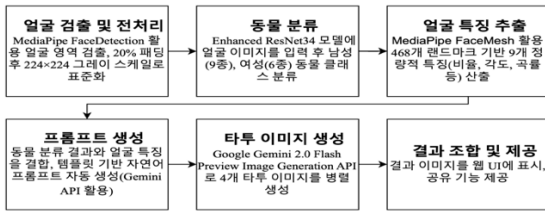
(그림 1) UML 다이어그램 시스템

교신저자: 이은서(eslee@gknu.ac.kr)

이 논문은 국립경국대학교 기본연구지원사업에 의하여 연구되었음

3.2 전체 파이프라인 개요

본 서비스는 (그림 2)와 같이 순차적 처리 파이프라인으로 구성된다.



(그림 1) 전체 파이프라인 개요.

3.3 얼굴 검출 및 전처리

MediaPipe FaceDetection 을 사용하여 입력 이미지에서 가장 큰 얼굴을 검출한다. 검출된 얼굴 영역은 20% 패딩을 추가하여 Cropping 하며, 이후 224×224 픽셀로 리사이즈 및 Grayscale 변환, 정규화 과정을 거친다. 이 과정은 딥러닝 모델의 입력 일관성 및 분류 성능 향상을 위해 설계되었다.

3.4 딥러닝 기반 동물 분류

동물 분류는 Enhanced ResNet34 구조를 기반으로 하며, Attention 및 FPN 모듈을 추가하여 다양한 스케일의 특징을 효과적으로 통합한다[2]. 남성/여성 각각에 대해 별도의 모델을 학습시켜, 남성(9 종), 여성(6 종) 동물 클래스를 분류한다.

3.5 얼굴 특징 추출

MediaPipe FaceMesh 를 활용하여 468 개 얼굴 랜드마크를 추출, 주요 특징 점 간의 거리 및 각도를 계산하여 9 개 정량적 특징을 0~1 범위로 정규화한다. 이 특징 값은 프롬프트 생성 단계에서 세부 묘사로 활용된다[3][4].

3.6 프롬프트 자동 생성

동물 분류 결과, 사용자 선택 스타일, 얼굴 특징 값을 결합하여, 규칙기반 프롬프트 엔지니어링 기법을 활용하여 자연어 프롬프트를 자동 생성한다[5]. 프롬프트는 Application Programming Interface(API) 호출을 통해 생성하며 활용한 모델은 Gemini 1.5 Flash 이다.

프롬프트 엔지니어링 예시(시크)	출력된 이미지 생성용 프롬프트
당신은 시크한 타투 디자인 프롬프트 전문가입니다. 아래 ... - 동물 유형:{animal_type}... 이미지 스타일 요구사항 - 스타일: 시크하고 칼과 꽃이 포함된 ... (이하생략)	A stylish, monochrome line art tattoo design of a male dog, featuring a long, rectangular face with small, closely set eyes, a downturned mouth, short nose, and oval face ... (이하생략)

<표 1> 프롬프트 입력 및 출력 예시

3.7 병렬 타투 이미지 생성

생성된 프롬프트를 기반으로 Google Gemini 2.0 Flash Preview Image Generation 모델을 통해 타투 이미지를 생성한다. ThreadPoolExecutor 를 활용하여 4 개의 타투 이미지를 병렬로 생성하며, 추가로 라틴어 레터링 타투 이미지를 별도로 생성하여 생성속도를 높였다.

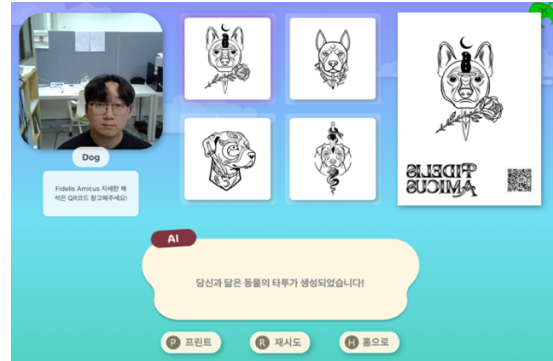
3.8 UI/UX 설계 및 공유 서비스

UI/UX 는 초반 요구사항 분석 및 설계를 기반으로 하였으며, 성별 선택, 사진촬영, 스타일 선택 등 알파/베타 테스트를 통한 사용자 피드백을 반영하여 키오스크와 같이 동작하게 하는 등 지속적으로 개선하였다.

또한 사용자가 생성한 타투 이미지를 1 시간 동안 임시로 공유할 수 있도록 QR 코드를 생성한다. 타투 생성 시 발급되는 고유 토큰을 주소로 하는 엔드 포인트를 통해 공유페이지로 접근이 가능하다. 공유 정보는 SQLite 데이터베이스에 저장되며, 만료 시 자동 삭제된다.

3.9 결과 및 UI/UX 설계화면

(그림 3)은 실제 사진 기반 맞춤형 타투 생성 서비스 화면을 나타낸다. 화면 좌측 상단에는 원본 이미지와 매칭된 동물 상이 표시되며, 하단의 버튼을 통해 다양한 타투 스타일을 선택할 수 있다. 선택된 타투는 동물에 대한 표현을 나타낸 라틴어 레터링과 함께 우측의 인쇄 레이아웃에 좌우 반전되어 표시된다. 또한 QR 코드를 통해 출력물에서 공유 페이지의 접속이 가능하다.



(그림 3) 실제 서비스 화면 예시

4. 결론 및 향후 연구

본 논문은 얼굴 이미지 기반의 개인화 생성형 AI 프롬프트 자동화 파이프라인을 제안하고, 실제 웹 서비스로 구현하여 그 효과를 입증하였다. 실험 결과, 제안 시스템은 기존 수동 방식 대비 품질과 만족도에서 우수함을 보였으며, 교내 각종 행사에 키오스크 형태로 투입되는 등 실시간 서비스로써 실용성도 입증하였다. 향후 연구로는 동물 클래스 및 스타일 다양화, 실시간 처리 최적화, 다중 모달 융합, 문화적 맥락 반영, 그리고 실사용자 기반의 정량적 평가를 중심으로 발전시킬 예정이다.

참고문헌

- [1] 황원준, "딥러닝 기반 얼굴 검출, 랜드마크 검출 및 얼굴 인식 기술 연구 동향," 방송과 미디어, 제 22 권, 제 4 호, pp.41-49, 2017..
- [2] A. Vaswani et al., "Attention Is All You Need," arXiv preprint arXiv:1706.03762, 2017.
- [3] 목진왕, 광노윤, "MediaPipe Face Mesh 를 이용한 얼굴 제스처 기반의 사용자 인터페이스의 성능 개선," 사물인터넷 융복합논문지, 제 9 권, 제 6 호, pp.125-134, 2023.
- [4] M. Meghana, M. Vasavi, and D. Shravani, "FACIAL LANDMARK DETECTION WITH MEDIAPIPE & CREATING ANIMATED SNAPCHAT FILTERS," Vol.11 Issue 06, Nov. 2021.
- [5] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM Comput. Surv. 55, 9, Article 195 2023