

최근 정보통신 기술 기반 On-Device AI 스마트 앨범 시스템에 관한 연구

강민서, 남윤창
한성대학교

vhehddl2212@naver.com, ycnham@gmail.com

A Study on the On-Device AI-Based Smart Album System Using Recent ICT

Kang Minseo, Nam Yoonchang

Hansung Univ

요약

본 논문은 On-Device AI 기술을 활용하여 스마트폰 내 사진을 자동으로 분류, 태깅, 검색, 시각화하는 '스마트 앨범 시스템' 개발에 관한 연구이다. 사용자의 여행 발자취를 시간과 위치 정보를 기반으로 지도상에 시각화하고, 객체 인식을 통해 자동 태깅하며, OpenAI CLIP 기반 자연어 검색 기능을 구현하였다. MobileNet V3, YOLO v8 과 같은 경량화된 AI 모델을 활용하여 기기 내에서 실시간 처리가 가능하며, 프라이버시 보호와 효율성 측면에서 우수한 성능을 보인다. SQLite 기반 로컬 데이터베이스를 활용하여 데이터를 안전하고 빠르게 관리할 수 있도록 설계하였다. 본 연구는 사용자 프라이버시를 보장하는 AI 기반 차세대 스마트 미디어 플랫폼으로의 확장 가능성을 지닌다.

I. 서론

현대 사회에서 스마트폰은 개인의 삶을 기록하는 주요 도구로 자리잡았으며, 사용자들은 일상과 여행 중 다양한 순간을 사진으로 저장하고 있다. 하지만 시간이 지남에 따라 방대한 사진 데이터가 누적되면서, 이를 효율적으로 분류하고 추억을 회상할 수 있는 체계적인 관리 시스템의 필요성이 제기되고 있다.

본 연구는 이러한 문제를 해결하기 위해 **On-Device AI 기술**을 기반으로 한 '스마트 앨범 시스템'을 설계하고 구현하였다. 해당 시스템은 사진의 메타데이터를 활용하여 시간과 장소에 따른 여행 발자취를 시각화하고, 객체 탐지를 통해 자동 태그를 생성하며, 사용자의 자연어 입력을 기반으로 원하는 이미지를 실시간으로 검색할 수 있는 기능을 제공한다. 기존의 클라우드 기반 시스템과 달리, 본 시스템은 모든 기능을 디바이스 내에서 처리하여 개인정보 유출 우려를 최소화하였으며, 실시간성과 사용 편의성을 동시에 확보하였다.

II. 본론

본논문에서는 On-Device AI 기술을 활용하여 스마트폰 내 사진을 효율적으로 분류, 태깅, 검색 및 시각화하는 스마트 앨범 시스템을 설계하고 구현하였다. 해당 시스템은 사용자 프라이버시를 보호하면서도, 고성능의 실시간 처리를 제공할 수 있도록 모든 기능을 디바이스 내에서 수행하도록 설계되었다. 본 절에서는 시스템의 주요 구성 요소와 기능, 기술 스택에 대해 구체적으로 기술한다.

1. 시스템 구성

본 시스템은 다음과 같은 세 가지 계층 구조를 바탕으로 설계되었다.

첫째, **사용자 인터페이스 (UI)**는 Android 기반으로 구현되어 사용자에게 직관적인 앨범 탐색, 검색, 편집 기능을 제공한다.

둘째, **On-Device AI 엔진**은 MobileNet V3, YOLO v8 등 경량화된 딥러닝 모델을 활용하여 이미지 분석 및 객체 인식을 수행하며, OpenAI 의 CLIP 모델을 이용하여 자연어 기반 검색을 가능하게 한다.

셋째, **데이터 관리 계층**은 SQLite 를 기반으로 하여 이미지 메타데이터 및 태그 정보를 효과적으로 저장하고 빠르게 검색할 수 있도록 설계되었다.

2. 핵심 기능

본 논문에서 구현한 스마트 앨범 시스템의 주요 기능은 다음과 같다.

(1) **여행 발자취 시작화**

사용자의 사진에 포함된 GPS 정보를 활용하여 사진 촬영 위치를 지도에 마커로 표시하고, Google Maps SDK를 활용해 시간 순 이동 경로를 시작화하였다. 캘린더 뷰와의 연동을 통해 날짜 기반 탐색도 지원한다.

(2) **객체 탐지 및 자동 태깅**

YOLO v8 기반 객체 탐지 기술을 통해 사진 속의 인물, 동물, 풍경 등 다양한 객체를 자동으로 식별하고 태그화하였으며, 사용자가 이를 기반으로 카테고리별 사진을 쉽게 분류할 수 있도록 하였다.

(3) **자연어 기반 이미지 검색**

OpenAI의 CLIP 모델을 기반으로 사용자의 자연어 입력(예: "눈 오는 산 풍경")을 임베딩하여 의미적으로 유사한 사진을 실시간으로 검색할 수 있는 기능을 구현하였다.

(4) **이미지 품질 향상 및 자동 분류**

ESRGAN을 활용하여 저해상도 이미지의 품질을 향상시켰고, AI 기반 유사도 분석을 통해 날짜, 장소, 인물별 자동 분류 기능을 제공하였다. 또한, 앨범별 대표 이미지를 자동 선정하는 기능도 포함하였다.

3. 기술 스택

본 논문에서 구현한 시스템의 주요 기술 스택은 다음과 같다.

- **AI 모델**: MobileNet V3, YOLO v8 Detection, OpenAI CLIP, ESRGAN
- **플랫폼**: Android (Java), Qualcomm AI Hub
- **데이터베이스**: SQLite
- **디자인 도구**: Figma를 활용한 UI/UX 설계
- **개발 환경**: macOS, Windows

이와 같은 기술적 구성과 기능을 바탕으로, 본 시스템은 모바일 환경에서도 안정적인 성능과 사용자 편의성을 동시에 제공할 수 있도록 구현되어 있다.

III. 결론

본 논문에서는 On-Device AI 기술을 기반으로 사용자의 사진 데이터를 자동으로 분류, 태깅, 검색 및 시작화할 수 있는 스마트 앨범 시스템을 설계하고 구현하였다. 본 시스템은 사진의 메타데이터를 활용한 여행 발자취 시작화, YOLO v8 기반 객체 인식 및 자동

태깅, CLIP 모델을 이용한 자연어 이미지 검색, 그리고 ESRGAN을 통한 이미지 품질 향상 등 다양한 기능을 통합하여 사용자 편의성과 정보 활용성을 극대화하였다.

특히 모든 기능이 디바이스 내에서 처리되도록 설계되어, 개인정보 보호와 실시간 처리 성능이라는 두 가지 측면에서 기존의 클라우드 기반 시스템보다 뛰어난 성능을 보였다. SQLite 기반 로컬 데이터베이스를 활용함으로써 저장 효율성과 접근성도 확보하였다.

향후 본 시스템은 AI 기반 개인 디지털 아카이빙 솔루션으로 발전할 수 있으며, 사용자 프라이버시를 보장하는 차세대 스마트 미디어 플랫폼으로의 확장 가능성을 지닌다. 이를 통해 데이터 중심 사회에서의 효율적 정보 자산화와 지능형 콘텐츠 관리 체계 구축에 기여할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW 중심대학사업의 결과로 수행되었음(2024-0-00049)..

참 고 문 헌

- [1] Qualcomm, "OpenAI CLIP Model," GitHub, https://github.com/quic/ai-hub-models/tree/main/qai_hub_models/models/openai_clip
- [2] Qualcomm, "YOLOv11 Detection," Hugging Face, <https://huggingface.co/qualcomm/YOLOv11-Detection>
- [3] Qualcomm, "MobileNet V3 Large," GitHub, https://github.com/quic/ai-hub-models/tree/main/qai_hub_models/models/mobilenet_v3_1arge
- [4] Qualcomm, "ESRGAN Model," GitHub, https://github.com/quic/ai-hub-models/tree/main/qai_hub_models/models/esrgan
- [5] Salesforce, "BLIP: Bootstrapping Language-Image Pre-training," GitHub, <https://github.com/salesforce/BLIP>
- [6] Google, "EfficientNet," GitHub, <https://github.com/tensorflow/tpu/tree/master/models/official/efficientnet>
- [7] Qualcomm, "YOLOv11 Detection (GitHub)," GitHub, https://github.com/quic/ai-hub-models/tree/main/qai_hub_models/models/yolov11_det
- [8] Qualcomm, "MobileNet V3 Large," Hugging Face, <https://huggingface.co/qualcomm/MobileNet-v3-Large>
- [9] Qualcomm, "ESRGAN," Hugging Face, <https://huggingface.co/qualcomm/ESRGAN>
- [10] Qualcomm, "OpenAI-CLIP," Hugging Face, <https://huggingface.co/qualcomm/OpenAI-Clip>