

Pic-Tag

소상공인 맞춤형 경량화 CCTV AI SaaS 개발 (하드웨어 제약 극복을 위한 재인식 알고리즘 기반 객체 탐지 시스템 구현)

시스템 아키텍처

모듈형 AI 파이프라인 설계를 통해 객체 탐지, 임베딩 생성, 재인식 로직을 분리하여 시스템의 확장성 및 재사용성 확보
처리율 향상을 위해 비동기 기반 고성능 처리 파이프라인 구축
서비스 목적에 따라 다중 통신 프로토콜을 분리 적용하여 통신 독립성 및 안정성 극대화

핵심 기능

재인식 알고리즘 개발 및 학습 고도화
한국인 재식별 이미지 데이터셋을 증강 처리 기법을 통해 학습 데이터의 다양성을 확보 및 AI 모델의 일반화 성능을 고도화
학습 효율을 극대화하기 위해 계통 샘플링 기법을 적용하여 데이터셋을 필터링하여 전체 학습 효율을 50% 이상 개선
동일 인물/다른 인물 간 유사도 값 차이를 분석하여 최적의 기준값 (Threshold)을 설정하고 ID 할당 로직 구현

실시간 객체 탐지 및 대시보드 통합

RTSP 프로토콜로 수신된 스트리밍 데이터를 Camera Frame Capture 스레드에서 처리하고, Django 기반 Presentation Layer와 Websocket 통신을 연결하여 실시간으로 객체 탐지 결과를 Web Dashboard에 반영
AI 파이프라인 후단 모델로 데이터를 전달할 때 버퍼 방식을 사용하여 데이터 병목 현상을 관리하고 시스템의 처리율 향상

문제 해결

알고리즘 최적화를 통한 데이터 병목 현상 해결

초기 재인식 알고리즘 설계 시 데이터 처리 과정에서 병목 현상이 발생할 가능성이 높아, 처리 속도 개선이 필요했습니다.

이에 Yolo 백본 분해 후, 성능 향상을 위해 특징 추출 (Feature Extraction) 알고리즘을 Linear → Pooling → Attention 순으로 비교 분석하였으며, 최종적으로 Attention 알고리즘을 선택했습니다. 이 알고리즘은 모든 은닉 상태 벡터에 Attention Score를 부여하고, 이를 기반으로 매 순간 새로운 벡터를 생성하는 방식으로 데이터의 병목 현상을 효과적으로 해결하고 재인식 정확도를 개선했습니다. 추가적으로 검증이 완료된 AI 모델의 추가적인 경량화 및 최종 응답 속도 향상을 위하여 Intel OpenVINO Toolkit을 이용하여 모델 최적화 (Model Optimization)를 수행했습니다.

회고

프로젝트를 수행하는 과정에서 RTSP 통신, Websocket 기반 실시간 통신 등 다양한 프로토콜을 구현하고, Yolo 백본 분해, OpenVINO와 같은 고난도 기술적 해법을 찾아내며 미지의 기술에 대한 도전 정신과 실제 사용 환경을 고려하는 솔루션 설계 역량을 확보했습니다. 특히, 하드웨어 제약 조건 하에서 성능을 최적화하기 위해 실시간 처리 파이프라인을 설계한 경험은 실제 사용 환경을 고려하는 솔루션 설계 역량을 확보하는 데 기여했습니다.

향후 프로젝트에서는 초기 모델의 성능과 정확도 검증을 더욱 체계적으로 수행하여 모델의 실질적인 현업 적용 가능성을 높일 계획입니다. 복잡한 시스템 구조를 다루면서 발생할 수 있는 잠재적 버그를 최소화하기 위해, 테스트 자동화 및 CI/CD 파이프라인을 구축하는 것을 다음 목표로 설정할 것입니다.

