2024년도 <종단형 PBL> 성과 발표

2024.12.6.(금) 10:30~14:30

관리번호	2024-

프로젝트 개요

프로젝트 명칭		
	국문명	AI 활용 안면 정보 노출 검색 시스템 : 능동적 개인정보보호와 초상권 관리의 시작
	영문명	AI-Based Facial Information Exposure Detection System : The Beginning of Active Privacy Protection and Portrait Rights Management

프로젝트 팀	팀명	ASAP (AI, Security, Analysis, Privacy)	지도교수	박후린 교수님

팀구성	학과	학번	이름	휴대폰번호	E-mail 주소
팀원1(팀장)	경영학과 · 정보보호학과		조은빛		eunbit2001@naver.com
팀원2	정보보호학과		김다은		daeun6913@gmail.co
팀원3	정보보호학과		노영주		noyongju@naver.com
팀원4	정보보호학과		김은지		eun02@swu.ac.kr
팀원5					
팀원6					

오픈소스 탑재 URL 명시 (GitHub, SourceForge)	https://github.com/ASAP-Find-My-Face/Find-My-Face

프로젝트 요약

① 개발 배경 (이 프로젝트를 왜 개발하게 되었나?)

본 프로젝트는 디지털 환경과 SNS 확산으로 인한 얼굴 정보 노출과 초상권 침해 문제를 해결하기 위해 개발되었다. 얼굴 이미지를 서버에 저장하지 않고 바이너리 파일(키 값)로 변환하여 개인정보 보호를 강화하며, 사용자가 자신의 얼굴이 포함된 콘텐츠를 탐지하고 관리할 수 있도록 지원한다. 더불어 기존에 얼굴 인식 기술보다 정확도와 속도를 제공한다.

② 전체 구성 (이 프로젝트는 어떻게 구성되어있나?)

프로젝트는 사용자 등록, 개인정보처리방침 및 이용 동의, 고유 키 생성, 영상 키 생성, 키 간의 검색 및 매칭, 결과 제공으로 구성되어 있으며, 얼굴 특징점을 바이너리 파일로 변환하여 저장 공간을 보안적이며, 효율적으로 관리한다. 또한 직관적인 웹 인터페이스를 통해 개발 지식이 없는 사용자도 쉽게 접근할 수 있다.

③ 구현 환경 (이 프로젝트의 구현을 위한 개발환경은?)

개발환경은 Python(Django 웹 프레임워크), SQLite 데이터베이스, HTML/CSS/JavaScript 기반의 프론트엔드로 구성되었으며, OpenCV와 face_recognition 오픈소스를 활용했다.

④ 구현 결과 (이 프로젝트는 어떤 상태로 최종 구현되었는가?)

최종 구현된 시스템은 얼굴 데이터 등록 및 고유 키 생성, 영상 키 생성, 매칭 및 검색 결과 제공 기능을 완성하였으며, 유사 얼굴 매칭 정확도 97%, 데이터 저장 공간 99.5% 절감을 달성했다. 추가적으로 개인정보보호 및 보안 강화를 위한 개인정보처리 방침 및 이용 사전 고지 및 동의 단계 진행, 본인 인증 절차를 기능을 구현하였으며, 타 플렛폼 적용을 위한 라이브러리와 API 개발 마무리 및 제안 예정이다.

프로젝트 상세 설명

프로젝트 명 칭	국문명	AI 활용 안면 정보 노출 검색 시스템 : 능동적 개인정보보호와 초상권 관리의 시작		
	영문명	AI-Based Facial Information Exposure Detection System : The Beginning of Active Privacy Protection and Portrait Rights Management		
오픈소스 URL		https://github.com/ASAP-Find-My-Face/Find-My-Face		
프로젝트 팀명		ASAP (AI, Security, Analysis, Privacy)	지도교수	박후린 교수님
팀원(이름·학과·학번)				

1. 프로젝트 개요

1.1 개발 필요성

디지털 환경의 발전과 SNS의 확산은 개인정보, 특히 얼굴 정보 노출 문제를 심화시키며 초상권과 프라이버시 침해로 이어지고 있다. 그러나 현행법과 개인과 기업 단계에서 예방하고 탐지하여 대응할 기술 부족으로 인해 신속히 해결하기 어려운 상황이다. 또한, 개인의 얼굴이 포함된 콘텐츠 관리에 대한 요구와 비주얼 검색 선호 및 온라인 평판 관리의 중요성도 높아지고 있다. 이러한 배경에 기존 얼굴 검색 시스템은 개인정보 보호 취약성, 데이터 관리 비효율성, 사용자 접근성 부족 등의 문제가 있음을 파악하여 본 프로젝트는 개인정보 유출 방지와 초상권 보호를 위한 안전하고 효율적인 기술적 대안을 제시하고자 한다.

1.2 목적

본 프로젝트는 **안전한 개인정보 보호, 효율적인 데이터 관리, 사용자 중심의 서비스 제공**이라는 세 가지 주요 목적을 가지고 있다. 첫째, 사용자는 자신의 얼굴 정보가 노출된 콘텐츠를 스스로 검색하고 필요한 경우 노출 제한이나 삭제 요청을 통해 개인정보를 능동적으로 보호할 수 있다. 얼굴 이미지는 서버에 저장하지 않고 고유 키로 변환하여 검색에 활용함으로써 데이터 유출 위험을 최소화하고 프라이버시를 효과적으로 보호한다. 둘째, 얼굴 특징점을 바이너리 파일로 저장하여 기존 이미지 저장 방식에 비해 약 99.5%의 저장 공간을 절약하여 대규모 데이터베이스에서도 효율적인 데이터 처리가 가능하도록 한다. 셋째, 직관적이고 간단한 사용자 인터페이스를 제공하여 개발 지식이 없는 일반 사용자도 자신의 얼굴 정보를 쉽게 검색하고 관리할 수 있는 사용자 중심의 서비스를 제공하고자 한다.

1.3 용도

본 시스템은 개인과 기업 단계에서의 범죄 선대응 및 예방, 법적 증거 수집, 온라인 평판 관리 등 다양한 상황에서 활용될 수 있다. 타 영상 플랫폼의 적용으로 사용자의 얼굴이 포함된 콘텐츠를 신속히 탐지하여 원하지 않는 노출을 제한하거나 삭제를 요청함으로써 초상권 침해와 개인정보 유출을 방지할 수 있다. 또한, 초상권 침해 사례에서 법적 대응에 필요한 증거 자료로 활용할 수 있으며 자신의 얼굴이 포함된 영상을 탐지해 법적 문제 해결을 지원한다. 마지막으로 기업이나 개인은 시스템을 통해 인터넷상에서 자신의 이미지가 어떻게 활용되고 있는지 모니터링하고, 부적절한 이미지 사용을 확인하여 온라인 평판을 효과적으로 관리할 수 있다.

1.4 유사 기술과의 차별성 및 독창성

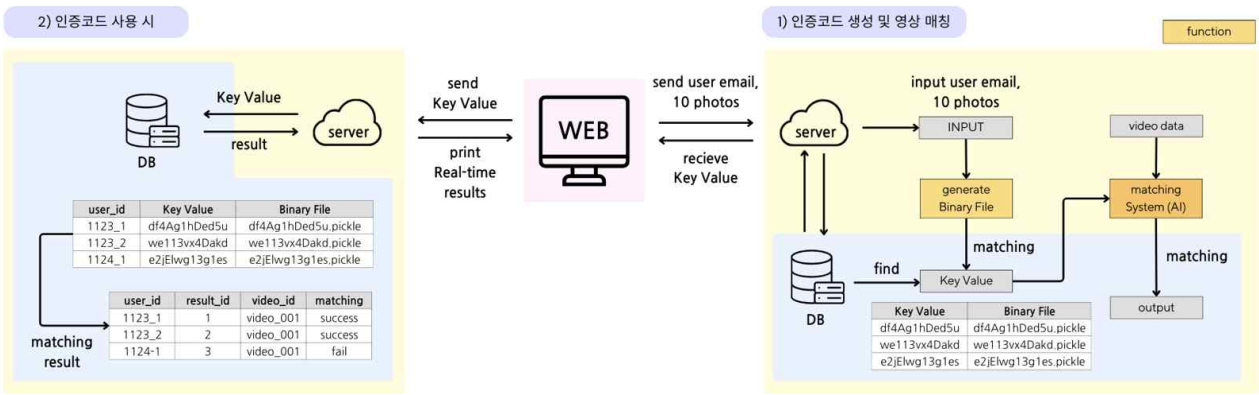
기존 얼굴 인식 및 검색 시스템은 얼굴 이미지를 직접 다루는 방식으로 개인정보 보호와 데이터 관리, 사용자 접근성에서 한계를 드러내고 있다. 본 프로젝트는 이러한 문제를 해결하고, 기존 기술과의 차별성을 명확히 하며 독창성을 제공한다.

기존 기술은 사용자 얼굴 이미지를 직접 저장하여 해킹, 유출 등의 보안 문제를 초래했다. **개인정보 보호** 측면에서 본 프로젝트는 사용자 얼굴 이미지를 직접 저장하지 않고, 얼굴 특징점을 고유 키로 변환해 저장 및 검색에 활용한다. 이를 통해 데이터 유출 위험을 최소화하고, 사용자가 자신의 얼굴 정보 노출 여부를 안전하게 탐지하고 관리할 수 있는 환경을 제공한다. 이는 Clearview AI와 같은 이미지 기반 서비스가 초래한 문제를 해결하고 PimEyes 같은 단순 이미지 기반 검색 서비스에 비해 개인정보 보호 수준을 한층 강화한다. 둘째, **데이터 관리 효율성**에서는 얼굴 특징점을 바이너리 파일로 저장함으로써 기존 이미지 저장 방식에 비해 약 99.5%의 저장 공간을 절감하고 대규모 데이터베이스에서도 안정적이고 빠른 데이터 처리가 가능하다. 이는 고해상도 이미지를 저장해야 하는 기존 기술의 데이터 관리 효율성을 혁신적으로 개선한다. 마지막으로 **사용자 중심 설계**를 통해 개발 지식이 없는 일반 사용자도 직관적인 인터페이스를 이용해 쉽게 자신의 얼굴 정보를 탐지하고 관리할 수 있다. 사용자는 간단히 고유 키를 입력해 자신의 얼굴이 포함된 콘텐츠를 탐지하고, 불필요한 노출을 제한하거나 삭제 요청을 할 수 있다. 이는 Amazon Rekognition, Microsoft Azure Face API와 같은 개발자 중심의 시스템과 차별화된 사용자 경험을 제공한다.

2. 프로젝트 구성

2.1 시스템 전체 구성도

본 프로젝트의 모든 작업은 사용자가 브라우저를 통해 접근하는 웹 플랫폼에서 수행된다. 시스템의 전체 다이어그램은 아래와 같다.



❶ 키 값(인증코드) 생성

사용자가 자신의 이메일과 얼굴 사진 10장을 업로드하면, 서버는 각 사진에서 얼굴의 특징점을 추출해 바이너리 파일(pickle 파일)로 변환하고, 이 파일의 이름을 고유한 키 값으로 설정하여 데이터베이스에 저장한다. 생성된 키 값은 사용자가 검색할 때 사용할 인증 코드로 사용된다. '인증 코드'라는 용어는 '키 값'이라는 용어가 사용자들에게 생소할 수 있기 때문에 더 쉽게 이해할 수 있도록 사용한 표현이다. 마지막으로, 서버는 사용자에게 인증 코드(키 값)를 부여하여, 이후 사용자가 이 인증 코드(키 값)를 통해 검색 결과를 조회할 수 있도록 한다.

❷ 영상 키값 생성 저장

영상에서 프레임 단위로 얼굴을 인식하면 얼굴의 특징점을 추출하여 키 값 생성과 동일한 과정을 거쳐 킷값을 생성하여 저장한다.

❸ 키 값(인증 코드)-키값 매칭 및 결과 확인

사용자는 부여받은 인증 코드를 "FIND MY FACE" 웹 페이지의 입력란에 입력한다. 인증 코드를 통해 서버는 해당 사용자(user_id)를 식별하고, 해당 사용자의 키값과 영상 키값을 매칭한 결과를 실시간으로 사용자에게 제공함으로써 사용자는 자신이 나온 영상을 찾을 수 있다. 해당 인증 코드는 3일 동안 유효하며, 유효기간이 지난 후 사용자는 새로운 인증 코드를 통해 매칭 결과를 수행할 수 있다.

2.2 기획 및 설계

❶ 사용자 흐름 및 설계

타 플렛폼 크롤링 시 발생하는 법적 문제 발생 예방을 위한 시현용 영상 플랫폼 제작하였으며, 홈 페이지는 기본 영상 플랫폼과 같이 업로드된 영상을 확인할 수 있다. 영상 업로드 페이지는 영상 업로드 시 이를 키 값으로 변환하여 DB에 저장한다. FIND MY FACE 페이지는 서비스의 핵심 기능을 제공하며, 사용자가 자신의 사진을 통해 키 값(인증코드)을 생성하고 이를 이용하여 자신이 나온 영상 결과를 조회할 수 있다.

❷ 얼굴 감지 및 특징점 추출

전처리는 사용자가 업로드한 얼굴 이미지를 처리하는 단계로 face_recognition 라이브러리를 활용하여 진행한다. 먼저 이미지 크기를 256x256 픽셀로 표준화하여 연산 속도와 인식 성능을 높인다. 다음으로 밝기와 대비를 조정을 통해 얼굴 경계를 명확히 하여 감지 성능을 높인다. 또한, 컬러 채널을 BGR에서 RGB로 변환하여 얼굴 감지 모델과의 호환성을 확보한다. 마지막으로 PIL 이미지를 NumPy 배열로 변환하는 과정을 포함한다.

얼굴 감지 및 특징점 추출은 OpenCV와 face_recognition 라이브러리를 사용하여 진행한다. 업로드된 이미지에서 얼굴 영역을 탐지한 후 감지된 얼굴에서 고유 특징점을 추출하여 이를 임베딩 벡터로 변환한다. 생성된 벡터는 사용자의 얼굴 정보를 수치화한 데이터로 다른 얼굴과 비교할 수 있는 기초 데이터로 활용된다.

❸ 고유 키 값 생성 및 저장

추출된 특징점을 바이너리 파일 형식으로 변환하여 저장 공간을 최소화한다. 이 방식은 기존 이미지 저장과 비교해 얼굴 정보 저장 공간을 약 99.5% 절약하며, 원본 이미지를 복구할 수 없도록 설계되어 개인정보 유출 위험을 낮춘다. 기존 기술과 달리 사진을 직접 DB에 저장하지 않기 때문에 보안성이 더욱 강화된다. 고유 키 생성 과정에서는 바이너리 파일의 이름을 랜덤 값(고유 키)으로 설정하여 보안성을 높인다. 생성된 고유 키는 바이너리 파일 이름으로 사용되며 사용자가 얼굴 데이터를 검색하는 데 활용된다.

❹ 영상 키 값 생성

사용자가 플랫폼에 영상 업로드 시 프레임 단위로 분리하여 각 프레임을 이미지로 변환하는 작업이 진행된다. 변환된 이미지로 키 값을 생성하는 과정은 고유 키 값 생성과 동일한 방식으로 진행되며, 영상에서 추출한 키 값도 랜덤 값으로 설정되어 저장되며 사용자의 얼굴 데이터 검색 시 활용된다.

❺ 고유 키 값(인증코드) - 영상 키 값 매칭 결과 출력

데이터베이스에 저장된 사용자 특징점과 영상에서 생성한 키 값을 비교한다. 유사도 측정은 유클리드 거리를 기반으로 이루어지며 거리가 임계값(0.35) 이하일 경우 동일 인물로 판단하여 결과를 반환한다. 매칭된 콘텐츠는 실시간으로 웹 플랫폼에 표시되며 사용자가 결과를 확인하고 관리할 수 있도록 지원한다. 결과는 사용자가 인식된 영상과 영상 이름, 인식된 시간, 매칭 유사도를 함께 사용자에게 보여준다.

2.3 주요 기능 설명

본 시스템은 AI 기반 얼굴 데이터 분석 및 검색 기술을 활용하며, 주요 기능은 다음과 같다.

❶ 고유 키 기반 얼굴 정보 검색

사용자는 고유 키를 통해 자신의 얼굴이 포함된 콘텐츠를 안전하게 검색한다. 키 기반 접근 방식은 개인정보를 직접적으로 다루지 않으며, 보안성과 효율성을 동시에 제공한다.

❷ 저장 공간 최적화

얼굴 정보를 바이너리 파일 형식으로 저장하여 기존 이미지 저장 방식 대비 약 99.5%의 저장 공간 절감 효과를 제공한다. 대규모 데이터베이스에서도 빠르고 안정적인 데이터 관리가 가능하다.

❸ 사용자 친화적 인터페이스

직관적인 웹 플랫폼 설계를 통해 사용자는 개발 지식 없이도 자신의 얼굴 정보를 검색하고 관리할 수 있다. 검색 결과는 실시간으로 제공되며, 능동적인 초상권 관리가 가능하다.

3. 구현 환경

3.1 팀 소개 및 역할

❶ 팀명: ASAP (AI, Security, Analysis, Privacy)

- o 조은빛(팀장): 프로젝트 전반 관리, 파일 구조 설계, 영상 매칭 파트 개발, 성능 평가 진행
- o 김다은: 데이터 전처리, 서버 전처리 개발, 영상 플랫폼 및 서비스 프론트앤드 개발
- o 노영주: 데이터베이스 설계, 프로젝트 서버 환경 구축, 영상 플랫폼 및 영상 업로드 개발
- o 김은지: 프로젝트 구현 관리, 웹 UI/UX 디자인, 키값 생성 및 영상 매칭 파트 개발

❷ 개발환경

프로젝트는 Python을 사용하여 Django 웹 프레임워크 기반으로 개발되었다. 초기에는 Docker 환경에서 실행하려 했으나 환경 설정 문제로 인해 로컬 개발 환경으로 전환하여 진행되었다. 데이터베이스는 SQLite를 사용하며, 프론트엔드는 HTML, CSS, JavaScript로 구성되었다. 개발은 Windows 11 기반의 로컬 서버에서 이루어졌고, GitHub를 통해 형상 관리를 진행하였다.

❸ 활용한 오픈소스

- o OpenCV: 얼굴 감지 및 영상 처리
- o face_recognition: 얼굴 특징점 추출 및 매칭
- o Django Rest Framework (DRF): REST API 설계

❹ 독자적 개발

- o 얼굴 이미지에서 특징점을 추출하여 고유 키 값으로 변환하는 알고리즘 개발
- o 얼굴 특징 데이터의 저장공간 최적화를 위한 바이너리 파일 관리 시스템 구축
- o 사용자 친화적 웹 UI 설계 및 인증 코드 기반 검색 시스템 구현

4. 구현 결과

4.1 프로젝트 진행 현황 및 결과 소개

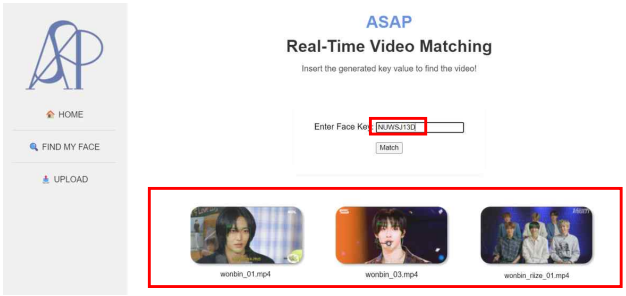
❶ 완성된 기능

o 개인정보 보호 강화를 위한 얼굴 데이터 고유 키 생성 및 바이너리 파일 저장

o 영상 업로드 시 키 값 생성 후 매칭 시 사용

o 사용자 얼굴 데이터 등록 및 고유 키 생성

o 동영상 업로드 후 프레임 단위 얼굴 검색 및 매칭 후 결과



Video Name	Match Time (Seconds)	Match Similarity (%)
wonbin_01.mp4	0.00	29.67%
wonbin_03.mp4	10.50	33.33%
wonbin_riize_01.mp4	0.50	31.10%

(참고) 다음은 woonbin의 사진으로 키 값(인증 코드)를 생성하였으며, woonbin이 나온 비디오는 woonbin_01, woonbin_03, woonbin_riize_01 밖에 없기에 올바른 결과 값을 출력하는 것을 알 수 있음

❷ 기술적 성과:

1. 기존 이미지 저장 대비 약 99.5%의 데이터 저장공간 절감
2. 유사 얼굴 매칭 정확도: 97% 이상
3. 검색 응답 시간: 동영상 1분 기준 평균 10초

4.2 활용 방법 및 사용법

❶ 사용자 플로우:

o **등록**: 사용자가 자신의 눈, 코, 입이 나온 사진 10장을 업로드하면 인증코드(고유 키 값)를 생성하여 사용자에게 제공한다.

o **인증코드를 이용한 매칭 결과 확인**: 사용자가 고유 키를 입력하면 매칭이 시작되고, DB에서 매칭된 동영상 이름, 처음 매칭된 시간과 매칭율을 결과를 출력한다.

❷ 응용 분야

o 초상권 침해 확인 및 관리

o 기업의 온라인 평판 관리

o 법정 증거 수집

❸ 오픈소스 URL

https://github.com/ASAP-Find-My-Face/Find-My-Face

5. 기대효과 및 향후 작업

5.1 기대효과

본 시스템은 개인, 기업, 사회 전반에 걸쳐 개인정보 보호 측면에 긍정적인 영향을 미칠 것으로 기대된다.

❶ 개인정보 보호 인식 고취

먼저 **개인 사용자의 관점**에서 사용자는 자신의 얼굴이 포함된 콘텐츠를 쉽게 탐지하고, 원치 않는 콘텐츠를 제한하거나 삭제 요청을 할 수 있다. 이를 통해 개인은 자신의 프라이버시를 보다 능동적으로 보호하고 초상권 침해를 예방할 수 있다. 또한, **기업 플랫폼의 관점**에서는 개인정보 보호에 대한 사용자 요구가 높아짐에 따라 플랫폼 기업들은 서비스의 신뢰성을 확보하고 개인정보 보호 정책을 강화하게 될 것이다. 이러한 변화는 플랫폼의 사용자 신뢰도를 높이고, 서비스 품질 개선으로 이어질 것으로 기대된다. 마지막으로 **사회적 관점**에서는 개인정보 보호와 초상권 관리의 중요성에 대한 인식이 확대되고, 디지털 시대에 적합한 개인정보 보호 문화를 형성하는 데 기여할 것이다. 이는 사회 전반에서 프라이버시와 데이터 보호를 성숙하게 다루는 문화를 촉진할 것이다.

❷ 개인정보 침해 범죄 대응

개인정보 침해 범죄 대응을 위해 본 기술**은** 딥페이크와 같은 개인정보 침해 범죄를 방지 및 대응하는 기술적 기반을 제공한다. 사용자는 자신의 얼굴 정보가 유해 콘텐츠에 불법적으로 노출되는 것을 탐지하여 예방할 수 있으며, 이러한 정보를 법적 대응에도 활용될 수 있다. 또한 사회적 경각심을 높이는 동시에 법적 규제와 처벌 강화를 통해 개인정보 침해 범죄의 예방과 대응에 효과적으로 기여할 것이다.

❸ 개인 및 기업의 온라인 평판 관리

본 시스템을 통해 개인 및 기업은 자신의 온라인 이미지를 모니터링하고 관리할 수 있다. 또한 부적절하거나 불만스러운 콘텐츠를 탐지하고 제거함으로써, 디지털 환경에서 개인 및 기업의 평판을 효과적으로 보호할 수 있다. 나아가 개인과 기업 모두가 신뢰할 수 있는 디지털 정체성을 유지할 수 있게 된다.

5.2 향후 작업

❶ 키 값 생성과 동시에 영상 키 값과 자동 매칭 기능 추가

고유 키 생성 후 자동으로 영상 매칭을 수행하여 더욱 빠르고 효율적인 검색 결과를 제공한다.

❷ 로그인 기능 및 본인 인증 절차 추가

보안성과 신뢰성을 강화하기 위해 로그인 기능과 본인 인증 절차를 추가할 예정이다. 개인의 키 값 생성 및 관리 과정에서 본인 인증 절차를 도입함으로써 비인가자의 접근을 방지하고, 시스템 남용을 예방할 것이다. 이를 통해 사용자 개인정보 보호와 시스템 신뢰성을 높일 수 있다.

❸ 상용화를 위한 플랫폼용 라이브러리 및 API 개발

본 시스템의 핵심 기술을 라이브러리 형태로 개발하여 유튜브, 네이버와 같은 대형 플랫폼에서도 쉽게 도입할 수 있도록 제공할 예정이다. 이러한 라이브러리와 API는 플랫폼 사용자들이 안전하게 얼굴 정보를 관리할 수 있도록 지원하며, 서비스 이용 약관에 따라 사용자의 권리를 보호할 것이다.