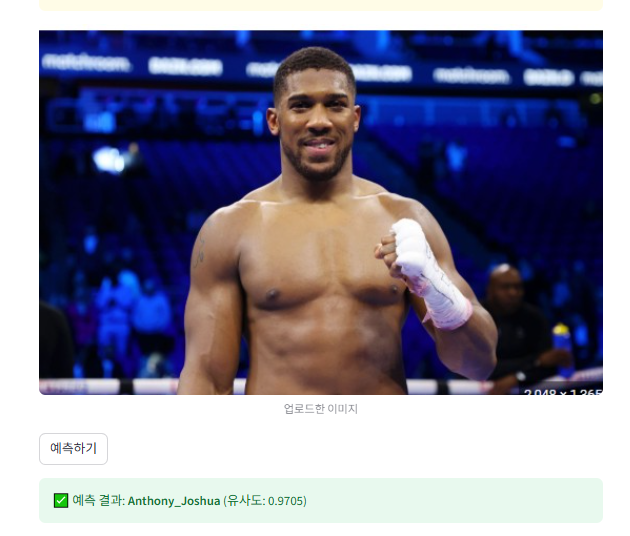
**실시간 얼굴 유사도 판별 (InsightFace + FAISS + Streamlit)**

****

****

웹에서 이미지를 업로드하면 **사전에 등록한 인물 임베딩들 가운데 가장 유사한 사람을 즉시 찾아주는** 데모 서비스입니다.

얼굴 임베딩은 **InsightFace**로 추출하고, **FAISS**로 고속 최근접 탐색을 수행합니다. **Streamlit** UI로 사용/시연성이 뛰어나며, 향후 사내 인증/미디어 검색/사진 정리(중복/유사 인물 군집화) 등으로 확장 가능합니다.

* 핵심 모델: InsightFace FaceAnalysis (buffalo\_l) – 얼굴 감지·정규화·128D 임베딩 추출
* 검색 엔진: FAISS IndexFlatIP – L2 정규화 후 내적(=코사인 유사도) 기반 Top-K 탐색
* 전처리: OpenCV 디코딩, (오프라인 임베딩 구축 시) Albumentations로 Flip/밝기/회전 증강
* 산출물: faiss\_index.index(인덱스), faiss\_labels.pkl(라벨)
* 시각화: streamlit

## **문제 정의와 해결 전략**

* **문제**: 여러 명의 얼굴 중, 입력 이미지와 **가장 비슷한 인물**을 빠르게 찾아야 함.
* **전략**:  
  1. 얼굴 임베딩을 **정확하게**(조명/각도 변화를 버티도록) 뽑는다.
  2. 모든 임베딩을 \*\*벡터 DB(FAISS)\*\*에 올려 **밀리초 수준**으로 최근접을 찾는다.
  3. 유사도 점수와 함께 \*\*라벨(사람 이름)\*\*을 반환해 사용자에게 직관적으로 보여준다.

## **아키텍처(논리 흐름)**

1. **데이터 준비**: dataset/person/사람명/이미지들 폴더 구조
2. **google colab에서 임베딩 구축 (**[**Insightface\_train\_test.ipynb - Colab**](https://colab.research.google.com/drive/1pVuvk2HPuCQh6qNvh3RxU9STpqu7drZN)**)**
   * InsightFace로 이미지별 **얼굴 임베딩** 추출
   * **증강(Flip/밝기/회전)** 후 임베딩을 평균 → 견고성 향상
   * 임베딩을 **L2 정규화**하고 **FAISS 인덱스** 생성, 라벨과 함께 저장
3. **local Streamlit에서 실제 사용**
   * 사용자가 이미지를 업로드 → 임베딩 추출 & L2 정규화
   * **FAISS Top-1 검색** → 최상위 라벨과 유사도 점수 표시

## **설계상의 선택과 근거**

* **InsightFace 선택**: 공개 벤치마크에서 강력한 성능, 실무 적용사례 풍부.
* **L2 정규화 + Inner Product**: 코사인 유사도와 동일한 효과로, 단순 인덱스(IndexFlatIP)와 궁합이 좋음.
* **증강 후 임베딩 평균화**: 단일 사진의 편향(표정, 각도)을 완화.
* **Streamlit**: 연구/시연/사내 PoC 속도 극대화, 데모 제작 비용 최소화.
* **CPU Execution Provider**: GPU 미사용 환경에서도 시연 가능(데모 확장성 확보).