라벨 파일이 없거나 바운딩 박스가 0개인 이미지를 제거하여 학습 데이터의 신뢰도

**SK네트웍스 Family AI 과정 13기  
 데이터 전처리 인공지능 데이터 전처리 결과서**



| **산출물 단계** | 데이터 전처리 |
| --- | --- |
| **평가 산출물** | 인공지능 데이터 전처리 결과서 |
| **제출 일자** | 2025.08.10. |
| **깃허브 경로** | https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN13-FINAL-2TEAM |
| **작성 팀원** | 박수빈 |

1. **문서 개요**

* 프로젝트명: AI 의류 쇼핑 어시스턴트 플랫폼 개발
* 사용 모델**:** YOLOv8l, Fashion CLIP
  + Fashion CLIP 개요:
* CLIP(Contrastive Language-Image Pre-training)은 OpenAI에서 2021년 발표한 멀티모달 학습 모델
* 인터넷에서 모은 4억여 쌍의 이미지-텍스트 데이터를 대규모로 학습
* 이미지와 텍스트를 동일한 임베딩 공간에 매핑하여 두 도메인 간의 유사도를 직접 계산할 수 있도록 설계 (“제로샷”/멀티모달 활용)
* 주요 패션 이미지 검색·추천, 멀티모달 AI모델 기반으로 사용
* 전처리 목적: 학습 데이터의 품질과 일관성을 확보하고, 데이터 다양성을 확장하여 YOLOv8I 및 Fashion CLIP 모델의 정확도와 일반화 성능을 향상시키기 위함
* 문제 정의:

1. **문제점**

원본 데이터의 크기, 방향, 명암이 제각각이고, 라벨이 누락된 이미지가 포함됨

이로 인해 모델이 일관된 특징을 학습하기 어렵고, 예측 성능이 저하됨

1. **개선 방향**

이미지 표준화 및 라벨 누락 데이터 제거로 데이터 품질을 향상

데이터 증강 기법을 활용하여 다양한 학습 패턴 확보 및 실제 환경 대응력 강화

1. **데이터셋 개요**

* 데이터 출처 및 수집 방법:

모델 학습용 이미지를 별도로 자체 수집 (출처:29cm/무신사/w\_concept/지그재그)

* 데이터 구성:

| **항목명** | **설명** | **예시** | **필요성**  **(YOLO → Fashion CLIP 연계)** |
| --- | --- | --- | --- |
| image\_id | 이미지 파일 명 | image\_001.jpg | YOLO 탐지 결과와 원본 이미지 매칭하여 CLIP에 해당 영역 이미지를 전달하기 위해 필요 |
| image\_path | 이미지 경로 | dataset/train/images/image\_001.jpg | CLIP 모델이 실제 이미지를 불러와 임베딩을 생성하는 입력 경로로 사용 |
| width | 이미지 가로 크기(px) | 640 | YOLO 바운딩 박스 좌표(비율)를 절대 좌표로 변환하여 CLIP에 크롭된 이미지 전달 시 필요 |
| height | 이미지 세로 크기(px) | 640 | YOLO 바운딩 박스 좌표(비율)를 절대 좌표로 변환하여 CLIP에 크롭된 이미지 전달 시 필요 |
| class\_id | 객체 클래스 ID (0, 1, 2, 3, 4) | 1 | YOLO 탐지 객체의 정수형 레이블로, CLIP 검색·추천 시 해당 카테고리 필터링에 활용 |
| class\_name | 객체 클래스명 (dress, pants, skirt&pants, skirt, top) | pants | CLIP의 텍스트 쿼리(예: “pants”)와 직접 비교하여 유사도 검색 수행 |
| bbox\_x\_center | 바운딩 박스 중심 X좌표 (이미지 비율) | 0.515625 | CLIP 입력 이미지를 크롭할 때, YOLO가 예측한 중심 좌표 정보 활용 |
| bbox\_y\_center | 바운딩 박스 중심 Y좌표 (이미지 비율) | 0.53828125 | CLIP 입력 이미지를 크롭할 때, YOLO가 예측한 중심 좌표 정보 활용 |
| bbox\_width | 바운딩 박스 너비 (이미지 비율) | 0.3453125 | YOLO 예측 객체의 가로 범위를 CLIP 크롭 이미지 생성에 반영 |
| bbox\_height | 바운딩 박스 높이 (이미지 비율) | 0.7609375 | YOLO 예측 객체의 세로 범위를 CLIP 크롭 이미지 생성에 반영 |

* 원본 데이터 샘플:

| **번호** | **images** | **labels(원본)** |
| --- | --- | --- |
| (a) |  | 1 0.515625 0.53828125 0.3453125 0.7609375 |
| (b) |  | 4 0.43671875 0.271875 0.5890625 0.396875  1 0.425 0.59375 0.4359375 0.3171875 |

* 라벨 해석:

| **번호** | **class\_name** | **class\_id** | **bbox\_x\_center** | **bbox\_y\_center** | **bbox\_width** | **bbox\_hegith** |
| --- | --- | --- | --- | --- | --- | --- |
| (a) | pants | 1 | 0.515625 | 0.53828125 | 0.3453125 | 0.7609375 |
| (b) | pants | 1 | 0.425 | 0.59375 | 0.4359375 | 0.3171875 |
| top | 4 | 0.43671875 | 0.271875 | 0.5890625 | 0.396875 |

1. **전처리 프로세스 개요**

* 전체 흐름도:  
    
  ① 데이터 수집 → ② 결측치 처리 → ③ 라벨 검증 → ④ 이미지 표준화 → ⑤ 데이터 분리 → ⑥ 데이터 증강
* 전처리 파이프라인 요약:

| **단계** | **목적** | **수행 작업** | **사용 도구/라이브러리** |
| --- | --- | --- | --- |
| 결측치 처리 | 누락값 제거 | 라벨(annotation) 파일이 없거나 바운딩 박스가 0개인 이미지 제거 | Roboflow |
| 라벨 검증 | 데이터 품질 확보 | 라벨 값이 클래스 목록에 포함되는지 검증 | Roboflow |
| 이미지 표준화 | 모델 학습 품질 향상 | Auto-Orient(방향 일치), Resize(imgsz=640), Adaptive Equalization(명암 대비 자동 보정) 적용 | Roboflow |
| 데이터 증강 | 모델 일반화 성능 강화 | Horizontal/Vertical Flip, Rotation(-10°~+10°) 적용 | Roboflow |
| 데이터 분리 | 학습/검증/테스트 데이터 분할 | 초기학습: train 4177, valid 400, test 200 / 추가 학습: train 1545, valid 150, test 77 | Roboflow |
| 모델 최적화 | 모델별 성능 개선 | YOLOv8I : 객체 검출 정확도 향상  Fashion CLIP : 이미지-텍스트 임베딩 품질 향상, 멀티모달 검색 성능 개선 | YOLOv8I : torch, numpy  Fashion CLIP : faiss |