

인공지능 학습용 데이터 구축 사업(2차)를 위한
환경오염 시데이터 검수용

Annotation Tools 매뉴얼

2020. 11

어노테이션 툴 사용방법

가. 실행하기



- 1) 바탕화면에 설치된 아이콘 ScalingAnnoter를 실행합니다.
- 2) Scaling Annoter Utility가 실행되고 로그인 합니다.
- 3) 위와 같이 Scaling Annoter Utility Main 화면이 실행됩니다.

나. 화면설명 (Scaling Annoter Utility 메뉴 구성)



1) 파일 리스트

- 촬영원본 데이터 불러오기
- 프로그램 종료
- 작업대상 목록 상태(현황)

2) 이미지 뷰어화면

- 파일 리스트에서 선택한 이미지 확인
- 이미지 회전, 확대, 축소
- 이미지 라벨링
- 라벨링(박싱)정보 확인

3) 이미지 정보 화면(Image Info)

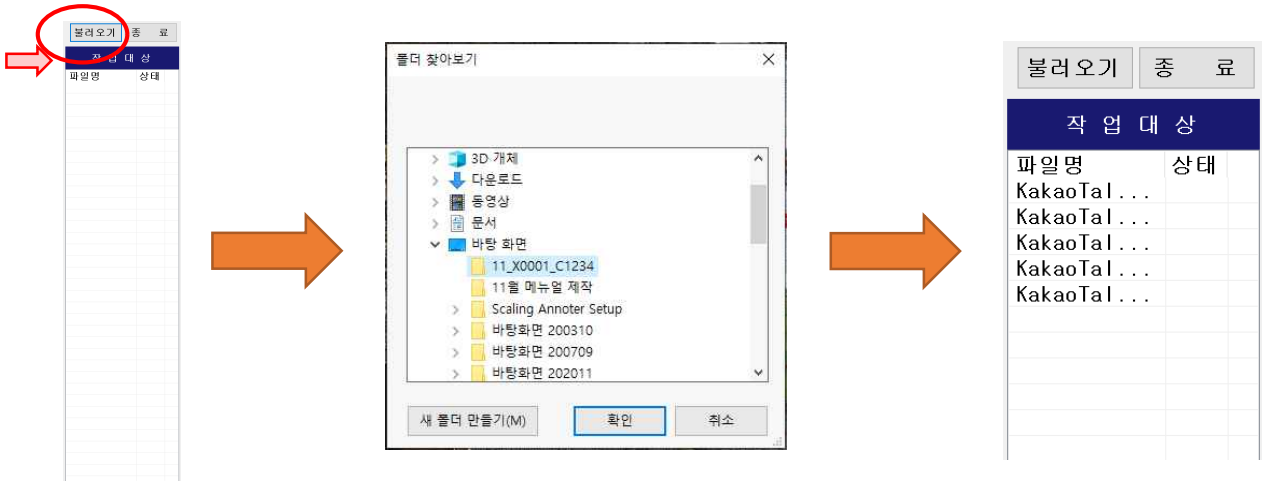
- 어노테이션정보 추가/수정

4) 이미지 정보 화면(Bounding Box)

- 라벨링(박싱)정보 추가/수정

다. 검수방법 (사용방법)

1) 작업 파일 불러오기



- 가) 파일 리스트 화면에서 [불러오기] 버튼을 클릭 합니다.
- 나) 「폴더 찾아보기」에서,촬영 원본 데이터 폴더를 지정한 후 [확인] 버튼을 클릭 합니다.
- 다) 파일 리스트 화면에 작업대상 데이터가 표시됩니다.

2) 작업 분량

- 가) 일별 작업분량 : 각 팀별 일별 작업분량을 할당합니다
- 나) 각 팀별 팀장은 팀별 작업량 배분 및 작업 결과물 수집 확인

3) 검수 체크리스트에 의한 검사



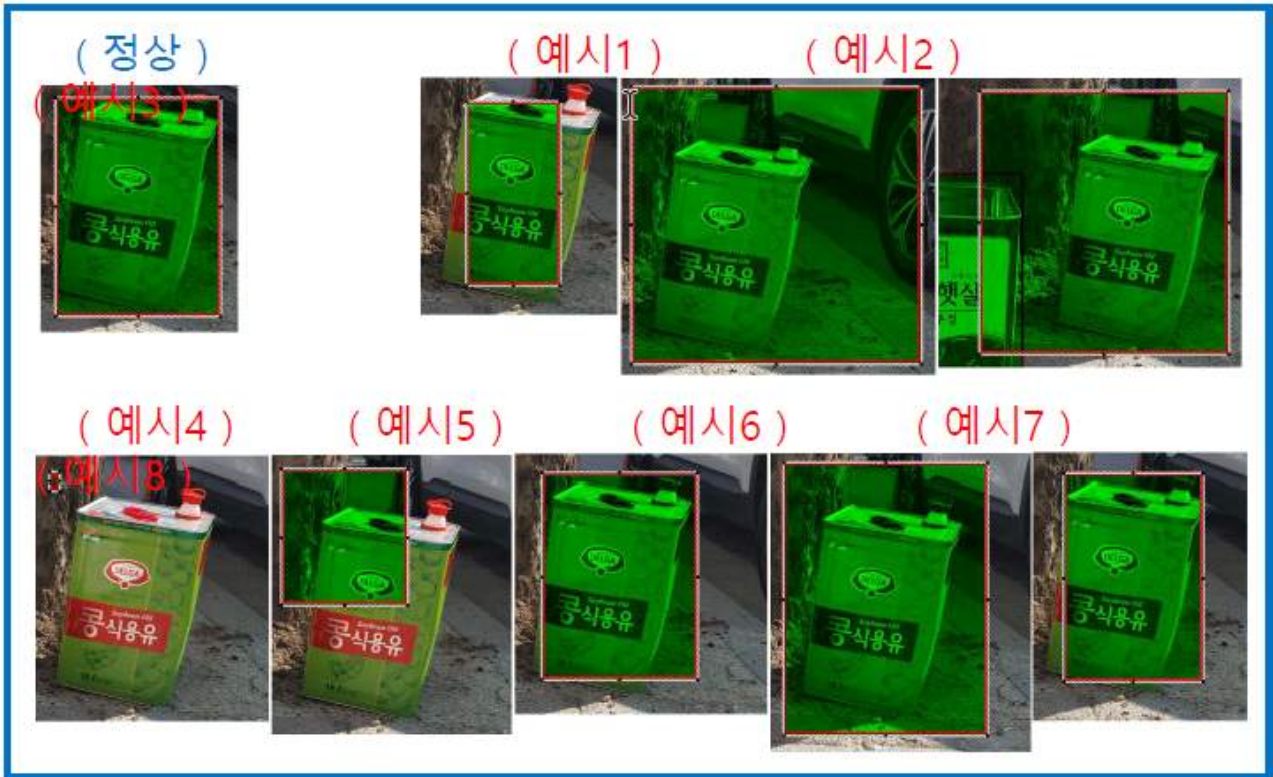
이미지 라벨링 작업은 위의 화면을 통하여 생성됩니다.

그림① : 해당 타켓에 대한 라벨링 영역이 설정된 상태

그림② : 이미지 라벨링 내용을 확인

그림③ : 해당 타켓의 바운딩 박스 내용이 일치하는지 확인

4) 검수 대상에 의한 검사 (바운딩 검사)



- (정상) 타겟의 테두리 외부에 라벨링 영역을 정확히 지정한 케이스 <합격>
- (예시1) 타겟의테두리내부에라벨링영역을지정한케이스 <불합격>
- (예시2) 라벨링영역한쪽구석에타겟을집어넣은케이스 <불합격>
- (예시3) 정확하지도않고, 다른타겟의영역을침범하여라벨링을지정한케이스 <불합격>
- (예시4) 실수로라벨링영역을지정하기전에마우스를놓아버린케이스 <불합격>
- (예시5) 라벨링영역을타겟의일부분만지정한케이스 <불합격>
- (예시6) 라벨링영역이타겟테두리외부에정확하지않게지정된케이스 <불합격>
- (예시7) 라벨링영역이타겟테두리외부를벗어나게지정한케이스 <불합격>
- (예시8) 라벨링영역을타겟의전체가아닌한쪽부분만기준으로지정한케이스 <불합격>

5) 검수 대상에 의한 검사 (라벨링 검사)

Image Info	
초기화	저 장 (F5)
과제구분	생활폐기물
수집구분	제작자료
촬영형태	이미지
주 / 야	주간
촬영장소	실외
촬영일시	2020-09-10 15:52:25
촬영위치	37.520926/126.71614
촬영자	C1234
해 상 도	4032*3024
초점거리	4.23mm
노출시간	1/2174sec
조리개값	1.16
ISO 감도	50
노출방식	normal program

ㄱ [코드 값참조]

코드구분	종류
과제구분	산업폐기물 / 생활폐기물
수집구분	제작자료 / 확보자료
주/야	주간 / 야간
촬영장소	실외 / 실내 / 스튜디오

라벨링 내용을 코드값을 참조하여 이미지와 일치하는지 확인하여
검사합니다

(참조) #.A3 – 바운딩 박스 [선택 구분 참조표]

분류 내용이 일치하는지 확인합니다.

<분류>

1)분류가 쉬운 종목

- 가구, 나무, 도기, 비닐, 스티로폼, 유리병, 의류, 자전거, 전자제품, 종이, 형광등

2)분류가 어려운 종목

(고철&캔)

- 특정 대상을 보관하기 위한 용도로 제작된 금속 케이스는 캔으로 분류
- 이외 철재는 고철로 분류

(페트병&플라스틱)

- 특정 대상을 보관하기 위한 용도로 제작된 플라스틱 케이스는 페트병으로 분류
- 이외 플라스틱 재질은 플라스틱으로 분류

<상세> : 타켓 의상 세품명 수기 입력상태 확인

<훼손 정도>

- 원형: 100% 근사치 원형 보존상태
- 일부훼손: 25% 미만 훼손 상태
- 상당훼손: 50% 미만 훼손 상태
- 완전훼손: 50% 이상 훼손 상태

5. 검수 기준 및 절차 (생활폐기물)

가. 생성데이터 검수 절차

- 생성데이터는 각 조의 조장을 통하여 할당 받은 분량에 대하여 검수 작업을 진행한다.

나. 생성데이터 검수 기준

- 정확성 시험 :
 - 1) 어노테이션 포맷 정확도 정합률 95% 이상
- 유효성 시험 :
 - 1) 생활 폐기물 인식 모델에 대한 인식률을 93% 이상

다. 라벨링 데이터의 검수 절차

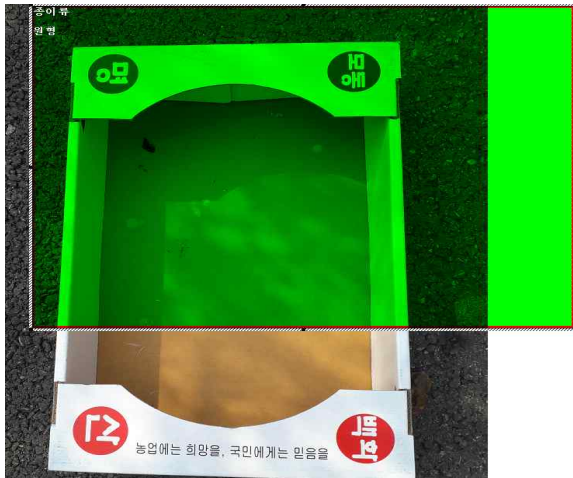
- 최종검수 : 크로수 체크에 의한 이중 검수

라. 라벨링 데이터의 검수 기준

1) Bounding Box의 정확도

① Bounding Box 오생성

- 작업 대상객체가 Bounding Box에 포함되지 않고 잘리거나 여백이 많이 생성된 경우 불합격



② Bounding Box 과생성

- 대상 객체가 훼손되어 흩어진 파편 또는 전자제품의 경우 전기선 등은 대상객체 벗어난 부분을 포함하지 않으며 포함된 경우 불합격

Don't



Do



Don't



Do



The screenshot shows a software interface with a bounding box around a bottle and a metadata panel on the right. The bounding box is labeled with a red '1'. The metadata panel has a '초기화' (Reset) button and a '자 장 (F5)' (Fullscreen) button. The metadata includes:

과제구분	산업폐기물
수집구분	채석자료
촬영형태	이미지
주 / 야	주간
촬영장소	실외
촬영일시	2020-09-26 15:34:49
촬영위치	37.5200383/126.7444947997
촬영자	C011
해 상 도	4032*3024
초점거리	4.32mm
노출시간	1/100sec
조리개값	2.52
ISO 감도	50
노출방식	normal program

Below the metadata is a 'Bounding Box' section with three columns: '추가 (F1)', '취소 (F2)', and '삭제 (F3)'. The '분 류' (Classification) dropdown is highlighted with a red '2' and shows '금속류' (Metals) selected. Other dropdowns include '상 세' (Details) and '훼손정도' (Damage level).

[실물이미지 페트병의 오분류 사례]

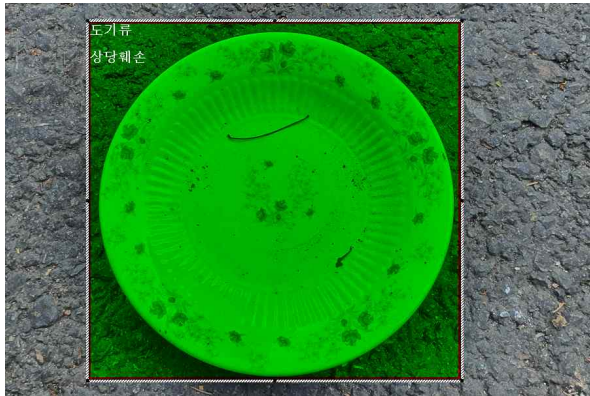
산업폐기물 X

생활폐기물 O

금속류 X

페트병류 O

- 훼손정도 기준과 상이하게 분류되어있는 사례에 대한 판별



[폐기물 오분류 사례]

*좌)도기류(상당훼손 → 원형) / 우)페트병류(원형 → 상당훼손)

2) 라벨링 데이터의 세부 검수 항목

라벨링 데이터의 검수는 11가지 항목으로 세분화 하여 check list 형식으로 검수 진행하며 검수자 및 작업자 추적을 위해 코드화 하여 관리를 한다.

- ① 이상없음 : 오류가 없는 정상 라벨링 데이터
- ② 여백많음 : 객체와 Bounding Box 사이 빈공간에 대한 허용기준에 부합되지 않는 데이터
- ③ 객체잘림 : 대상 객체가 Bounding Box내 포함되지 않고 잘린형태로 작업된 데이터
- ④ 객체누락 : 라벨링 대상 객체인데 Bounding Box 작업이 되어 있지 않은 데이터
- ⑤ 분할작업오류 : 한가지 형태의 대상객체를 불필요하게 나누어 Bounding Box 작업하여 Bounding Box 내 객체가 명확히 식별이 어려운 데이터
- ⑥ 군집작업오류 : 다수의 객체가 흩어진 상태인데 한 개의 Bounding Box 로만 처리하여 배경이 지나치게 많이 포함되는 데이터
- ⑦ 불필요 박스 : 대상 객체(폐기물)이 아닌 사물에 Bounding Box 작업된 데이터
- ⑧ 과제구분오류 : 생활 폐기물로 식별되는 객체가 산업폐기물로 잘못 분류된 데이터
- ⑨ 분류정보오류 : 식별되는 대상객체와 상이하게 폐기물 분류가 정의된 데이터
- ⑩ 훼손정도오류 : 훼손정도 속성정보가 식별되는 대상객체와 상이하게 정의된 데이터
- ⑪ 기타오류 : 위 항목에 포함되지 않는 기타 라벨링 오류 데이터

6. 검수 기준 및 절차 (산업폐기물)

가. 생성데이터 검수 절차

- 생성데이터는 각 조의 조장을 통하여 할당 받은 분량에 대하여 검수 작업을 진행한다.

나. 생성데이터 검수 기준

- 정확성 시험 :
 - 1) 어노테이션 포맷 정확도 정합률 95% 이상
- 유효성 시험 :
 - 1) 생활 폐기물 인식 모델에 대한 인식률을 90% 이상

다. 라벨링 데이터의 검수 절차

- 최종검수 : 크로tm 체크에 의한 이중 검수

라. 라벨링 데이터의 검수 기준

1) Bounding Box의 정확도

① Bounding Box 미생성

- 작업 대상 객체의 Bounding Box가 누락된 경우. (주로 복합촬영 이미지에서 발견)



② Bounding Box 과생성

- 양 끝에 다른 사물(나뭇가지, 밧줄등)이 있을경우 대상 객체의 양 끝부분이 잘리더라도 다른사물이 포함되지 않도록 대상 객체만 포함될수 있게 Crop형태의 Bounding Box 작업필요.

Don't



Do

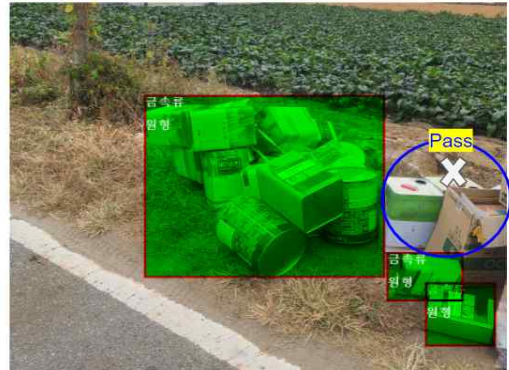


- 다른 종류의 폐기물이 섞여 포함되지 않도록 Bounding Box 작업 필요.
(한개의 Bounding Box내에는 동일한 폐기물만 포함)

Don't



Do



2) Class분류의 정확도

: 10종의 산업폐기물 분류기준에 맞게 각 Class가 올바르게 선택이 되었는지 여부 판별.

① 과제구분 및 ② 분류항목 오류

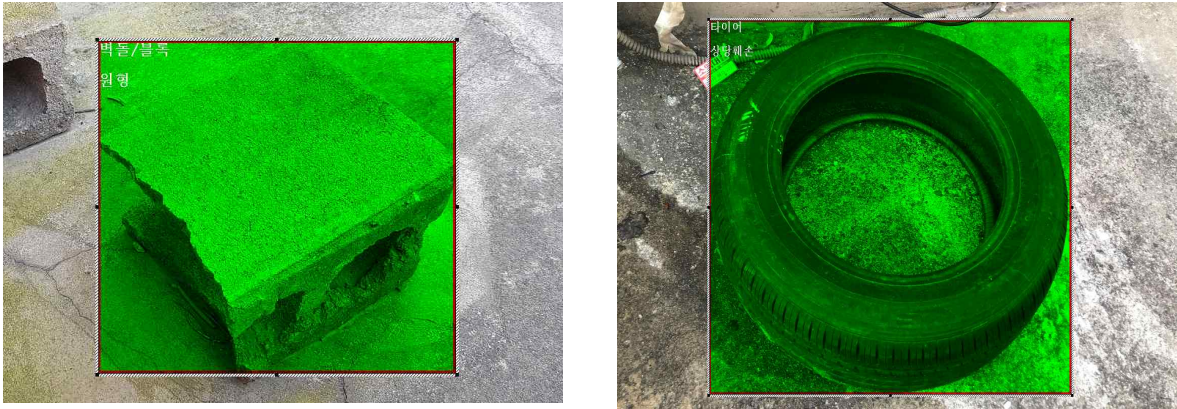
-실제 촬영된 이미지와 상이하게 구분이 되어 있는 사례에 대한 판별

과제구분	생활폐기물	
수집구분	제석치류	
촬영형태	이미지	
주 / 야	주간	
촬영장소	실외	
촬영일시	2020-10-13 13:31:22	
촬영위치	37.5112267/126.70630069972	
촬영자	C162	
해상도	4032*3024	
초점거리	5.40mm	
노출시간	1/119sec	
조리개값	1.69	
ISO 감도	160	
노출방식	normal program	
Bounding Box		
추가 (F1)	취소 (F2)	삭제 (F3)
분류	가구류	
상세		
훼손정도	원형	
분류	상세	훼손정도

[실물이미지 타이어의 오분류 사례]

③ 훼손정도의 오류

- 훼손정도 기준과 상이하게 분류되어있는 사례에 대한 판별



[폐기물 오분류 사례]

*좌)벽돌(원형 → 상당훼손) / 우)타이어(상당훼손 → 원형)

3)라벨링 데이터의 세부 검수 항목

라벨링 데이터의 검수는 11가지 항목으로 세분화 하여 check list 형식으로 검수 진행하며 검수자 및 작업자 추적을 위해 코드화 하여 관리를 한다.

- ① 이상없음 : 오류가 없는 정상 라벨링 데이터
- ② 여백많음 : 객체와 Bounding Box 사이 빈공간에 대한 허용기준에 부합되지 않는 데이터
- ③ 객체잘림 : 대상 객체가 Bounding Box내 포함되지 않고 잘린형태로 작업된 데이터
- ④ 객체누락 : 라벨링 대상 객체인데 Bounding Box 작업이 되어 있지 않은 데이터
- ⑤ 분할작업오류 : 한가지 형태의 대상객체를 불필요하게 나누어 Bounding Box 작업하여 Bounding Box 내 객체가 명확히 식별이 어려운 데이터
- ⑥ 군집작업오류 : 다수의 객체가 흩어진 상태인데 한 개의 Bounding Box 로만 처리하여 배경이 지나치게 많이 포함되는 데이터
- ⑦ 불필요 박스 : 대상 객체(폐기물)이 아닌 사물에 Bounding Box 작업된 데이터
- ⑧ 과제구분오류 : 산업 폐기물로 식별되는 객체가 생활폐기물로 잘못 분류된 데이터
- ⑨ 분류정보오류 : 식별되는 대상객체와 상이하게 폐기물 분류가 정의된 데이터
- ⑩ 훼손정도오류 : 훼손정도 속성정보가 식별되는 대상객체와 상이하게 정의된 데이터
- ⑪ 기타오류 : 위 항목에 포함되지 않는 기타 라벨링 오류 데이터