

데이터 수집 & 모델 훈련

지도 교수 : 이도훈

팀번호 : 13

팀명 : 디텍터

팀원 : 김윤재 201824446

이영한 201824552

23.07.28

1) 데이터 수집

2) 데이터 전처리

3) 모델 훈련

4) 다음 주 진행할 사항

데이터 수집

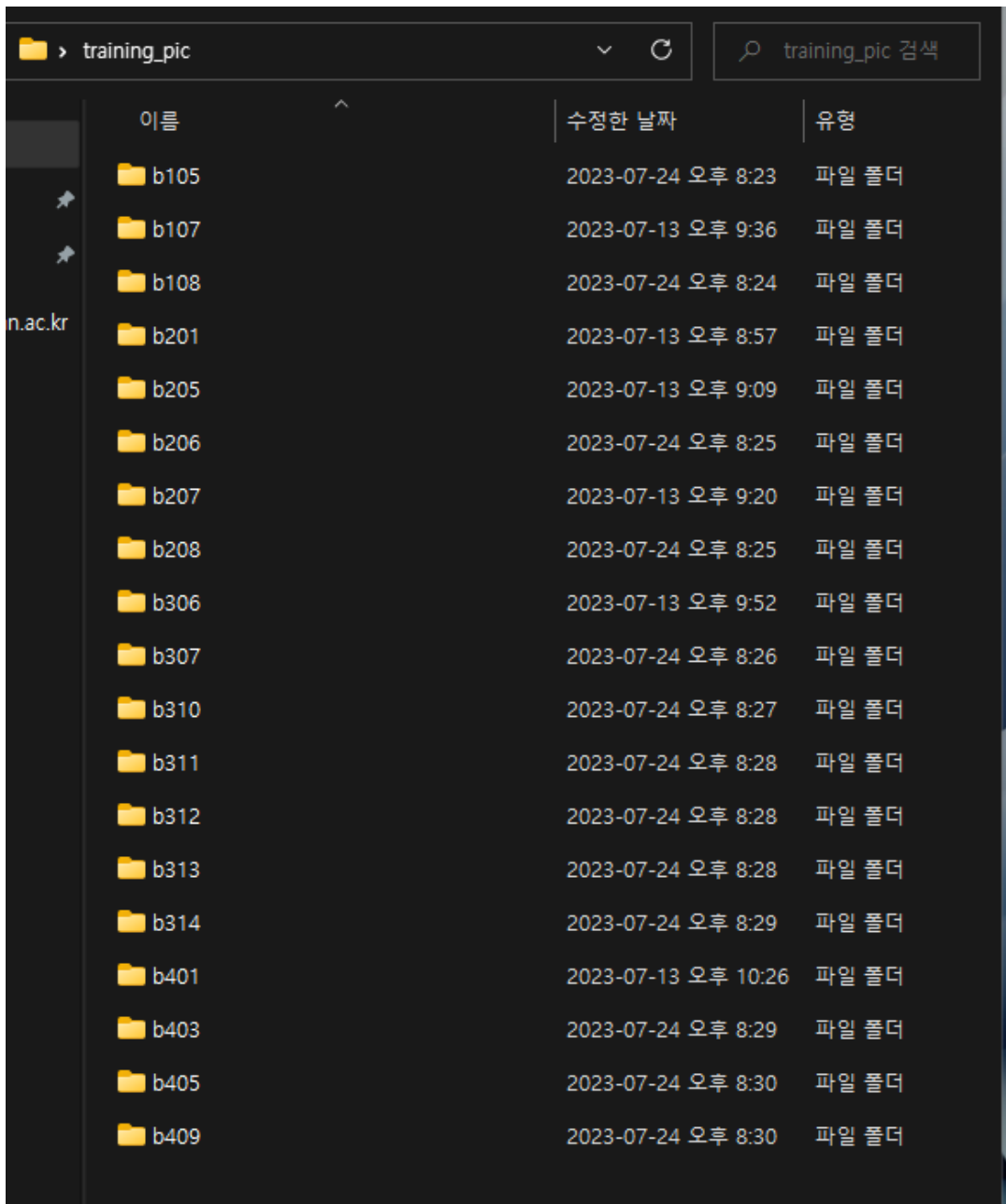


수집 완료

미수집

제외 대상

데이터 수집



이름	수정한 날짜	유형
b105	2023-07-24 오후 8:23	파일 폴더
b107	2023-07-13 오후 9:36	파일 폴더
b108	2023-07-24 오후 8:24	파일 폴더
b201	2023-07-13 오후 8:57	파일 폴더
b205	2023-07-13 오후 9:09	파일 폴더
b206	2023-07-24 오후 8:25	파일 폴더
b207	2023-07-13 오후 9:20	파일 폴더
b208	2023-07-24 오후 8:25	파일 폴더
b306	2023-07-13 오후 9:52	파일 폴더
b307	2023-07-24 오후 8:26	파일 폴더
b310	2023-07-24 오후 8:27	파일 폴더
b311	2023-07-24 오후 8:28	파일 폴더
b312	2023-07-24 오후 8:28	파일 폴더
b313	2023-07-24 오후 8:28	파일 폴더
b314	2023-07-24 오후 8:29	파일 폴더
b401	2023-07-13 오후 10:26	파일 폴더
b403	2023-07-24 오후 8:29	파일 폴더
b405	2023-07-24 오후 8:30	파일 폴더
b409	2023-07-24 오후 8:30	파일 폴더

- 동영상 촬영 후 사진으로 변환
- 건물 당 최소 60장 ~ 최대 100장
-> 전처리 시 augmentation을 통해
x3배 확보 가능

--	--



b206 1



b206_3



b206_4



b206_6



b206_8



b206_9



b206_99

데이터 전처리

데이터 전처리

The screenshot displays the Roboflow web interface for a dataset named 'PUSAN NATIONAL UNIV building'. The interface includes a sidebar on the left with navigation options: Universe Page, Upload, Assign, Annotate, Dataset (highlighted with 275 items), Generate, Versions (1), Deploy, and Health Check. The main area shows a 'Browse' view of the dataset. At the top, there's a search bar and filters for 'Split' (ALL, TRAIN, VALID, TEST), 'Classes' (All Classes), 'Tags' (All), and 'Sort By' (Newest). Below these, a grid of 32 image thumbnails is shown, each with a bounding box and a 'VALID' label. The images are of various buildings. At the bottom, there's a pagination bar showing '1 - 50 of 275' and a dropdown for 'Images per page' set to 50.

roboflow Projects Universe Documentation Forum

Browse How to Search

De-Select Select All 0 Images Selected

Search your images by subject (e.g. "brown dog" or "green apple") Search

Filename: Split: ALL TRAIN VALID TEST Classes: All Classes Tags: All Sort By: Newest

Show Public View >>

PUSAN NATIONAL UNIV

building

Object Detection

Universe Page

Upload

Assign

Annotate

Dataset 275

Generate

Versions 1

Deploy

Health Check

020_000003.jpg VALID

011_000002.jpg VALID

021_000003.jpg VALID

018_000001.jpg VALID

016_000024.jpg VALID

029_000016.jpg VALID

024_000030.jpg VALID

031_000024.jpg VALID

016_000028.jpg VALID

025_000011.jpg VALID

015_000012.jpg VALID

031_000012.jpg VALID

010_000018.jpg VALID

026_000002.jpg VALID

027_000017.jpg VALID

027_000037.jpg VALID

018_000012.jpg VALID

019_000020.jpg VALID

028_000016.jpg VALID

017_000001.jpg VALID

013_000002.jpg VALID

020_000025.jpg VALID

030_000012.jpg VALID

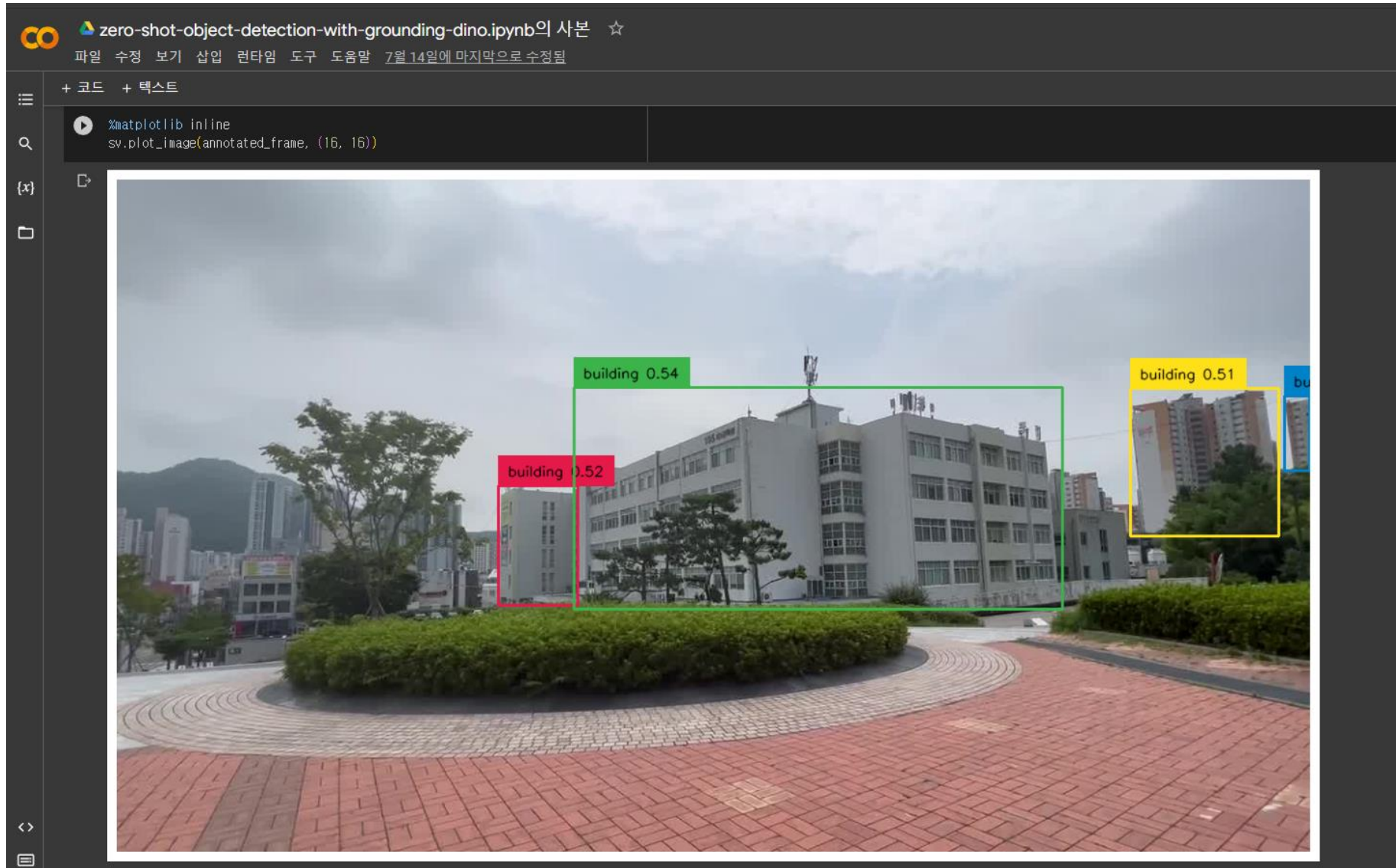
012_000003.jpg VALID

Images per page: 50

1 - 50 of 275

Roboflow

수동 라벨



Grounding Dino

자동 바운딩 박스



Train/Test Split

Training Set: 192 images

Validation Set: 43 images

Testing Set: 40 images



Preprocessing

Auto-Orient: Applied

Resize: Stretch to 640×640



Augmentation

Create new training examples for your model to learn from by generating augmented versions of each image in your training set.

Flip

Horizontal, Vertical

Edit

×

Rotation

Between -15° and +15°

Edit

×

Shear

±15° Horizontal, ±15° Vertical

Edit

×



Add Augmentation Step

Continue



Augmentation Options



Augmentations create new training examples for your model to learn from.

IMAGE LEVEL AUGMENTATIONS



90° Rotate



Crop



Grayscale



Hue



Saturation



Brightness



Exposure



Blur



Noise



Cutout



Mosaic

BOUNDING BOX LEVEL AUGMENTATIONS ?



Flip



90° Rotate



Crop



Rotation



Shear



Brightness



Exposure



Blur



Noise

Cancel

Roboflow 장점

- Resizing
- Augmentation
- Yolo와의 호환성

모델 훈련

- Yolo v8 모델 전이 학습
- 3개 건물 대상으로 진행
-> b105(기계관), b107(항공관), b108(기전관)
- 반복 학습(epochs)을 3번, 10번, 30번, 50번 진행한 4가지 학습 모델 비교

2

Train/Test Split

Here is how you split your images when you added them to the dataset:

Training Set

70%

192 images

Validation Set

16%

43 images

Testing Set

15%

40 images

Continue

⚙️ Rebalance

- 건물 당 60장 이상의 train, 13장의 valid, 13장의 test
- Trainingset 의 경우 Augmentation을 적용하면
x3 배 -> 576장의 training set

TRAIN / TEST SPLIT

Training Set

87%

576 images

Validation Set

7%

43 images

Testing Set

6%

40 images

https://github.com/PNU-Detector/AI/tree/main/training_practice/yolov8_basic_test

epochs=3 -> Accuracy : 75%

epochs=10 -> Accuracy : 97%

epochs=30 -> Accuracy : 97%

epochs=50 -> Accuracy : %

다음 주 진행 할 사항

1) 모든 건물 데이터 확보

2) 많은 건물을 대상으로 Yolo v8 훈련 진행하여 정확도 확인