

# P stage < Object Detection >

팀: 하나둘셋Net()

팀원: 공은찬, 곽민구, 김준섭, 김진용, 심용철, 오재석, 최현진

## 1. 프로젝트 개요

### ○ 프로젝트 주제

- 주어진 이미지에서 쓰레기를 인식, 분류하는 모델을 개발하여 “매립지 부족”과 같은 사회문제 해결에 기여

### ○ 학습데이터 소개

- 이미지 개수: 9754장 (학습: 4883, 평가: 4871), 크기: (1024,1024), Type: COCO
- Annotations: annotation id, bbox, area, category\_id, image\_id
- Class: General trash, Paper, Paper pack, Metal, Glass, Plastic, Styrofoam, Plastic bag, Battery, Clothing

### ○ 개발환경

- OS: Linux Ubuntu
- GPU: Tesla V100
- Runtime: Python 3.8.5
- Library: MMDetection, YOLOv5

### ○ 구현 내용

- Object detection 모델을 활용해 쓰레기를 탐지하고 10종류의 쓰레기 class를 분류하는 시스템을 구현
- MMDetection, detectron2 등 여러 Object detection library들을 활용
- Jupyter Notebook을 활용한 EDA

### ○ 기대효과

- 최종 모델이 다양한 환경에 버려진 쓰레기를 LB점수를 준수한 성능으로 분류했다. 따라서 직접 분류 이전 1차적 분류에 활용하여 인적자원의 절약에 활용할 수 있을 것으로 보임

## 2. 프로젝트 수행 절차 및 방법

- 1주차 : 개인적으로 여러가지 Test를 진행함으로써 각종 라이브러리 활용
- 2주차: 1-stage, 2-stage model별 성능 비교
- 3주차: 1-stage, 2-stage model ensemble 및 pseudo labeling적용

### 3. 프로젝트 수행 결과

- 제출 결과: Public LB: 0.668, Private LB: 0.653 (최종 순위 10위)
- EDA
  - i. Check Point
    - 각 image 별 bbox 분포의 불균형 확인
    - 각 쓰레기 class 별 분포 불균형 확인
    - bbox 별 image내 중심점 위치 확인
    - bbox 별 class 크기 확인
- Augmentation
  - i. Basic Augmentation
    - Resize(1024, 1024)
    - RandomFlip
    - Normalize
    - Pad
  - ii. Heavy Augmentation
    - Mosaic
    - ColorBrightness
    - Mix-up
    - PhotoMetricDistortion
- TTA(Test Time Augmentation)
  - Resize
  - RandomFlip
  - Normalize
- Ensemble
  - Seed: 모델의 팀원 다양성을 위해 seed를 랜덤으로 설정하고 모델 생성 후 결과 ensemble (0.650 -> 0.654)
  - Weight box fusion: 2-Stage 모델과 1-Stage 모델 ensemble 하여 성능 크게 향상 (0.654 -> 0.668)

## ○ 모델 선정 및 분석 결과

- 초기에 모델 선정을 위해 팀원별로 모델을 나누어 다양한 실험을 진행: Faster R-CNN, Cascade R-CNN, Retinanet, EfficientDet, YOLOv4, v5, ...

## ○ 최종 선정 모델

### i. Cascade R-CNN

Model	Backbone	Pretrained	Resolution	LB score
Cascade R-CNN	Swin Transformer	Imagenet	1024 * 1024	0.652

- Epochs: 29
- Batch : 2
- Optimizer : AdamW (lr=3e-5), LR Scheduler
- Neck: fpn
- Loss: CorssEntropy, LabelSmoothingLoss

### ii. Faster R-CNN

Model	Backbone	Pretrained	Resolution	LB score
Faster R-CNN	ResNet-152	Imagenet	1024 * 1024	0.560

- Epochs: 30, ealry-stop: 18
- Batch : 6
- Neck: fpn
- Optimizer : AdamW (lr=3e-4), LR Scheduler
- Loss: CorssEntropy, LabelSmoothingLoss

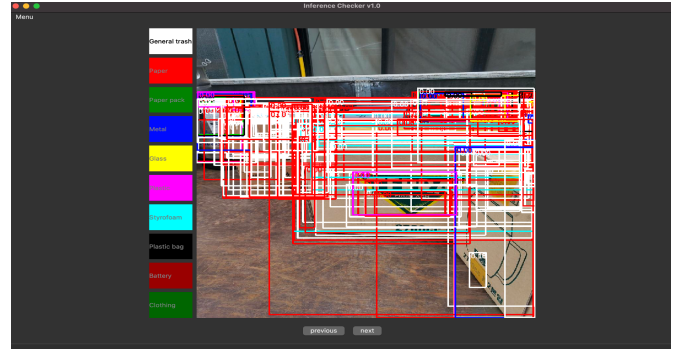
### iii. YOLOv5x6

Model	Backbone	Pretrained	Resolution	LB score
YOLOV5 X6	CSP-Darknet	COCO	1024 * 1024	0.474

- Epochs : 50
- Batch : 8
- Optimizer : SGD (lr=2e-2), LR Scheduler
- Loss: CorssEntropy, Focal Loss

- Tools(Python, PyQt5)

- Inference viz: 모델이 test data에 대한 추론 결과를 시각화 하는 프로그램 작성하여 결과로부터 insight
- Train set viz: 모델이 train data에 대한 시각화 프로그램을 작성하여 EDA 과정에서 좀 더 쉽게 insight



## 4. 자체 평가 의견

- 잘한점

- Inference Tool을 만들어 팀원들이 쉽게 모델의 예측 결과 사진을 쉽게 공유
- 각자 모델을 학습 시켜보며 결과를 표로 작성하여 공유
- 1-Stage 모델과 2-Stage 모델을 Ensemble 하여 성능 향상
- 코드 공유를 통해 팀원들이 좋은 모델을 쉽게 사용 가능하도록 함
- 다양한 관점으로 실험하기 위해 역할을 분배

- 아쉬운점

- Wandb를 사용하여 Augmentation이 어떻게 적용되는지 확인 못함
- 큰 모델에 Pseudo labeling을 적용해보지 못함
- 학습과 평가 모든 과정을 수행할 수 있는 pipe line 코드를 작성하지 못함
- 실험했던 내용(bbox 사이즈 다양하게 추가, augmentation, pseudo labeling, ensemble,...)에 대한 성능 개선으로 잘 이어지지 않아 모델 크기와 Hyperparameter를 중점적으로 실험