

HW#8

학과 : 소프트웨어학과

학번 : 2020039062

이름 : 김민주

Github link : <https://github.com/alswnsp411/ml-yolov8>

YOLOv8(You Only Look Once) Object Detection

- 이미지 전체를 $S \times S$ 그리드로 나누고 각각의 그리드 셀에서 객체 예측을 한 번에 처리하여 실시간 객체 검출에 매우 적합한 알고리즘이다.
- 단일 컨볼루션 신경망(CNN)을 사용하여 이미지의 여러 객체를 동시에 예측하여 한 번의 처리로 모든 영역을 동시에 예측한다.
- 이미지 전체를 한 번에 처리하기 때문에 이미지 내의 객체 간 관계나 배경 정보 등을 고려하여 더욱 정확한 검출 결과를 제공한다.
- 비 최대 억제 : 여러 개의 경계 상자가 동일한 객체를 나타낼 때, 신뢰도가 가장 높은 경계상자만 남기고 나머지는 제거하는 비 최대 억제를 사용하여 중복된 검출을 줄이고 정확성을 높인다.
- YOLOv8은 YOLO 시리즈의 최신 버전으로 속도가 아주 빠르고 성능이 우수하다.
- Ultralytics 라이브러리를 통해 쉽게 모델을 로드하고, 학습시키며, 예측할 수 있다.

과제 수행 과정

데이터 수집

핀터레스트에서 '봉지라면', '음료수', '컵라면' 3개 부류 관련 100개 이상의 이미지 수집

핀터레스트 :

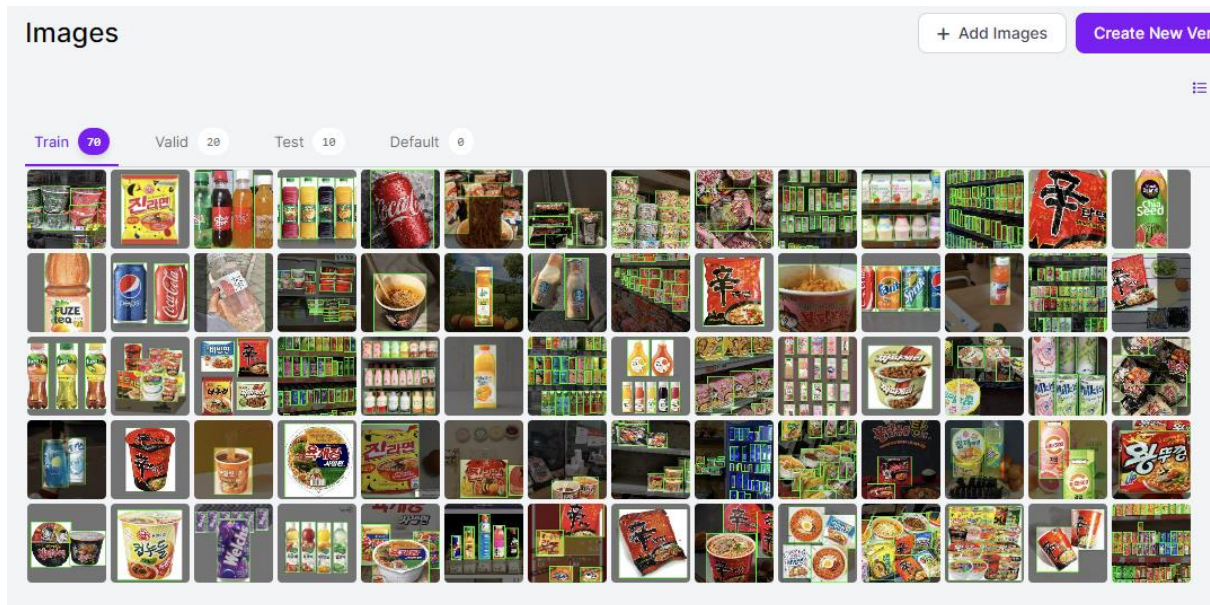
<https://kr.pinterest.com/search/pins/?q=%EB%B4%89%EC%A7%80%EB%9D%BC%EB%A9%B4&rs=typed>

데이터 세트 생성

수집한 이미지들을 '봉지라면', '음료수', '컵라면' 3개 부류에 대한 데이터 세트로 생성

Roboflow : [Sign in to Roboflow](#) 활용

Train set 70개, valid set 20개, test set 10개 지정

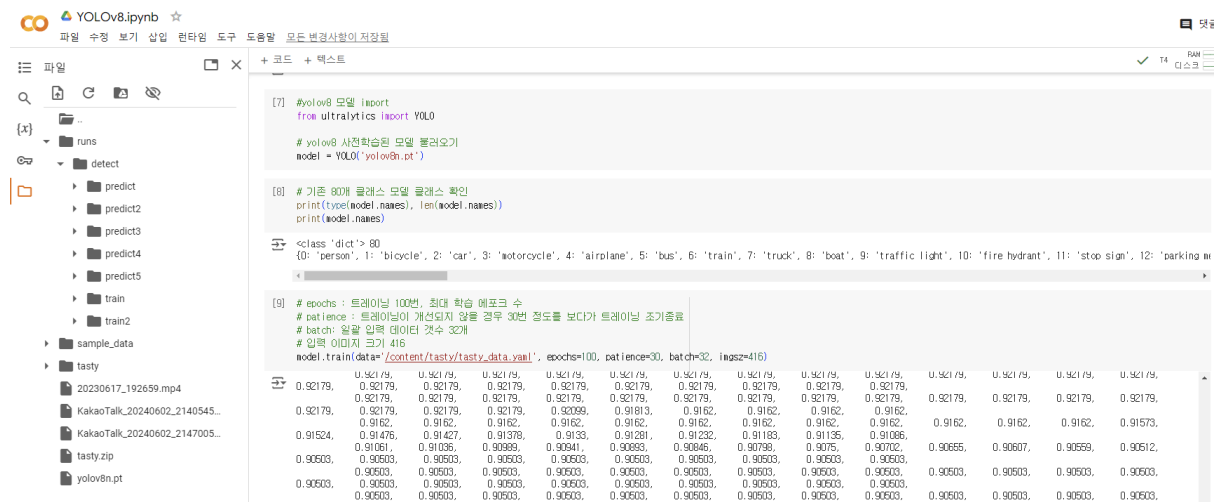


3개 부류 : '봉지라면', '컵라면', '음료수' labeling



Google colab을 사용하여 YOLOv8 모델을 직접 라벨링한 데이터셋을 사용하여 학습

```
# epochs : 트레이닝 100 번, 최대 학습 에포크 수
# patience : 트레이닝이 개선되지 않을 경우 30 번 정도를 보다가 트레이닝 조기종료
# batch: 일괄 입력 데이터 갯수 32 개
# 입력 이미지 크기 416
model.train(data='/content/tasty/tasty_data.yaml', epochs=100,
patience=30, batch=32, imgsz=416)
```



비디오 파일을 통해 3개의 부류에 대한 object detection 모델 학습이 잘 되었는지 확인

```
#학습된 모델 불러와서 비디오 파일 object detection
model = YOLO("/content/runs/detect/train/weights/best.pt")
model.predict(source = '/content/KakaoTalk_20240602_214700592.mp4',
show = True, save=True)
```



성능 평가

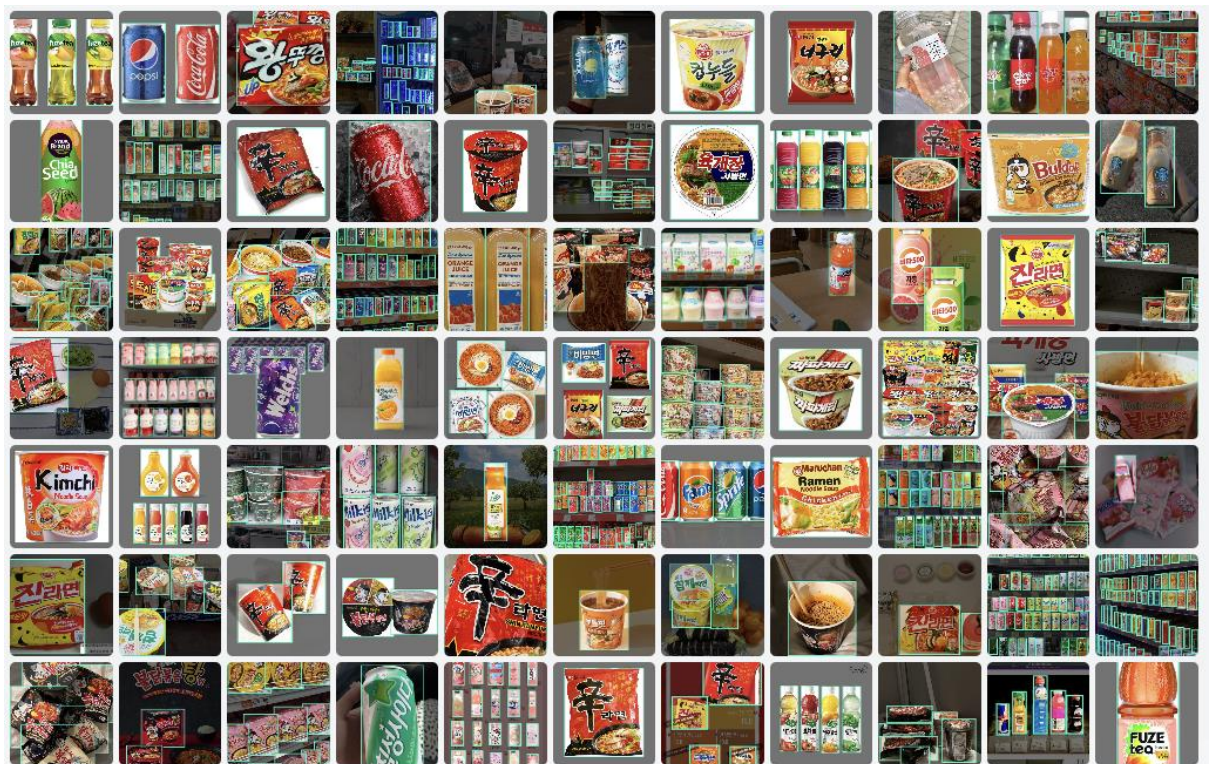
YOLOv8을 사용한 Object detection 결과 모든 객체가 '봉지라면' 으로만 탐지되었다. 이를 통해 모델의 성능이 좋지 않음을 확인할 수 있었으며, 이러한 결과에 대해 성능을 향상 시키기 위해 다음과 같은 이유를 고려해보았습니다.

1. 학습, 검증, 테스트 데이터 수의 부족

- A. 객체 탐지 모델을 학습시키기에는 100장의 이미지가 충분하지 않다.
- B. 봉지라면과 컵라면과 음료수의 표지가 다양하여 모델이 각 카테고리의 다양한 특성을 학습하기에는 데이터 수가 부족하다.

2. 데이터 세트 라벨링 문제

- A. Roboflow에서 데이터 세트를 만드는 과정에서 모든 클래스가 0번으로 통일되어 라벨링되었다. 라벨링 할 때에는 3개의 부류로 나누어 라벨링하였는데 데이터 세트 생성 후 라벨링된 결과를 확인한 결과, 모든 객체가 0번 클래스('봉지라면')로 라벨링 되어 있었다.
- B. 잘못된 라벨링을 통해 모델이 각 카테고리를 구분하지 못하고 학습한 것으로 예상된다.



3. 이미지 크기 불일치

- A. 학습시 이미지 크기를 416으로 설정했으나, 데이터 세트를 생성할 때는 640으로 설정했다.
- B. 추후 이미지 전처리 과정에서 이미지를 416크기로 조정하여 일관성있는 학습을 진행이 필요하다.

3

Preprocessing

[? What can preprocessing do?](#)

Decrease training time and increase performance by applying image transformations to all images in this dataset.

Resize

Fit (white edges) in 416×416

Edit

×



Add Preprocessing Step

Continue