

그룹명 : 66기 비전팀 (5)차 주간보고서

활동 현황

작성자	정성실	장소	온라인
모임일자	2022년 10월 16일 일요일	모임시간	17:00 ~ 18:30 (총 90 분)
참석자	여윤기, 이근호, 정성실	결석자	없음

학습 내용

학습주제 및 목표

1. YOLOv7 모델 학습 성능 개선
2. YOLOv7 모델 학습 Output 비교 및 TEST Dataset 수집
3. Confusion Matrix 이해

학습내용

◇ Yolov7 모델 학습 성능 개선 및 Output 결과 비교

1) Yolov7 모델 학습 성능 개선

지난주 YOLOv7 학습에 이어 모델의 성능을 개선하기 위해 학습에 여러 가지 변형을 주었다. 에폭값을 55와 300으로 다르게 학습을 시켰고, weights(yolov7.pt / 사전 설정된 weights를 사용하지 않고 모델 학습을 통해 weights 값을 저장)에 차이를 두었다.

에폭값을 55, weights를 yolov7.pt를 준 모델을 [1번 모델], 에폭값을 300, 사전 설정된 weights를 사용하지 않고 모델 학습을 통해 weights 값을 저장한 모델을 [2번 모델]로 정의하겠다.

```
%cd /content/drive/MyDrive/yolov7/  
!python train.py --batch 16 --cfg cfg/training/yolov7.yaml --epochs 55  
-data /content/drive/MyDrive/yolov7/data/data.yaml --weights 'yolov7.pt' --device 0 --name yolov7_result
```

※ img : 입력 이미지 크기 / data : 학습 데이터셋 설정 파일 경로 / cfg : 모델 구성 지정 / weights : 가중치에 대한 사용자 정의 경로 / hyp : 하이퍼파라미터)

2) YOLOv7 모델 학습 Output 결과 비교 및 TEST Dataset 수집

에폭값을 300으로 하고, weights를 사용하지 않고 모델 학습을 통해 weights값을 저장한 [2번 모델]에서 성능이 좀 더 우수한 결과를 얻었다. 선명한 이미지에서의 Output 결과값에서는 큰 차이는 없지만, 뒤에 있는 이미지와 희미한 이미지에서의 성능 차이는 존재했다.



[그림1.1 - 2번 모델 결과값]



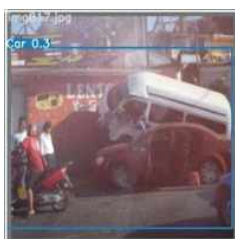
[그림 1.2 - 1번 모델 결과값]



[그림2.1 - 2번 모델 결과값]



[그림 2.2 - 1번 모델 결과값]



[그림3.1 - 2번 모델 결과값]



[그림 3.2 - 1번 모델 결과값]

하지만 일부 그림3.1, 그림3.2에서는 [모델 1번]에서만 Motorcycle을 인식하며 그림3.1에서는 이미지에 나온 차량 전부를 Car로 인식하는 차이를 보임을 알 수 있었다.

- 그 후 TEST Dataset을 만들어 어떤 데이터에 대해 못 찾았는지에 대한 분석을 진행했다. 각 클래스마다 5개의 TEST Dataset을 만들어서 Detect하는 과정을 진행했다.

- 한 가지 아쉬운 점은 한 사진에 클래스가 여러 개일 때, 평가 진행 시 Car 100개, Motorcycle 100개 등의 비슷한 데이터 셋의 수를 놓아야했지만 TEST Dataset의 Car가 압도적으로 많았다는 점이 아쉬운 점(Data Class Imbalance문제)이었다.

3) Confusion Matrix 이해

- Confusion Matrix란?

분류 모델을 학습하는 것의 목적은, 주어진 데이터를 의도에 맞게 잘 분류해내기 위한 것이다. 그렇다면 이러한 모델을 평가하는 기준이 필요할 것이다.

모델을 평가할때는 모델이 얼마나 정밀한지, 얼마나 실용적인 분류를 해내었는지, 얼마나 정확한 분류를 했는지를 평가해야 한다.

이러한 내용들을 모두 포함하고 있는 것이 Confusion Matrix이다.

레이블 0,1을 가진 데이터를 분류한다고 할 때 관심 범주를 1이라고 한다.

- True Positives : 1인 레이블을 1이라 하는 경우를 True Positives라고 한다. -> 관심 범주를 정확하게 분류한 값.
- False Negatives : 1인 레이블을 0이라 하는 경우를 False Negatives라고 한다. -> 관심 범주가 아닌 것으로 잘못 분류함.
- False Positives : 0인 레이블을 1이라 하는 경우를 False Positives라고 한다. -> 관심 범주라고 잘

못 분류함.

- True Negatives : 0인 레이블을 0이라 하는 경우를 True Negatives라고 한다. -> 관심 범주가 아닌 것을 정확하게 분류.

이 4가지 정보를 바탕으로 3가지의 척도를 평가할 수 있다.

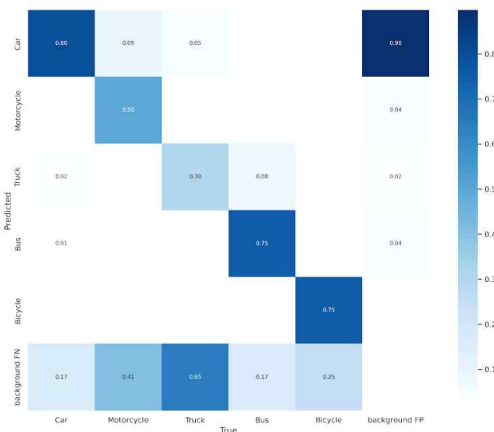
첫번째 척도로, **정확도(Accuracy)**가 있다. 정확도는 1을 1로, 0을 0으로 정확하게 분류해낸 것을 의미한다. 모델이 얼마나 정확한지를 평가하는 척도라고 할 수 있다.

두번째 척도로, **정밀도(Precision)**가 있다. 모델을 통해 1이라고 분류해낸 그룹 A가 있을 때, 모델이 얼마나 믿을만한 정도로 A를 만들어 냈는지를 평가하는 것이다. 어부가 그물을 던져 물고기를 잡을 때, 그물 안에 1이라는 물고기가 얼마나 들어있는가에 대한 척도이다.

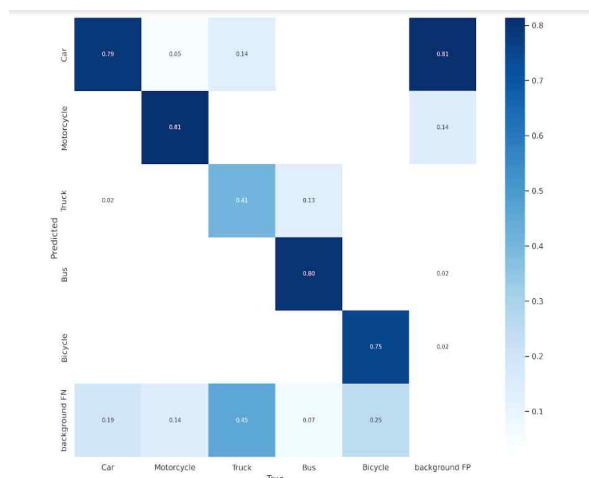
세번째 척도로, **재현도(Recall)**가 있다. 정밀도와 비교되는 척도로서, 전체 예측 중에 TP가 얼마나 많은가에 관한 것이다. 관심 영역만을 얼마나 추출해냈는지를 말한다. 즉, 모형의 실용성과 관련된 척도라고 할 수 있다.

일반적으로 Precision, Recall은 상충되는 수치라기 보다는, 독립적으로 평가되는 수치이다. 두 가지를 함께 평가하는 수치로 **F-score**라는 것이 있는데, Precision과 Recall의 곱을 둘의 합으로 나눈 뒤, 2를 곱해준 수치이다.

이렇게 총 4가지의 평가 척도를 각각 상황에 맞게 사용해야 한다.



[그림4.1 - 2번 모델 결과값]



[그림4.2 - 1번 모델 결과값]

그림4.1는 2번 모델을 학습시키고 나온 Confusion Metrix에서 Car는 80%, Motorcycle은 50%, Truck은 30%, Bus 85%, Bicycle 75%의 확률로 정답을 찾은 점을 확인할 수 있다. Background FN은 정답 레이블을 찾지 못한 값을 의미한다.

그림4.2는 1번 모델을 학습시키고 나온 Confusion Metrix에서 Car는 79%, Motorcycle은 81%, Truck은 41%, Bus 80%, Bicycle 75%의 확률로 정답을 찾은 점을 확인할 수 있었다.

그룹 운영 기록사항

팀 성찰

GPU 성능이 부족하다보니 모델 학습 시 시간이 매우 오래 걸린다는 것에 늘 어려움이 있다.
또한 개인이다보니 Dataset에 한계가 있어서 비슷한 Dataset의 갯수로 학습을 진행하지 못해 특정 Class값이 압도적으로 많았다는 점(Data Class Imbalance)에 아쉬움이 있다.

다음 모임 계획

모임일자	2022년 10월 23일 일요일	모임시간	17:00~19:30 (총 120 분)
역할분담	1. Confusion Metrix 분석 (accuracy, recall, mAP, f1score) 2. lstm을 빼고 기본 자연어가 무엇인지 알아오기(텍스트 모델 찾기)	장 소	온라인