

Level-3 P-stage : Model Optimization

| Wrap-UP Report by Team AI-TEEN(CV18)

1. 프로젝트 개요

1-1. 프로젝트 주제

분리수거 로봇에 가장 기초 기술인 쓰레기 분류기를 만들면서 실제로 로봇에 탑재될 만큼 작고 계산량이 적은 모델을 만들어볼 예정입니다. 기존에 쓰레기 분류기 모델이 있다는 가정하에 해당 모델에 경량화 기법을 도입하여 동일 파라미터 대비 고성능인 모델, 동일 성능 대비 더 적은 파라미터를 갖는 모델 등을 찾습니다. 소기의 목적 달성을 위해 AutoML 라이브러리인 **Optuna** 를 사용하여 search space를 설정하여 원하는 모델을 찾아볼 예정입니다.

1-2. 활용 장비 및 협업 도구

- AI Stages 서버 내 V100, [Github](#), [Notion](#), WandB

1-3. 프로젝트 구조

- 학습 데이터 입력 및 출력
 - 입력 : COCO format의 TACO Dataset 쓰레기의 이미지
 - 출력 : 입력된 쓰레기 이미지에 대한 분류 성능(f1 score)과 추론 속도(submit time)
- 사용한 최종 모델
 - MobileNetV3 구조에서 레이어를 축소한 모델

1-4. 기대 효과

현업에서 마주하는 서비스 요구사항들은 각 사항에 따라 제약 조건들이 존재하고 이는 대체로 추론 시간을 줄이면서 모델 성능 유지 혹은 높이는 데 중점을 둡니다. 이번 대회에서 주어진 시간 동안 조건에 맞는 모델을 찾는 방법론을 익혔고 모델 경량화 기법을 최종 프로젝트에서 적용할 수 있습니다.

2. 프로젝트 팀 구성 및 역할

이름	역할
김서기	MobileNetV3 실험
김승훈	MobileNetV3 실험
배민한	MobileNetV3 실험
손지아	MobileNetV3, ShuffleNetV2 실험, WandB 설정
이상은	MobileNetV3 실험
조익수	MobileNetV3 실험

3. 프로젝트 수행 절차 및 방법

3-1. 프로젝트 사전 기획 및 문제 정의

- **TACO** 데이터셋은 segmentation/detection task를 수행하기 위해 제작되었기 때문에, 이를 classification task로 변환하기 위해 bounding box를 crop 한 이미지 사용
- 학습 데이터셋
 - 총 20,851 장의 .jpg format images
 - 6개 카테고리 **Metal**, **Paper**, **Paperpack**, **Plastic**, **Plasticbag**, **Styrofoam**

- class imbalance 문제, image's width와 height 모두 제각각
- 외부데이터셋 사용 금지, 기학습 모델의 가중치 사용 허용

• 평가 데이터셋

- 총 5,217 장의 .jpg format images
 - private: 2,611개
 - public: 2,606개
- 학습 데이터셋과 동일한 6개 카테고리

• 평가 방법

최종 score는 모델의 **f1-score**와 모델의 **Submit time**를 통해 계산

- $score = 0.5 * score_{submit\ time} + score_{F_1}$
 - $score_{submit\ time} = \frac{thismodel_{submit\ time}}{baseline_{submit\ time}}$
 - $score_{F_1} = \text{sigmoid}(20 * (baseline_{F_1} - thismodel_{F_1}))$

3-2. 프로젝트 수행 절차 및 과정

- 1주차 : 강의 수강 및 베이스라인 코드 실습
- 2주차 : 모델 경량화 실험 및 Wrap-up report 작성

4. 프로젝트 수행 결과

4-1. MobileNetV3의 레이어 및 파라미터 개수, 이미지 크기, Epoch 변화에 따른 성능 변화

모든 실험에서 RandAugment (n_level=2) 적용

#	레이어 개수	파라미터 개수	이미지 크기	Epoch	Score / F1 / Time
1	179	2,127,374	224	100	1.5460 / 0.6880 / 84.9850
2	163	1,741,254	224	100	1.7950 / 0.6530 / 92.9260
3	79	294,854	156	100	1.2950 / 0.6870 / 63.0780
4	79	294,854	156	150	1.1732 / 0.7005 / 58.3412
5	79	294,854	156	200	1.1676 / 0.7012 / 58.1225

4-2. 최종 선정 모델

Key	Value	비고
Backbone	MobileNetV3	[1]
Optimizer	SGD	
Init Learning Rate	0.1	
LR Scheduler	OneCycleLR	
Loss	Custom Loss (<u>fixed</u>)	
Image Size	156	
Image augmentation	RandAugment (n_select=2)	[2]
Batch Size	64	
Epochs	200	
Val Ratio	0.1	
FP16	True	

[1] MobileNetV3의 파라미터를 축소해서 사용

[2] 학습 데이터에만 적용

4-3. 대회 결과

해당 대회에서는 대회 목표 성능을 초과하는 모델을 구축하는 것이 목표였으며, 앞서 구축된 모델을 바탕으로 대회 목표 성능을 달성할 수 있었습니다.

- **최종 제출 모델** : Public LB에서 가장 성능이 좋았던 모델

	Target LB	Public LB	Private LB
Score	1.1950	1.1676	1.1804
F ₁ score	0.7000	0.7012	0.6986
Time	60.0000	58.1225	58.1225

5. 자체 평가 의견

- 김서기 : 프로젝트랑 같이 진행하는 대회라서 프로젝트를 중점으로 두고 대회를 진행했었다. 대회를 중점으로 두지 않았다보니, 최적화 기법들을 baseline 단계에서만 적용 해 봤었다는게 아쉬웠다.
- 김승훈 : 강의에 등장하는 수식 이해도 어려웠고 프로젝트와 병행하다보니 이도저도 아닌 대회 참여가 된 느낌이 들기는 하지만 최적화 대회를 하면서 모델 최적화로 좋은 성능, 빠른 추론 속도가 나는 것을 보면서 수행 중인 프로젝트에 적용해 볼 수 있지 않을까 하는 생각을 갖게되어 만족스럽다.
- 배민한 : baselinecode를 이해하는 것에 시간이 많이 소요되어 적극적인 참여를 하지 못한 점이 아쉽다. 또한, 네트워킹데이 준비에 비중을 두어 해당 부분에 있어 시간분배를 제대로 하지 못한 부분이 있다.
- 손지아 : MobileNetV3를 바탕으로 레이어와 초매개변수를 수정하면서 목표 성능에 도달하였다. 학습에 사용되는 데이터 수를 늘리기 위해 검증셋 비율을 너무 적게 설정한 것이 아닌가 싶다. 검증셋에서의 성능을 바탕으로 모델을 저장하였기에 검증셋 비율을 잘 맞추면 지금보다 높은 F1 스코어를 얻었을 것 같다. Public LB 상에서 파라미터가 15,000 가량 더 많은 모델이 속도는 살짝 느려도 F1 점수가 더 높았다. 예상은 했지만 Public LB 상에서 성능이 좋지 않기에 최종 제출하지 않았는데 이 점 또한 아쉽다.
- 이상은 : 생소한 분야였고, 분야에 따라 목적이나 중요도가 다를음을 새삼 느낄 수 있었다.
- 조익수 : 부스트캠프 초반에 배웠던 개념들이 재등장해서 복습할 수 있는 기회로 삼기로 했다. 시간 관리가 중요하다는 점을 새삼 깨달았다.