

# 부스트캠프 AI Tech Lv.1 P-Stage

## Mask Image Classification

1주 - One AI

### <프로젝트 개요>

COVID-19 Pandemic 상황 속 마스크 착용 유무 판단 시스템 구축

### <프로젝트 목표>

- 전체 4500명의 데이터 중 60%의 train set을 학습, F1 Score(macro) 통해 평가 및 순위 매김.
- 마스크 착용 여부, 성별, 나이총 세가지 column이 존재하여 총 18개의 class로 구성
- 사진이 주어지면 이를 각 class에 맞게 분류하는 모델을 만드는 것이 목표

### <팀 구성 및 역할>

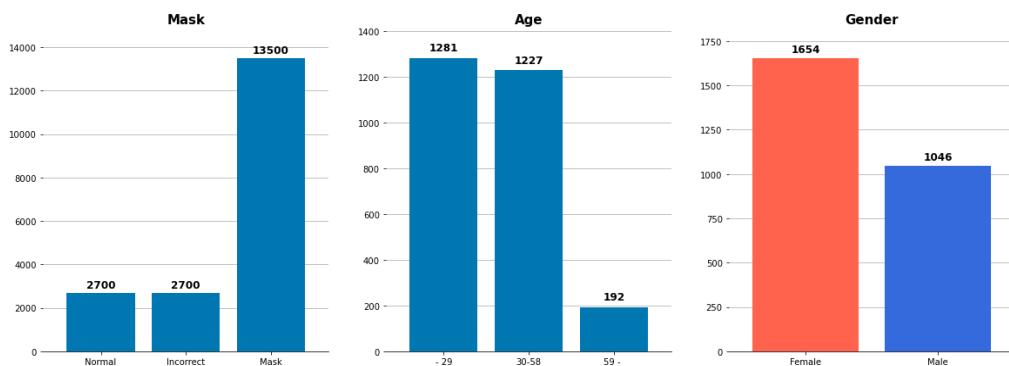
- 정유석(팀장) : 모델 개발
- 박상하(팀원) : 논문 리뷰 및 자료 조사
- 박진형(팀원) : 모델 개발 및 Wandb
- 백운경(팀원) : 자료 조사 및 모델 학습
- 이상은(팀원) : 모델 개발
- 이홍규(팀원) : 모델 개발 및 다양한 실험 진행

### <프로젝트 수행 절차 방법>

Date	To Do	Date	To Do
08/23 (월)	<u>서버 생성</u> , EDA	08/30 (월)	crop, 논문 리뷰
08/24 (화)	Data Augmentation	08/31 (화)	overfitting 잡기, wandb sweep
08/25 (수)	Dataset, Data generation, Model	09/01 (수)	overfitting 잡기, label, kfold
08/26 (목)	마지막 개인 제출, Training & Inference	09/02 (목)	<u>최종 제출</u>
08/27 (금)	베이스라인 코드, loss	09/03 (금)	python code
08/28 (토)	모델 확정	09/04 (토)	리포트 작성

### <프로젝트 수행 결과>

#### • EDA



- Mask : [5:1:1] 모두 같은 비율, 마스크 색이 다 다르고, 손수건 착용, 턱스크, 코스크, 눈스크 등 다양한 사진 존재
- Age : 30, 59 비율 비슷, 60대 비율 매우 적음, data imbalance 심하다.
- Gender : 여성 데이터가 많음, data imbalance까지는 아니다.
- Insight through EDA : Age가 특히 불균형이 심해 Age 불균형 해소에 집중할 필요가 있어 보인다.

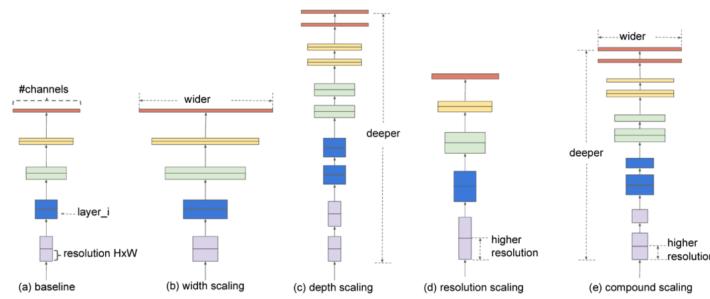
#### • Data Augmentation

Perspective, Rotate(limit=20), Normalize, RandomBrightness, HueSaturationValue, RandomContrast, RandomGridShuffle, GaussNoise 사용

#### • K-fold

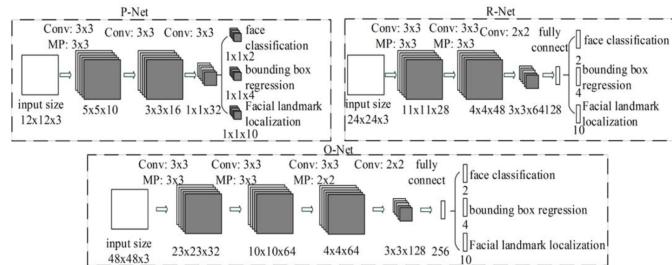
사람별 돌려 overfitting 확인 후, best epoch으로 label별로 돌리기.

- EfficientNet



- 목표 : 높은 성능을 가진 CNN 기반의 모델을 만든다.
- 방향 : Baseline의 EfficientNet B3, 4 코드에서 모델의 너비, 깊이, 입력 해상도의 크기를 균형 있게 증가시킴으로써 높은 성능과 효율성을 가진 모델을 만든다.

- MTCNN



- 목표 : 이미지에서 얼굴을 검출한다.
- 방향 : 얼굴 검출과 얼굴 특징점 검출의 연관성을 증폭시키기 위해서 Deep Cascade Multi-Task Framework를 사용한다.

- 모델 학습: 모델을 mask, age, gender로 나눠 각각 학습

Model	Mask (Batch Size / Learning Rate / epoch)	Age (Batch Size / Learning Rate / epoch)	Gender (Batch Size / Learning Rate / epoch)	Acc/F1	추가사항
EfficientNet B4	32 , 3e-4 , 10	32 , 1e-4 , 10	32 , 3e-4 , 10	78.20/0.726	Crop: MTCNN 적용, Optim: Adam
EfficientNet B4	32 , 3e-4 , 10	64 , 1e-4 , 10	32 , 3e-4 , 10	79.68/0.746	age: batch size 변경, Scheduler: StepLR => ReduceLROnPlateau로 변경
EfficientNet B4	32 , 3e-4 , 10	64 , 1e-4 , 15	32 , 3e-4 , 10	80.50/0.752	age: epoch 변경
EfficientNet B4	32 , 3e-4 , 6	64 , 1e-4 , 13	32 , 3e-4 , 6	77.33/0.705	age: Focal Loss / Random Sample 적용
EfficientNet B3	64, 3e-4, 9	64, 1e-4 , 4	64, 1e-4 , 8	76.06/0.680	B4 => B3 / mask, gender: batch size, lr 변경 / Input Image Size 312로 변경
EfficientNet B3	64, 1e-4, 10	64,1e-4,4*5	64,1e-4,10	81.19/0.753	mask lr 변경, age 반복, crop: 30 => 15 간격 줄임
EfficientNet B3	64, 1e-4, 2*5	64,1e-4,4*5	64,1e-4,4*5	81.76/0.761	mask, gender: epoch 변경, 반복, age 59, 60 한 클래스로 둑음
EfficientNet B3	64, 1e-4, 2*5	64,1e-4,4*5	64,1e-4,4*5	81.62/0.763	valid[4]로 통일

Model	acc / F1	Data augmentation	img_size	problem & solution (진행 중) & Hyperparameter	problem & solution about the renewed result	비고
ResNet 50	53%, 0.43	Resize(224, 224), ColorJitter, Normalize	224 * 224	문제점: class 불균형 해결 방안: Stratified K-fold 5회	부족한 점: epoch 를 3회만 수행, 모델 하나로 class 18종 분류 시도 개선 방안: epoch 늘리고, class 를 세 파트로 구분	직접 구현
EfficientNet B4	77%, 0.722	CenterCrop(380, 380), Normalize	380 * 380	문제점: label imbalance 해결 방안: stratified K-fold 4회, class_weight 부여 Loss: CE loss Optim: Adam Scheduler: ReduceLROnPlateau	class 파트 구분: Mask, Gender, Age epoch 10회	
EfficientNet B3 for Mask, Gender	81.095%, 0.758	RandomBrightnessContrast, CenterCrop(350, 300), Normalize	350 * 300	Loss : CE loss Optim : ADAM	Stratified K-fold 4회 / multi sample dropout 추가로 일반화 시도	3번째 아키텍처 Ensemble
EfficientNetB3+ViT Base Patch16_24 for Age		RandomBrightnessContrast, CenterCrop(350, 300), Normalize + Resize(224, 224)	224 * 224	Loss: Focal loss Optim: Adam	Focal loss로 label imbalance 완화 multi sample dropout 추가로 일반화 시도	
EfficientNet B3	76.6825% / 0.697	CenterCrop(380,350) Normalize Resize(312, 312)	312 * 312	Optim: Adam Loss: CE loss LabelSmoothing	cutmix, label smoothing을 통한 오버피팅 방지	

- Best Model

1. Data Augmentation

- a. ALL: Perspective, Rotate(limit=20), Resize(312, 312), Normalize
- b. Mask: RandomBrightness, HueSaturationValue, RandomContrast
- c. Age: RandomGridShuffle, GaussNoise
- d. Gender: GaussNoise

2. 5\_fold, label별

3. Model : EfficientNet B3

- a. Mask : CrossEntropy Loss / 64 / 1e-4 / 2 \* 5
- b. Age : Focal Loss / 64 / 1e-4 / 4 \* 5
- c. Gender : CrossEntropy Loss / 64 / 1e-4 / 4 \* 5

4. Optimizer : Adam

5. Scheduler : ReduceLROnPlateau(mode='min', factor=0.2, patience=3)

6. Input Image Size : (312, 312)

<자체 평가 의견>

1. 달성도 및 완성도 : 전체적으로 80%의 성능을 가졌으며 최종 8등을 하면서 무난한 성적을 거두었다..
2. 팀이 잘한 부분
  - a. 모든 팀원이 대회를 포기하지 않고 공부했던 점
  - b. 모델 뿐만 아니라 다양한 기법을 찾아서 시도했던 점
  - c. 이미지 분류와 관련된 다양한 논문 분석했던 점
3. 팀이 아쉬운 부분
  - a. 리눅스에 대한 공부가 부족했던 점
  - b. AutoML과 관련해서 많이 공부하지 못했던 점
  - c. 체력을 잘 조절하지 못했던 점

## <Reference>

1. Paper
  - a. CutMix : <https://arxiv.org/pdf/1905.04899.pdf>
  - b. ResNet : <https://arxiv.org/pdf/1512.03385.pdf>
  - c. EfficientNet : <https://arxiv.org/pdf/1905.11946.pdf>
  - d. MTCNN : <https://arxiv.org/pdf/1604.02878.pdf>
  - e. Vision Transformer : <https://arxiv.org/pdf/2010.11929.pdf>
  - f. Focal Loss : <https://arxiv.org/pdf/1708.02002.pdf>
2. Fine Tuning
  - a. <https://jungnamgyu.tistory.com/34>
3. Image Classification
  - a. [https://pytorch.org/tutorials/beginner/finetuning\\_torchvision\\_models\\_tutorial.html](https://pytorch.org/tutorials/beginner/finetuning_torchvision_models_tutorial.html)
  - b. <https://paperswithcode.com/sota/image-classification-on-cifar-100>