



Wrap-Up Report

1. 프로젝트 개요

1) 프로젝트 주제 : 마스크 착용 상태 분류

- 카메라로 촬영한 사람 얼굴 이미지의 마스크 착용 여부를 판단

2) 프로젝트 개요

- 코로나 전파 경로를 원천 차단하기 위해 올바른 마스크 착용이 선행
- 하지만 마스크 착용상태를 검사하기 위해서는 추가적인 인적자원이 필요
→ 카메라로 비춰진 사람 얼굴 이미지를 통해 마스크 착용 상태를 자동으로 판단하는 시스템이 필요
- 마스크 착용 상태 뿐만 아니라, 성별, 나이까지 분류하는 모델을 목표
 - 마스크 착용상태, 성별, 나이에 따라 총 18개 class로 분류

for Class 1	Mask	Gender	Age
0	Wear	Male	< 30
1	Wear	Male	>= 30 and < 60
2	Wear	Male	>= 60
3	Wear	Female	< 30
4	Wear	Female	>= 30 and < 60
5	Wear	Female	>= 60
6	Incorrect	Male	< 30
7	Incorrect	Male	>= 30 and < 60
8	Incorrect	Male	>= 60
9	Incorrect	Female	< 30
10	Incorrect	Female	>= 30 and < 60
11	Incorrect	Female	>= 60
12	Not Wear	Male	< 30
13	Not Wear	Male	>= 30 and < 60
14	Not Wear	Male	>= 60
15	Not Wear	Female	< 30
16	Not Wear	Female	>= 30 and < 60
17	Not Wear	Female	>= 60

3) 개발 환경

- Aistage 서버
 - GPU : Tesla V100-PCIE-32GB
- IDE : VS code, Jupyter Lab
- Deep Learning Framework : PyTorch

4) 프로젝트 구조

- Domain understanding
- Data Mining : AI Stages 에서 제공
- Data Analysis
- Data Processing
- Modeling
- Training

5) 기대 효과

- 공공장소나 음식점 등 유동인구가 많은 장소의 입구에 갖춰진다면 적은 인적자원으로도 충분히 올바른 마스크 착용 검사가 가능

2. 프로젝트 팀 구성 및 역할

- 팀장 : 이나경 : 모델 생성 및 학습, 외부 데이터 수혈
- 팀원(7명)
- 김재현 : 모델 생성 및 학습, 앙상블
- 서동진 : 모델 생성 및 학습, 캐글 자료 수집 및 배포
- 안성민 : 모델 생성 및 학습
- 양재욱 : 모델 생성 및 학습
- 정민규 : 모델 생성 및 학습

3. 프로젝트 수행 절차 및 방법

- backbone 모델 선정하고 학습하여 competition에 적합한 pre-trained model 탐색
 - model list: efficientnet, vit, inception, vgg, swin, mobilenet etc
- 이후 앙상블로 더 좋은 성능 추구

4. 프로젝트 수행 결과

1) 탐색적 분석 및 전처리 (학습데이터 소개)

- 데이터
 - Train: 2,700명, Test: 1,800명
 - 1인당 7장
 - 올바른 마스크 착용 5장, 마스크 미착용 1장, 잘못된 마스크 착용 1장
 - Train: 18,900장, Test: 12,600장
- 데이터 노이즈 존재
 1. gender labeling error
 - ID : 006359, 006360, 006361, 006362, 006363, 006364
 2. mask labeling error
 - ID : 000020, 004418, 005227



- 여성 61.41%, 남성 38.59%
- 나이 그룹에선 고령층 데이터가 유난히 부족



- 마스크 중에서 다양한 마스크 존재

✓ 데이터 불균형을 고려해야하고 경계에 있는 나이대 구분이 어려울 것이라 예상!

2) 모델 개요

- vgg
 - N개 계층의 convolution layer로 구성된 neural network
 - efficientnet
 - compound scaling 방법으로 성능 향상
 - swin
 - transformer 구조를 object detection에 적용한 모델
- 세 종류의 모델을 앙상블 적용

3) 모델 선정 및 분석

1. 아키텍처 : swin_base_patch4_window12_384

- a. training time augmentation
HorizontalFlip, RandomFog, RGBShift,
RandomBrightnessContrast(brightness_limit=(-0.1, 0.1), contrast_limit=(-0.1, 0.1)),
GaussNoise, cutmix
- b. img_size : 384 x 384
- c. Loss function : CrossEntropy
- d. Optimizer : Adam, || lr_scheduler : CosineAnnealingWarmRestarts
- 2. 아키텍처 : tf_efficient_b4_ns
 - a. Augmentation : (CenterCrop, Resize, ColorJitter, ToTensor, Normalize, RandomHorizontalFlip)
 - b. Image Size : 380 x 380
 - c. Optimizer : Adam, || lr_scheduler : StepLR
- 3. 아키텍처 : vgg19
 - a. img_size : 384 x 384
 - b. Optimizer : Adam
 - c. Scheduler : CosineAnnealingWarmRestarts

4) 모델 평가 및 개선

- face net을 이용해 얼굴 위치 추적
- 정확한 평가를 위해 사람 id별로 k fold Cross validation
- 다양한 Learning Rate scheduler 적용
- Weight sampling을 이용해 데이터 불균형 문제 해결
- Auto cast를 이용한 효율적인 학습
- ensemble을 활용해 모델의 일반화 성능 향상

5) 시연 결과

- 모델 성능 Private LB score(accuracy/F1-score): 79.33%/ 0.737

5. 자체 평가 의견

- 잘한 점
 1. competition을 통해 배운 이론을 구현해보고, 각자 부족한 부분을 부단히보며 앞으로의 학습 방향을 세밀하게 설정할 수 있는 계기가 됨
- 시도했으나 잘 되지 않았던 점
 1. 한 모델에서 18개의 class로 바로 분류하지 않고 나이, 성별, 마스크 착용 상태를 분류하는 모델을 각각 만들어 학습시킨 후 최종모델로 병합하려 했지만 하위 모델의 성능이 좋지 않아 오히려 최종모델의 성능이 하락함
 2. 추가데이터가 normal밖에 없어서 incorrect와 mask데이터를 추가하기 위해 transforms의 randomerasing의 위치와 크기를 고정하는 방식으로 데이터를 생성하였습니다. 하지만 적절한 데이터의 역할을 하지 못했는지 성능 향상으로 연결되진 못했습니다.
 3. 배경을 지운 이미지를 학습해봤으나 성능향상에 큰 도움이 되지 못함
- 아쉬웠던 점
 1. 마지막에 validation없이 데이터를 모두 학습한 모델을 제출하려 했으나, 서버제출시간 변경 공지를 받지 못해 계획에 차질이 생겼습니다. 공지를 받지 못했다는 것을 증명해 기회는 생겼지만 시간이 촉박해 원하던 결과를 제출하지 못함
 2. 기존의 model architecture를 그대로 사용, 나만의 architecture를 만드는 시도를 못해봄

