

FACE IMAGE CLASSIFICATION

33 조 BOOSTERSHOT

SUBJECT

얼굴 사진 데이터를 가지고 마스크 착용 여부, 성별, 나이를 분류할 수 있는 Model 을 개발

CONCEPT

- ◆ Pytorch 전체 프로젝트를 개발함으로써 Deep Learning 개발 프로세스를 익히고 교육 과정에 배운 여러 기법들을 적용하여 효과가 있는지 직접 실험을 거쳐 확인하며 최대한 스스로 배울 수 있는 프로젝트
- ◆ 토론과 개별 실험을 거친 효과적인 협업 프로세스를 익히고 다양한 협업 툴을 활용하여 배울 수 있는 프로젝트

DEV ENVIRONMENT

- ◆ AI Stages V100 Instance - 팀원 당 각 한개
- ◆ VS Code
- ◆ WanDB - 결과 시각화

PROJECT STRUCTURE

- args.json - Train 과 Inference 실행 시 필요한 Argument 들을 파일로 저장
- dataset.py (dataset_final_edit.py)
- model.py
- train.py (train_kfold.py)
- inference.py (inference_age.py)
- ensemble.py

TEAM

- ◆ 남세현 : Baseline 구축, Train-Validation Set 구성, Model Test
- ◆ 강지우 : EDA, Data Augmentation
- ◆ 김성민 : Project Managing, Label smoothing, Pseudo labeling, DataLoader
- ◆ 유영재 : K-fold, Ensemble, DataLoader
- ◆ 박세진 : Model 테스트
- ◆ 정세종 : Face crop, DataLoader

PROJECT

EDA & DATA

모든 데이터셋은 아시아인 남녀로 구성되어 있고 나이는 20 대부터 70 대까지 다양하게 분포하고 있습니다.

데이터 구성

ORIGINAL DATASET

- ◆ 전체 사람 명 수 : 4,500
- ◆ 한 사람당 사진의 개수 : 7 [마스크 착용 5 장, 이상하게 착용(코스크, 텍스크) 1 장, 미착용 1 장]
- ◆ Train data 수 : 18,900
- ◆ Eval data 수 : 12,600(public : 6,300 / private : 6,300)
- ◆ 이미지 크기: (384, 512)

ALL-AGE-FACES-DATASET

아시아인. 2~80 세의 얼굴 13,322 장

AGE LABEL ENCODING

대회의 원래 목표는 30 살과 60 살을 기준으로 3 개의 범주로 분류하는 것이었습니다.

하지만 60 세 이상인 사람은 60 살만 존재하였기에 58 세를 기준으로 Labeling 하였습니다.

모델 개요

torchvision.models 와 Time 라이브러리에 있는 Pretrained 모델들을 backbone 으로 사용했습니다.

모델 선정 및 분석, 개선 방안

Validation 전략

한 사람 당 7 장의 데이터(wear 5 장, unwear 1 장, incorrect wear 1 장)로 구성되어 있었고, 아무 기준 없이 나눌 경우 같은 사람의 데이터가 Train_set 과 Validation_set 에 나뉠 수 있는 상황이었습니다. 그렇게 되면 Crown 에서 학습된 age, gender 를 바탕으로 Validation 에서 cheating 이 발생할 가능성이 높다고 판단되었고, 동일한 사람의 데이터는 Train 또는 Validation 에만 들어가도록 grouping 을 해주어야 했습니다. 뿐만 아니라, 데이터가 imbalance 하기 때문에 모든 Fold 가 Stratified 해야 했기 때문에 'pandas_streaming'이라는 library 를 통해 위와 같은 사안들을 고려하여 해당 기능을 구현하였습니다.

MODEL

- Backbone candidate(VGG, ViT, [Resnet152](#))
 1. VGG 는 baseline 의 모델이므로 시도
 2. ViT 는 특성상 데이터가 많이 필요한데, 데이터가 부족해서 성능이 좋지 않다.
 3. [ResNet](#) 이 VGG 의 overfitting, vanishing gradient 이 해결되어 최종 Backbone 으로 결정
- Loss : Cross Entropy(Label Smoothing)
- Optimizer : Adam
- lr : 0.0001
- lr_scheduler : LambdaLR(lr_lambda=lambda epoch: 0.95**epoch)

모델 평가 및 개선

AUGMENTATION

일괄적으로 CentorCrop 을 적용하면 얼굴이 잘못 잘려서 Data 가 훼손되는 경우가 꽤 있으므로 FaceNet 으로 Crop 한 뒤 처리되지 않은 데이터를 RetinaFace 를 활용해서 Crop

LOSS

Label smoothing 시 정답이 아닌 나머지 Class 에 모두 Smoothing 이 적용되는 것이 아니라 경계가 붙어있는 나이 Class 들만 Smoothing 을 적용

TEST








- Eval Data 의 이미지를 Lable Smoothing, Resnet152 을 이용한 모델로 Inference 하여 결과를 CSV 파일로 내보낸 후 다시 불러오는 방법으로 Pseudo Labeling 적용
- Ensemble 을 적용시 Hard voting 은 답을 맞춘 모델들의 결과를 무시할 수 있으므로 Soft voting 시도

최종 결과(모델 성능 평가)

<https://wandb.ai/ksm0517/age?workspace=user-ksm0517>

2nd prize / 38

Backbone - Resnet152, Label Smoothing(Customized), Psuedo Labeling, Face Crop, Normalization, Customizing Age Class, K-Fold(Ensemble X) 적용

순위	팀 이름	팀 멤버	F1 (Rank)	Accuracy
			0.777 → 0.771	81.952 → 81.524
2 (4 ▲)	이미지분류_33조	      	0.773 → 0.774	81.762 → 81.667
			0.770 → 0.776	81.492 → 81.841

OPINION

GOOD

1. 역할을 분배해 모듈들에 적용하는 시도를 함
2. 어떠한 것을 적용할 때 적용하는 이유가 분명했음
3. 다들 맡은 바 역할을 최선을 다해 구현 하려 했고, 대부분 구현에 성공함

BAD

1. 팀프로젝트 진행 시작이 다소 늦어지게 되어서 못 마무리한 일들이 많았음
2. K-Fold Ensemble 이 코드 구현은 되었으나 적용하지 못함
3. Age Class 를 세분화
4. Data Augmentation 기법이 효과적이지 못함

5. 다양한 Hyper Parameter 을 적용 및 시도(Optimizer, lr_scheduler 등)
6. 베이스라인 코드를 만들 때 협업을 했으면 좋았을 것 같음
(구조 협의 후 Dataset, Train, Inference, Model, Loss 와 같은 형태)
7. 좀 더 빠른 데이터셋 사용 여부 제출로 피드백을 받았으면 좋았을 듯함

REFERENCE

1. <https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html?highlight=crossentropy#torch.nn.CrossEntropyLoss> (Wiegth 에 관한 수식 참고)
2. [debugging-practice/5_Locate_face_bounding_box_facenet_pytorch.ipynb at main · snoop2head/debugging-practice \(github.com\)](#) (Face Crop 참고)
3. [JingchunCheng/All-Age-Faces-Dataset: All-Age-Faces \(AAF\) Database. \(github.com\)](#)