Toy project

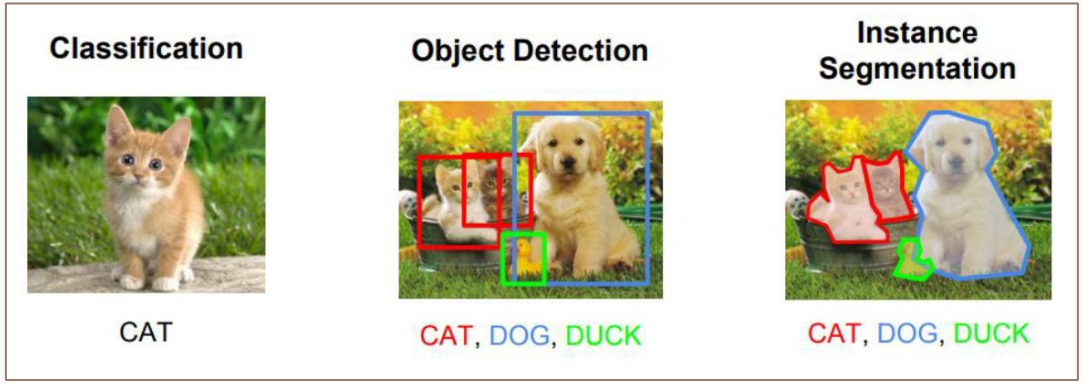
YAI

기초심화 CV팀

Table of contents

01. Before project – contents covered so far

- Task – image classification / object detection / segmentation



주차	일자	내용
1주차	24.07.22	ResNet
2주차	24.07.29	U-Net
3주차	24.08.05	R-CNN, ViT
4주차	24.08.12	Fast R-CNN, Faster R-CNN
5주차	24.08.19	GAN
6주차	24.08.26	Toy Project

01. Before project - criteria

Applicability

Practicality

Versatility

01. Before project – previous topics

제가 찾아본 데이터셋이랑 모델 보낼게요

-동영상 스트리밍 중 얼굴을 실시간으로 감지하여, 촬영자의 얼굴을 제외한 나머지 얼굴을 모자이크하는 시스템
YouTube Faces DB를 이용한 CNN (FaceNet 등)
https://www.cs.tau.ac.il/~wolf/ytfaces/index.html#download

-음식사진을 분석하여 칼로리 계산하기
Nutrition5k를 이용한 CNN (ResNet 등)
https://github.com/google-research-datasets/Nutrition5k

-저해상도 이미지를 고해상도로 변환하기
DIV2K 데이터셋을 이용한 GAN 모델 (SRGAN 등)
https://data.vision.ee.ethz.ch/cvl/DIV2K/

- 옷장을 보고 입을만한 코디 추천, 있으면 좋을만한 옷 종류 확인

무신사 코디맵 (데이터 크롤링, 다른 의류 사이트도 활용 가능)
https://www.musinsa.com/app/codimap/lists

Mask R-CNN (모델, image segmentation 용?)
https://github.com/matterport/Mask_RCNN?tab=readme-ov-file

- 얼굴 사진에서 object detection으로 눈 위치 감지, 눈 이미지로 졸음 판별 (R-CNN, ResNet 등)
- 감은 눈 사진에서 눈 위치를 감지, 뜬 눈 이미지로 대체 (R-CNN, GAN 등)

eye detector dataset
https://universe.roboflow.com/ai-project-t1xm8/eye-detector-01g2k/dataset/1

눈 close/open dataset
https://www.kaggle.com/datasets/tauilabdelilah/mrl-eye-dataset/data

네요ㅠㅠ 비디오에서의 특정 인물 추적과 관련된 내용을 보내드립니다.

Dataset
- LaPa dataset — 얼굴의 랜드마크와 마스크
https://github.com/jd-opensource/lapa-dataset
- COCO dataset — 다양한 객체에 대한 세분화 마스크와 주석
https://cocodataset.org/#home
- WIDER Face dataset — 얼굴의 바운딩 박스와 랜드마크
http://shuoyang1213.me/WIDERFACE/
- CelebAMask-HQ dataset — 얼굴 세분화 마스크
https://github.com/switchablenorms/CelebAMask-HQ
- YouTube Faces DB — 비디오에서의 얼굴 바운딩 박스와 랜드마크
https://www.cs.tau.ac.il/~wolf/ytfaces/

Model
- DeepSORT
https://github.com/nwojke/deep_sort
- SiamMask
https://github.com/foolwood/SiamMask
- Detectron2
https://github.com/facebookresearch/detectron2
- Bytetrack
https://github.com/ifzhang/ByteTrack

01. Before project – previous topics

Face Tracker

- ✓ Com[illegible]inputs…(video + targe[illegible]e image)
- ✓ Challenging for supervised learning…
- ✓ Difficulty in tracking the same individual…

02. Project Topic

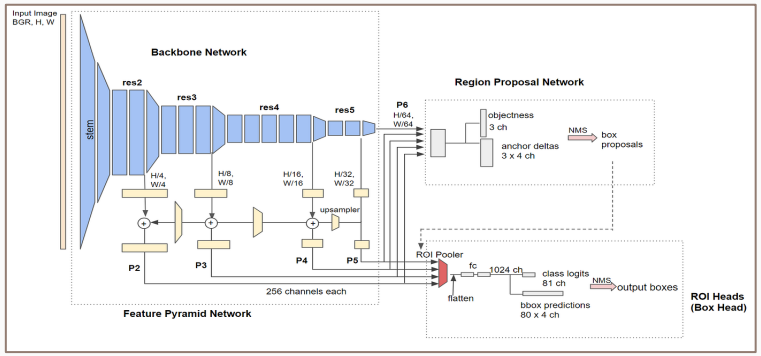
Face Mosaic

based on face detection

03. Method

Detectron2

- **Task – “Detect all faces in the image and then apply a mosaic effect”**
- Library – Detectron2 (Facebook AI Research)
- Backbone – R101-FPN Mask R-CNN model



Training & Inference

① Load a COCO-pretrained R101-FPN Mask R-CNN model
② Fine-tune with face dataset
③ Forward pass with the input image for inference
④ Utilize a mosaic filter on the faces in the output image (bounding boxes)

04. Implementation

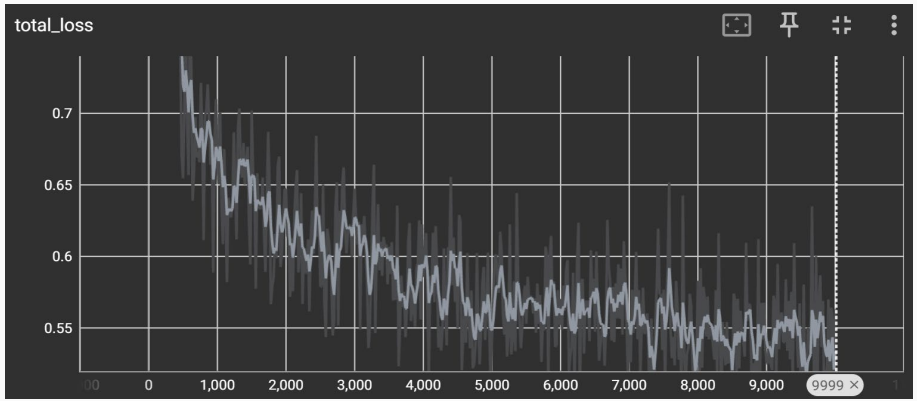
Dataset

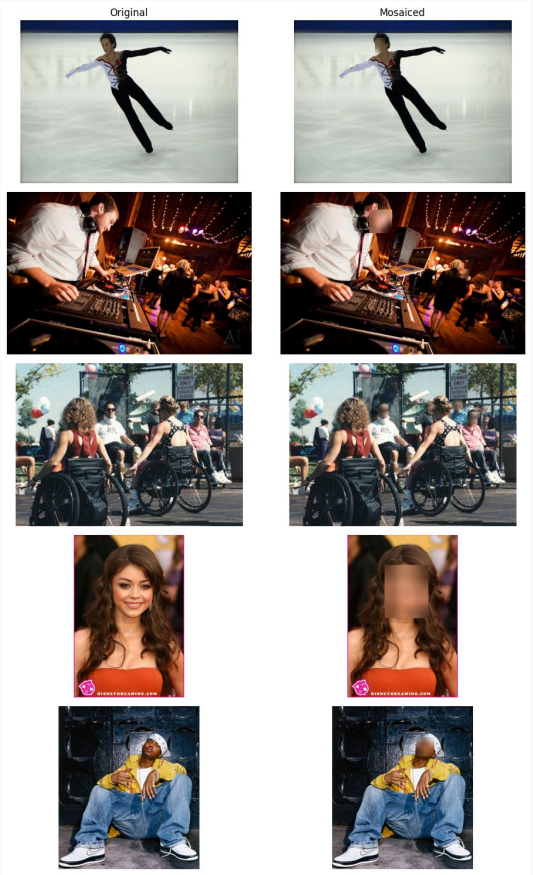
- WIDER FACE dataset
- 32,203 images and 393,703 face labels (40%/10%/50% for training/validation/testing)

Implementation details for fine -tuning

- Batch size = 16
- Iteration = 10,000
- Learning rate = 0.0025
- ROI head batch size = 512
- Used default settings in detectron2 for the other hyperparameters

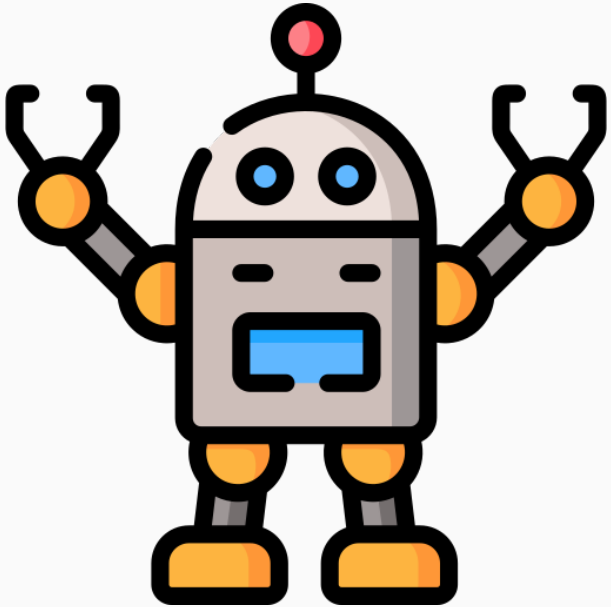
05. Results - image





05. Results - video

05. Results - demonstration



06. Conclusion

Limitations

- Inaccurate mosaic due to the use of bounding boxes (segmentation is required for better quality)
- Lack of consistency and information sharing between frames during video processing
- Unable to receive additional inputs such as text

Future works

- Selective face mosaic based on the input text or image
- Background blurring
- Real-time anonymization in public surveillance system

References & Roles

References

- WIDER FACE – http://shuoyang1213.me/WIDERFACE/
- Detectron2 – https://github.com/facebookresearch/detectron2
  https://detectron2.readthedocs.io/en/latest/index.html

Roles

- 김동윤 - 데이터 수집 및 데이터 로더 구현
- 김민규 – 데이터 수집 및 데이터 로더 구현
- 김수란 – 모델 학습
- 박민우 – 모자이크 필터 구현
- 신상우 (멘토) – 조언 및 방향성 제공

Thanks!

Any questions?