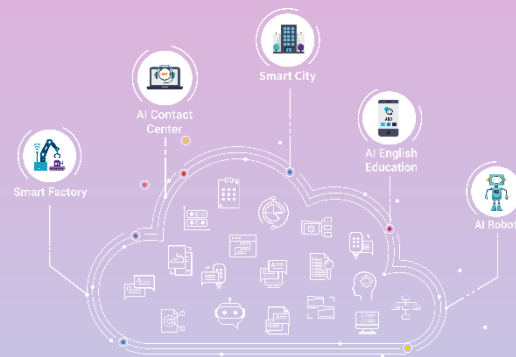


Level 2. How to customize AI engines



<http://mindslab.ai>

1. Text Removal 란?

2. Text Removal Model

3. Object Removal Dataset

4. Object Removal 실습



Text Removal (혹은 Object Removal) 란

- 주어진 그림 속 글자(혹은 임의의 물체)를 찾아내어 제거해주는 모듈
- Detection / Segmentation + Inpainting
 - 단 하나의 모델을 사용하는 것이 아니라, 두 개의 모델을 독립적으로 적용시켜 확장성을 증가시킴
 - Detect 혹은 Segment하고자 하는 대상이 달라진다면, 앞 단만 학습을 별도로 진행하면 됨



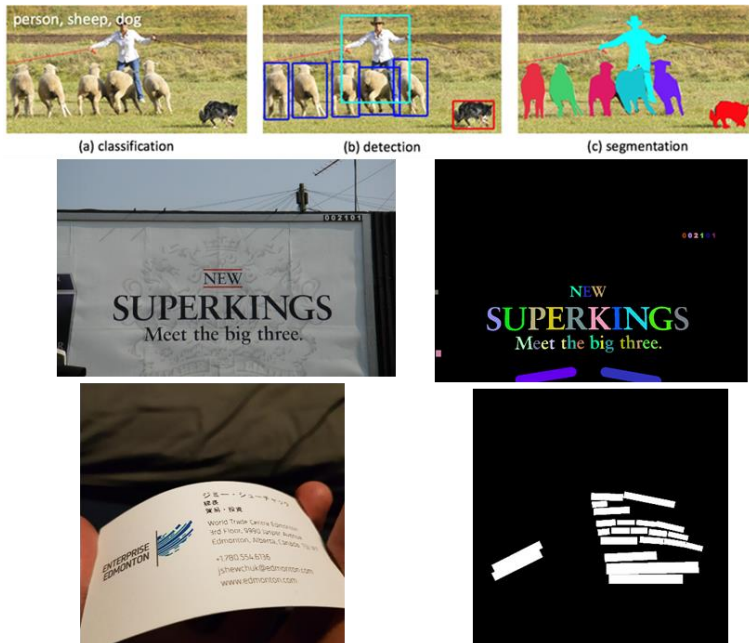
Text Removal

Detection/Segmentation & Inpainting

Text Removal (혹은 Object Removal) 란

- **Detection** : 사진 속 물체를 탐색하여, 네모 박스(Bounding Box)를 쳐주는 것
- **Segmentation** : 사진 속 물체를 탐색하여, pixel 단위로 그 물체의 위치를 예측하는 것
- **Inpainting** : 주어진 mask에 대하여 빈 공간을 자연스럽게 채워주는 로직

Detection / Segmentation



Inpainting



Text Removal

Detection/Segmentation & Inpainting

Text Removal (혹은 Object Removal) 란

- Detection/Segmentation + Inpainting

: 앞에서 찾은 물체의 위치를 기반으로 사진 내에서 물체를 지우는 방법

: 비단 글자 뿐만 아니라, 학습에 따라 임의의 물체를 Detection 하여 Inpainting 하는 것 역시 가능

: 현재는 Text Segmentation을 사용하나, Segmentation으로 고도화 가능

Detection / Segmentation

+

Inpainting



1. Text Removal 란?

2. Text Removal Model

3. Object Removal Dataset

4. Object Removal 실습



Text Detection

- 기본 Detection 과는 다르게 단어 구조를 인식해야함
 - e.g. 그림에 o가 있을 때, 디자인 o인지 글자 o, o, 0 인지 구분하기 위해선 주변 정보가 필요!
- Text Detection – PSENet (CVPR 2019)
 - IC17-MLT Dataset (Multi-language dataset)에 대해 학습되어, 영어/일본어/중국어/한국어 등에 대해 인식 가능

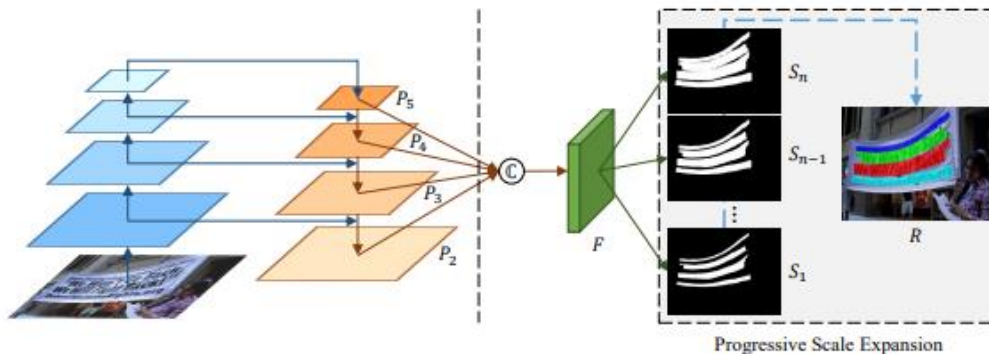


Figure 3. Illustration of our overall pipeline. The left part of pipeline is implemented from FPN [21]. The right part denotes the feature fusion and the progressive scale expansion algorithm.



(a) ICDAR 2015

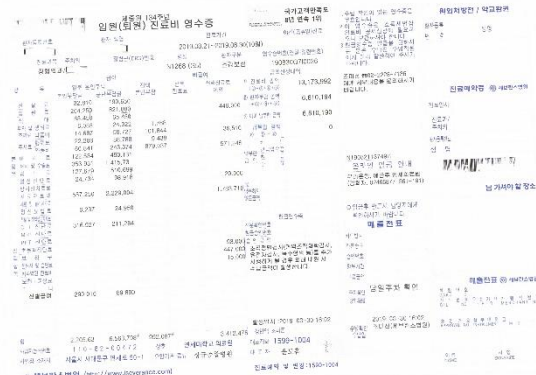
(b) ICDAR 2017 MLT

Text Removal Model

Text Segmentation

Text Segmentation (고도화 가능)

- Wild Image에 비해, **특정 Domain의 Dataset**의 경우 글자 단위로 인식해도 인식이 충분히 잘됨. 이 경우에는 segmentation network를 이용하여 detect 성능을 최대한 끌어올리는 것이 더 도움이 될 수 있음
 - 이 경우 inpainting에 넘겨주는 mask의 크기가 더 정확해 지기 때문에 removal의 성능을 올릴 수 있음
 - e.g. : 만화 속 글자 / 문서 속 글자
- 단 특정 domain을 벗어나면 성능 감소 할 수 있음
 - e.g. : 인쇄된 글자가 아닌 경우 / 휘어진 형태를 가진 경우)



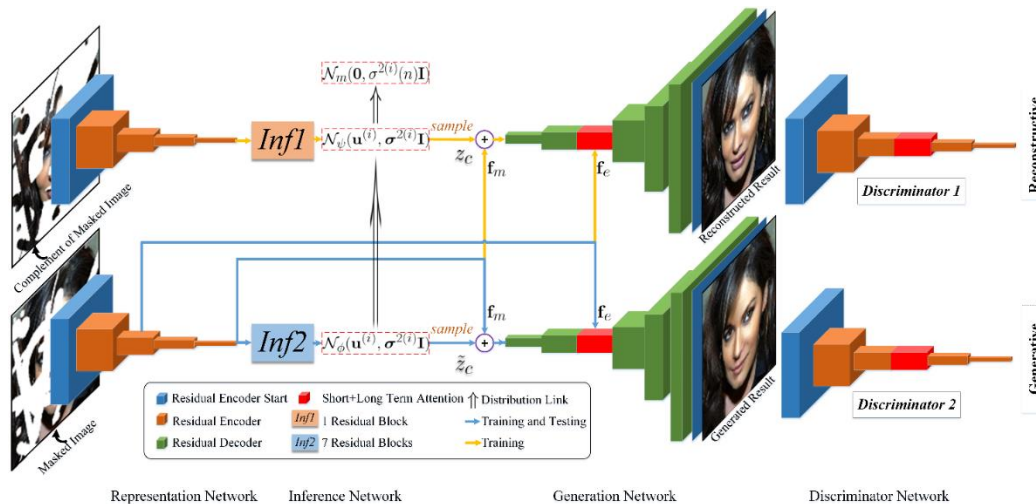
Inpainting

- 앞단에서 찾은 박스 혹은 mask에 대해서 지워주는 task를 수행
- Mask로 지워진 부분을 채웠을 때 나올 수 있는 결과는 유일하지 않으므로, 여러 번 수행할 수 있도록 구성
 - Pluralistic image Completion



Inpainting

- 현재 학습된 모듈은
 - 사람 얼굴 Dataset (CelebA)
 - 파리 풍경 (Paris)
 - 자연 풍경 (Place2)
 - 방대한 양의 일반적인 사물 데이터 (ImageNet)
- 특정 프로젝트에 최적화 시킬 시 inpainting 모듈 역시 변경 가능



1. Text Removal 란?

2. Text Removal Model

3. Object Removal Dataset

4. Object Removal 실습



Object Removal Dataset

Dataset 종류

- Segmentation 학습 Dataset에 따라 지우는 대상을 달리 할 수 있음
 - Step1 의 모듈을 쉽게 갈아끼울 수 있음
- Segmentation: 데이터 맞춤 제작이 상당히 비싸고 오래 걸림
- Detection: 데이터 제작이 저렴하고 간편하나, 섬세하지 못함

Text
Segmentation



Person
Segmentation



Brand Logo
Segmentation



Trash
Segmentation



1. Text Removal 란?
2. Text Removal Model
3. Object Removal Dataset
4. Object Removal 실습



Inpainting 테스트

- 1차: Nvidia-api
 - <https://www.nvidia.com/research/inpainting/>
 - 장점: 접근성이 좋음 / 간편한 ui
 - 단점: 내부 모델 / 데이터 비공개 / 최적화 불가
- 2차: 임의의 segmentation Mask로 Inpainting
 - <https://github.com/mindslab-ai/brain-ip>
 - python {project_home}/test_plu.py 실행
 - ./figure/tmp에 들어있는 sample img에 대해서 실행
 - python {project_home}/test.py 실행
 - Option:
 - img_file – input img 파일 경로
 - mask_file – mask img 파일 경로
 - 기타 : python {project_home}/test.py --help 로 조회 가능

Segmentation Customize

- 현재는 임의의 dataset 에 대한 finetuning pipeline은 존재하지 않음
 - Text Detection에 최적화 된 모델 사용
- 앞선 Inpainting 1, 2차 테스트 후 상용화 필요성이 검증된다면 해당 dataset 에 대해 detection/segmentation network 최적화 시키는 것이 최적의 성능을 뽑아 내는데 도움이 될 것으로 보임

예:



관련 자료



- http://image.ntua.gr/iva/datasets/flickr_logos/
- <https://github.com/lyndonzheng/Pluralistic-Inpainting>
- <https://github.com/whai362/PSENet>
- <https://github.com/garythung/trashnet>
- <http://cocodataset.org/>

감사합니다.

MINDs Lab. AI Platform Company