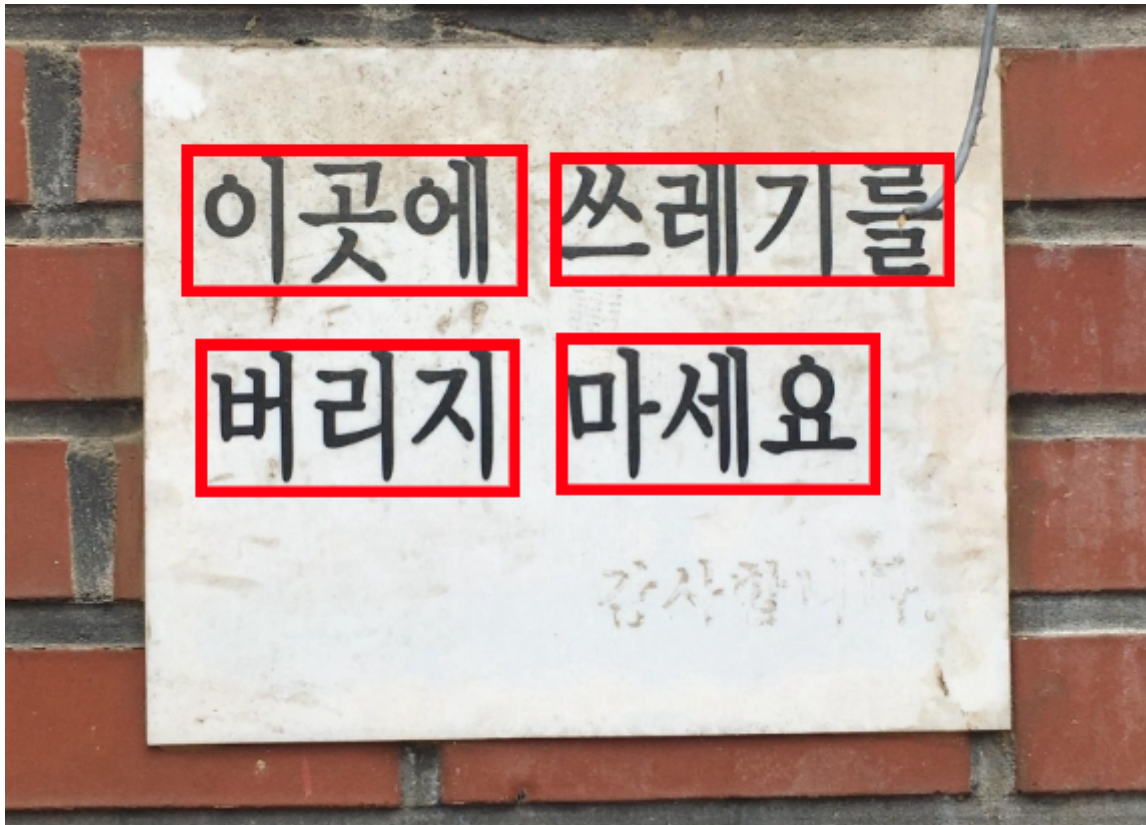


[부스트캠프] level-02 데이터제작 팀랩업리포트

프로젝트 개요



- 스마트폰으로 카드를 결제하거나, 카메라로 카드를 인식할 경우 자동으로 카드 번호가 입력되는 경우가 있다. 또 주차장에 들어가면 차량 번호가 자동으로 인식되는 경우도 흔히 있다.

이처럼 OCR (Optimal Character Recognition) 기술은 사람이 직접 쓰거나 이미지 속에 있는 문자를 얻은 다음 이를 컴퓨터가 인식할 수 있도록 하는 기술로, 컴퓨터 비전 분야에서 현재 널리 쓰이는 대표적인 기술 중 하나이다.

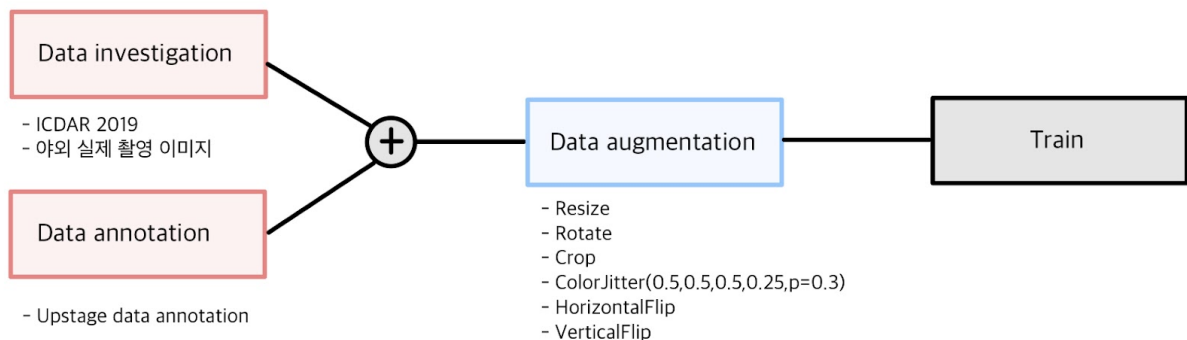
- OCR task는 글자 검출 (text detection), 글자 인식 (text recognition), 정렬기 (Serializer) 등의 모듈로 이루어져 있습니다.

이번 ‘글자검출대회’에서는 모델 학습 시 데이터 중요도 확인을 위해 **학습 모델은 고정된 상태로** 대회를 진행했다.

프로젝트 팀 구성 및 역할

- 김승현: data annotation
- 노창현: data annotation
- 최홍록: data annotation
- 최진아: data investigation
- 최용원: data augmentation

프로젝트 수행 절차 및 방법



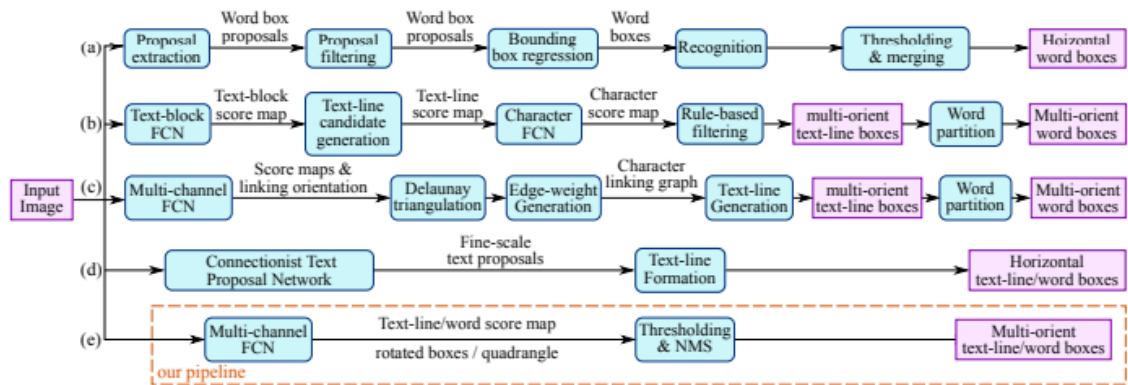
프로젝트 수행 결과

1. 학습 데이터

- ICDAR 2017 / 2019
- 야외 실제 촬영 한글 이미지([링크](#))

2. 모델(EAST: An Efficient and Accurate Scene Text Detector)

- <https://arxiv.org/abs/1704.03155>
- 데이터 제작 실습을 위해 모델은 고정하여 사용








3. Augmentation

- **horizon Flip:** f1 score, recall, precision 낮아짐
→ 데이터 셋이 정방향으로 많이 구성됨으로 효과가 없었던 것으로 추정
- **vertical Flip:** f1 score, recall, precision 낮아짐
→ 데이터 셋이 정방향으로 많이 구성됨으로 효과가 없었던 것으로 추정
- **Hue:** f1 score, recall, precision 낮아짐
→ 전체 밝기가 밝거나 어두워져 글자와 배경을 구분하지 못해 효과가 없었던 것으로 추정
- **saturation:** f1 score, recall, precision 낮아짐
→ 채도가 밝거나 어두워져 글자와 배경을 구분하지 못해 효과가 없었던 것으로 추정
- **Sharpen:** f1 score, recall, 낮아짐 precision은 높아짐
→ 이미지가 선명하게 됨으로써 precision이 증가한 것으로 추정 하지만 overfitting 이 발생해 recall이 낮은 것으로 추정
- **blur:** f1 score, recall, precision 낮아짐
→ OCR Data에서는 굵은 선들이 많기 때문에 blur가 오히려 글자 영역을 더 넓게 잡아 역효과가 나왔던 것으로 추정

실험 결과

Aa 데이터셋	≡ 언어	≡ 변경사항	≡ 결과
<u>ICDAR17</u>	1만장 En Ko others	baseline epoch 200 Adam crop size:512	f1 score: 0.5068 recall: 0.4121 precision: 0.6579
<u>제목 없음</u>		baseline + Sharpen epoch 200 Adam crop size: 512	f1 score: 0.4705 recall: 0.3572 precision: 0.6891
<u>제목 없음</u>		baseline + blur epoch 200 Adam crop size:512	f1 score: 0.4363 recall: 0.3272 precision: 0.6546
<u>제목 없음</u>		baseline + horizontal,vertical Filp epoch 200 Adam crop size:512	f1 score: 0.3907 recall: 0.2894 precision: 0.6008
<u>제목 없음</u>		baseline + Hue , saturation epoch 200 Adam crop size:512	f1 score: 0.2941 recall: 0.2164 precision: 0.4587
<u>ICDAR19</u>	1만장 En Ko others	AdamW epoch 80 crop size : 1024 batch 12 MultiStepAR [40, 70]	f1 score: 0.6509 recall: 0.5608 precision: 0.7754
<u>제목 없음</u>		baseline augmentation AdamW epoch 50	f1 score : 0.6612 recall: 0.5719 precision: 0.7837
<u>ICDAR17 + ICDAR19</u>	1500장 En Ko	baseline epoch 200 AdamW crop size: 412	f1 score: 0.4545 recall: 0.3755 precision: 0.5756
<u>ICDAR19 +야외 실제 촬영 한글 이미지</u>	22000장 En Ko others	baseline augmentaiton Epoch 30 AdamW	f1 score : 0.4121 recall 0.3385 precision 0.5266

Leader Board(Final)

순위	팀 이름	팀 멤버	f1 ↕	recall ↕	precision ↕	제출 횟수	최종 제출
11 (4 ▲)	CV_12조	    	0.6686	0.5760	0.7967	18	1d

자체 평가 의견

• 잘한 점

- 추가로 데이터를 수집하여 학습 형태에 맞게 annotation을 진행하였다.
- 다양한 augmentation을 추가하여 성능 향상을 할 수 있었다.
- 외부 데이터를 사용하여 모델의 성능을 높일 수 있다는 점을 알게 되었다.
- 데이터의 중요성을 깨달았다.

• 시도했으나 잘 되지 않은 점들

- 데이터를 추가하였으나 geometry정보 부재로 인해 성능이 좋게 나오지 않았다.