

2021 직업계고 AI 전문교육

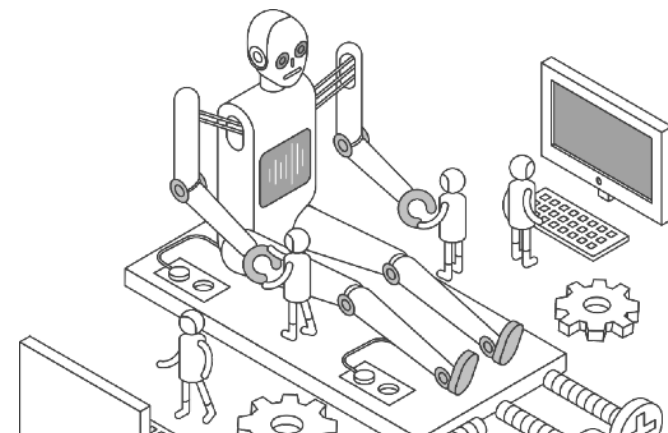
ARTIFICIAL INTELLIGENCE
BIG DATA
SMART FACTORY

AI·빅데이터 심화과정

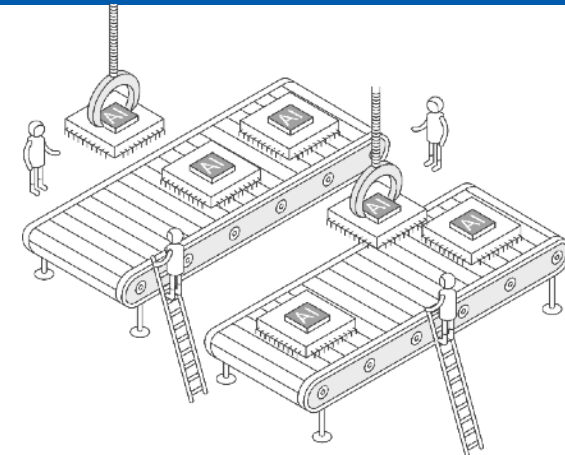
직접 만들어보는 OCR

박수철

github.com/scpark20
GaudioLab, 모두의연구소



RNN으로 음성 인식하기



Seq2Seq

9/24 - RNN으로 소설쓰기 (Aiffel 외)

9/29, 10/1 - 26. 뉴스 요약봇 만들기

10/6, 10/8, 10/15 - 27. 트랜스포머로 만드는 대화형 챗봇

CNN/GAN

10/20, 10/22 - 22. 난 스케치를 할테니 너는 채색을 하거라

10/27, 11/3 - 21. 흐린 사진을 선명하게

RNN+CNN

11/5 - RNN으로 음성인식하기 (Aiffel 외)

11.10 - 19. 직접 만들어보는 OCR

Text-to-Speech

11/12, 17 - Tacotron2: 딥러닝으로 만드는 TTS

OCR = Text Detection + Text recognition

- OCR은 Optical Character Recognition의 약자이며 text detection과 text recognition 두 모듈로 구성됩니다.
- Text detection은 주어진 사진에서 text를 가리키는 bounding box들의 좌표를 알아냅니다.
- Text recognition은 bounding box 내의 text를 인식하여 디지털화된 text로 출력합니다.

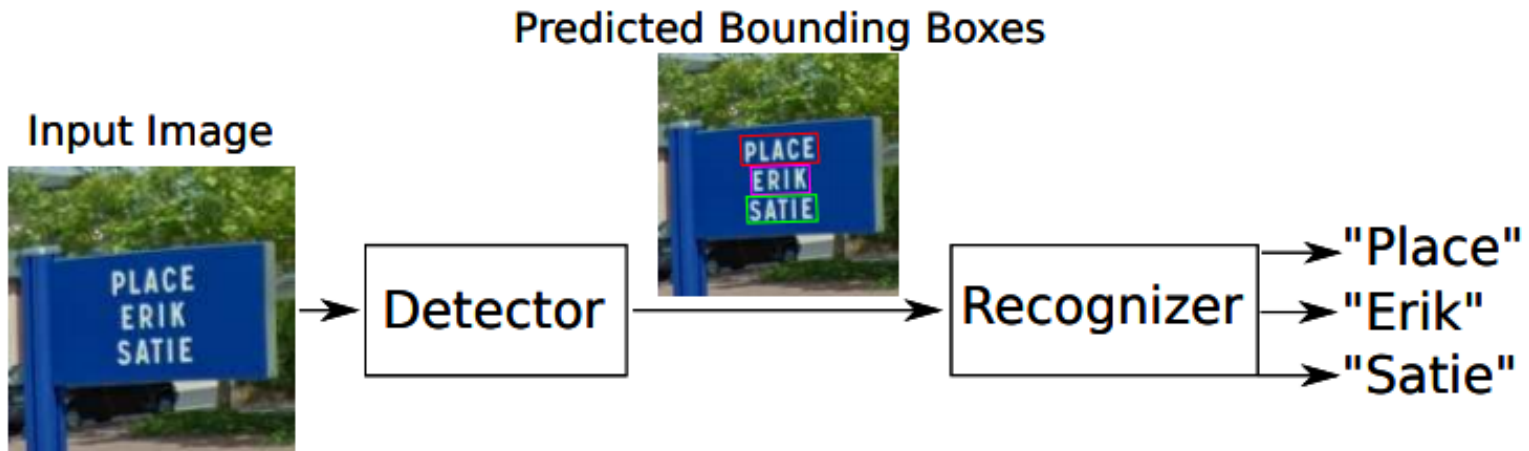


Figure credit : STN-OCR: A single Neural Network for Text Detection and Text Recognition

Text Detection

- Text detection은 일반적인 object detection과 비슷하게 convolution networks를 통해 bounding box의 좌표를 예측하는 방법으로 이루어집니다.
- 예를 들어 TextBoxes라는 모델은 VGG-16 network를 사용하며, 추가적으로 text-box layers를 통해 text의 유/무와 text가 있는 곳의 좌표를 출력합니다.

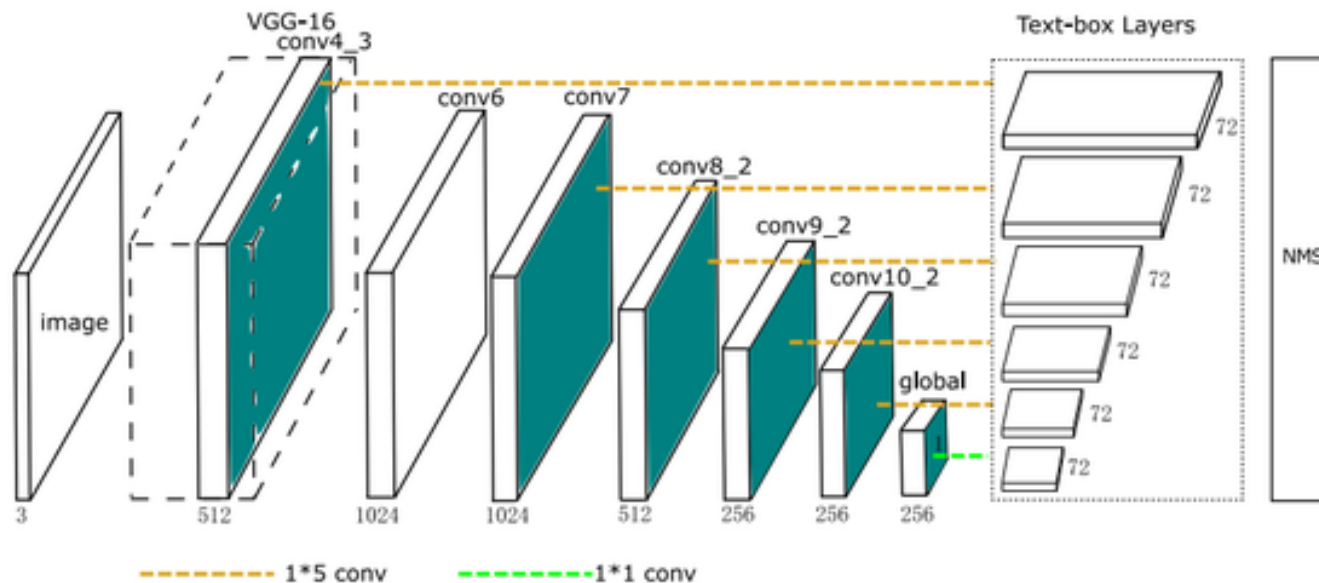


Figure credit : TextBoxes: A Fast Text Detector with a Single Deep Neural Network

Text Detection

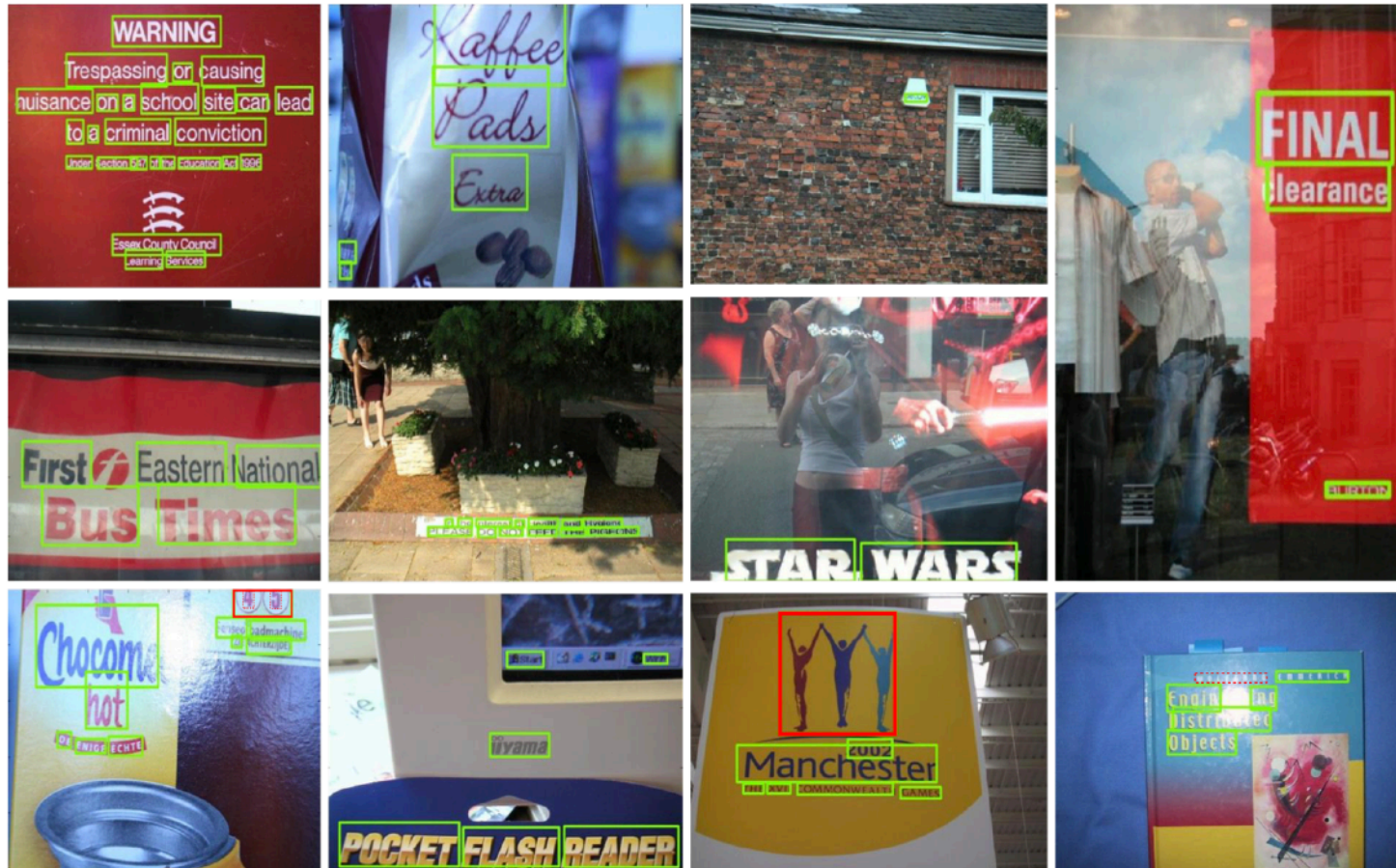


Figure 3: Examples of text localization results. The green bounding boxes are correct detections; Red boxes are false positives; Red dashed boxes are false negatives.

Figure credit : TextBoxes: A Fast Text Detector with a Single Deep Neural Network

Text Recognition

- Text recognition은 speech recognition과 비슷하게 CNN과 RNN을 통해 이미지를 분석하고 text의 각 token들에 대한 확률값을 출력하는 방법으로 이루어집니다.
- 그림에서 STATE가 적힌 그림이 주어지고 이를 2D convolution layers를 통해 spatial context를 분석합니다. 그 다음 LSTM layers를 통해 sequential context를 분석한 다음 각 frame당 어떤 token에 대응하는지 출력합니다.

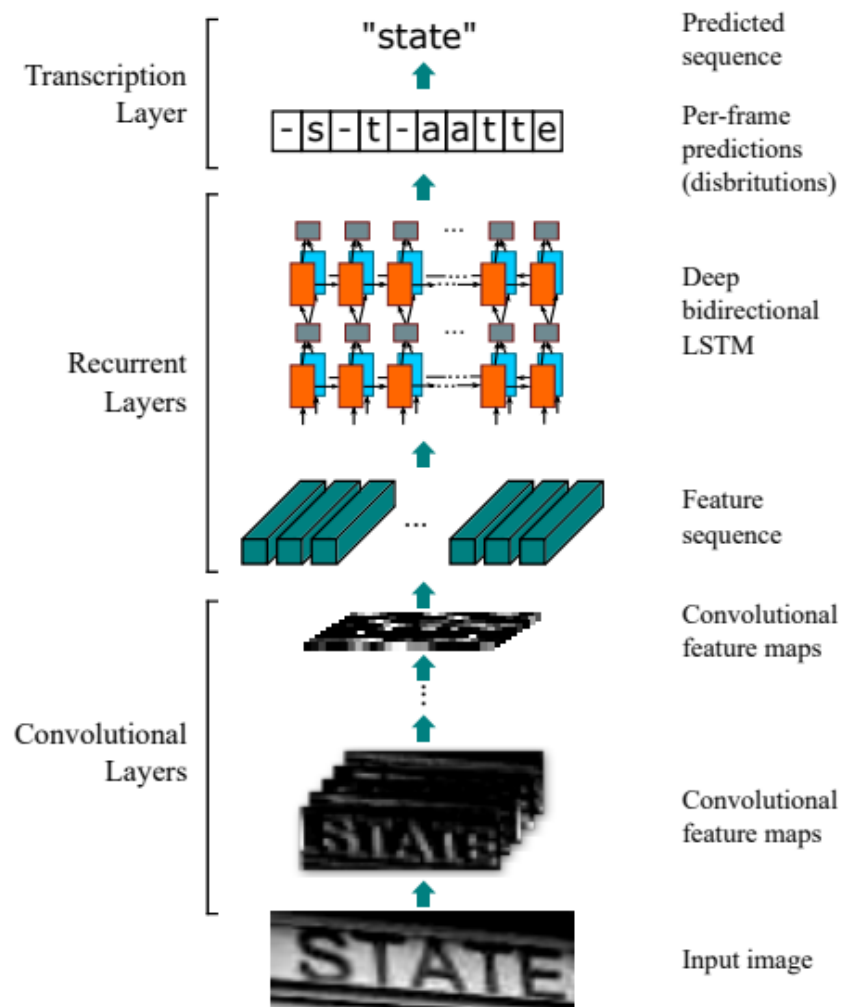


Figure credit : An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition

모두모두 파이팅!!