

# 물체 감지 모델을 활용한 화면 속 글자 감지 기법

곽철영, 박세웅 (서울대학교)

cykwak@netlab.snu.ac.kr, sbahk@snu.ac.kr

## I. 서론

최근 딥러닝 기술의 발달로, 화면 속 글자 감지 연구가 많은 각광을 받고 있다. 글자 감지 기술은 자율주행 차량 및 장애인을 위한 길안내 등 다양한 분야에서 활용될 수 있다. 본 논문에서는 물체 감지 모델 기반의 글자 감지 기법을 제안한다. 글자 감지에 필요한 제한조건을 맞추기 위해 목표 대상과 손실 함수를 변경하였고, 최종적으로 학습된 모델의 감지 결과를 제시한다.

## II. 본론

### 1. 물체 감지 모델 (Object detection)

최근 이미지 속 물체 감지를 위한 딥러닝 기반의 물체 감지 기법이 많이 제안되고 있다. 최근 제안된 Efficientdet [1]은 빠르면서도 높은 정확도를 보여주는 모델이다. 본 논문에서는 Efficientdet 을 화면 속 글자 검출을 위한 모델로 사용하였다.

### 2. 학습 데이터

글자 감지 모델의 경우 일반적인 물체 감지와 다르기 때문에, SynthText [2] 데이터셋을 이용하여 사전 학습을 한다. 본 논문에서도 SynthText 데이터셋을 이용하여 사전 학습 한 후 ICDAR 2015 [3] 데이터셋을 활용하여 학습을 진행하였다.

### 3. 목표 대상 및 손실 함수 변경

글자 감지의 경우, 감지된 글자를 바탕으로 글자 인식에 활용하는 경우가 많다. 이 때, 글자 형태에 맞춰 제대로 감지를 하지 못한다면, 글자 인식의 정확도에 영향을 미치게 된다. Efficientdet 의 검출 결과인 고정된 모양의 직사각형으로는 다양한 모양의 글자를 정확히 감지하는데 한계가 있다. 이를 해결하기 위해 목표 대상을 고정된 모양의 직사각형 대신 다양한 모양의 글자를 감지할 수 있는 일반적인 사각형으로 변경하였다.

Efficientdet 의 경우 anchor box 와 ground truth 사이의 차이를 학습한다. 기존의 모델의 경우 두 직사각형의 중심 좌표와 사각형의 높이 및 너비의 차이를 손실 함수로 사용한다. 그러나, 변경된 사각형은 직사각형이 아니기에 기존의 손실 함수를 그대로 사용할

수 없다. 따라서, 손실 함수를 anchor box 와 ground truth 네 꼭지점의 차이를 학습하도록 변경하였다.

### 4. 실험 결과

SynthText 데이터셋으로 사전 학습 시킨 후 ICDAR 2015 데이터셋을 학습시켰다. 최종적으로 학습된 모델을 이용한 성공 사례와 실패 예시는 아래 그림과 같다.



그림 1. 성공 및 실패 사례

위 그림과 같이, 큰 글씨가 화면 안에 산개한 경우는 잘 감지를 하는 반면, 작은 글씨로 여러 줄 적힌 경우에는 성능이 다소 떨어지는 것을 확인할 수 있었다. 이는, anchor box 의 크기 및 비율을 데이터에 맞춰 설정함으로써 해결할 수 있을 것으로 생각된다.

## III. 결론

본 논문에서는 물체 감지 모델 기반의 이미지 속 글자 감지 방법을 제안하였다. SynthText 데이터셋으로 사전 학습을 진행하였고, 최종적으로 ICDAR 데이터셋으로 학습 후, 학습된 모델의 검출 성능을 확인하였다.

## ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2020R1A2C2101815).

## 참 고 문 헌

- [1] T. Mingxing, R. Pang, and Q. V. Le. "Efficientdet: Scalable and efficient object detection." in *Proc. IEEE/CVF CVPR*, 2020.
- [2] D. Karatzas et. al., "ICDAR 2015 competition on robust reading," in *Proc. IEEE ICDAR*, 2015.
- [3] A. Gupta, A. Vedaldi, A. Zisserman, "Synthetic data for text localisation in natural images," in *Proc. IEEE CVPR*, 2016.