Preprocessing for Detecting Small Signs from Large Image

1. Introduction

Object detection은 현재 많은 분야에서 그 활용도가 높아지고 있다. 현재까지 이러한 object detection을 수행하는데 있어 CNN(Complex Neural Network)을 이용하여 왔다. CNN을 이용한 연구결과로 R-CNN, Fast-CNN, Faster-CNN, YOLO(You Look Only Once) 그리고 SSD(Single Shot Multibox Detector)와 같은 매우 효율적인 object detection system이 개발되었다.

그러나 큰 Input image에 대해 작은 object를 찾고자 할 경우 앞서 소개한 system들이 잘 찾아내지 못하였다. 더구나 input image의 사이즈가 크면 network에 넣고 수행함에 있어 memory부족이 발생하였다.

앞서 말한 memory문제를 해결하기 위해서는 network를 이전보다 간단하게 하여 성능에 저하가 발생할 수 있지만 뉴런 수를 줄이는 방법과 원래의 input image를 down sampling하여 사용하는 방법이 있다. 이 두 방법으로 필요한 memory를 줄일 수는 있지만 small object를 찾아내야 한다는 문제에는 답을 주지 못한다.

이를 해결하기 위해 아래와 같은 방법을 하려고 한다.

1) Input image로부터 학습을 통해 small object가 있을 확률이 높은 region을 먼저 찾는다.

2) 1)에서 구한 자료를 mask처럼 사용하여 Input image에 masking하여 small object가 있다고 생각하는 부분 외의 부분은 무시한다.

3) 2)에서 masking한 image를 maxpool을 통해 size를 줄이고 이를 YOLO와 같은 기존 network를 이용하여 찾고자 하는 object를 찾아낸다.

이 논문에서는 1)에 해당하는 작업을 수행하려고 한다.

2. Process

1) Data Set

앞으로 사용할 데이터 셋은 TT100K(Tsinghua-Tencent 100K)라는 교통표지판 detection & classification을 위해 만들어 놓은 데이터셋으로 아래의 <그림 1>에 있는 여러 종류의 교통표지판을 포함한 2048x2048 크기의 <그림 2>와 같은 야외 사진들로 이루어져 있다.

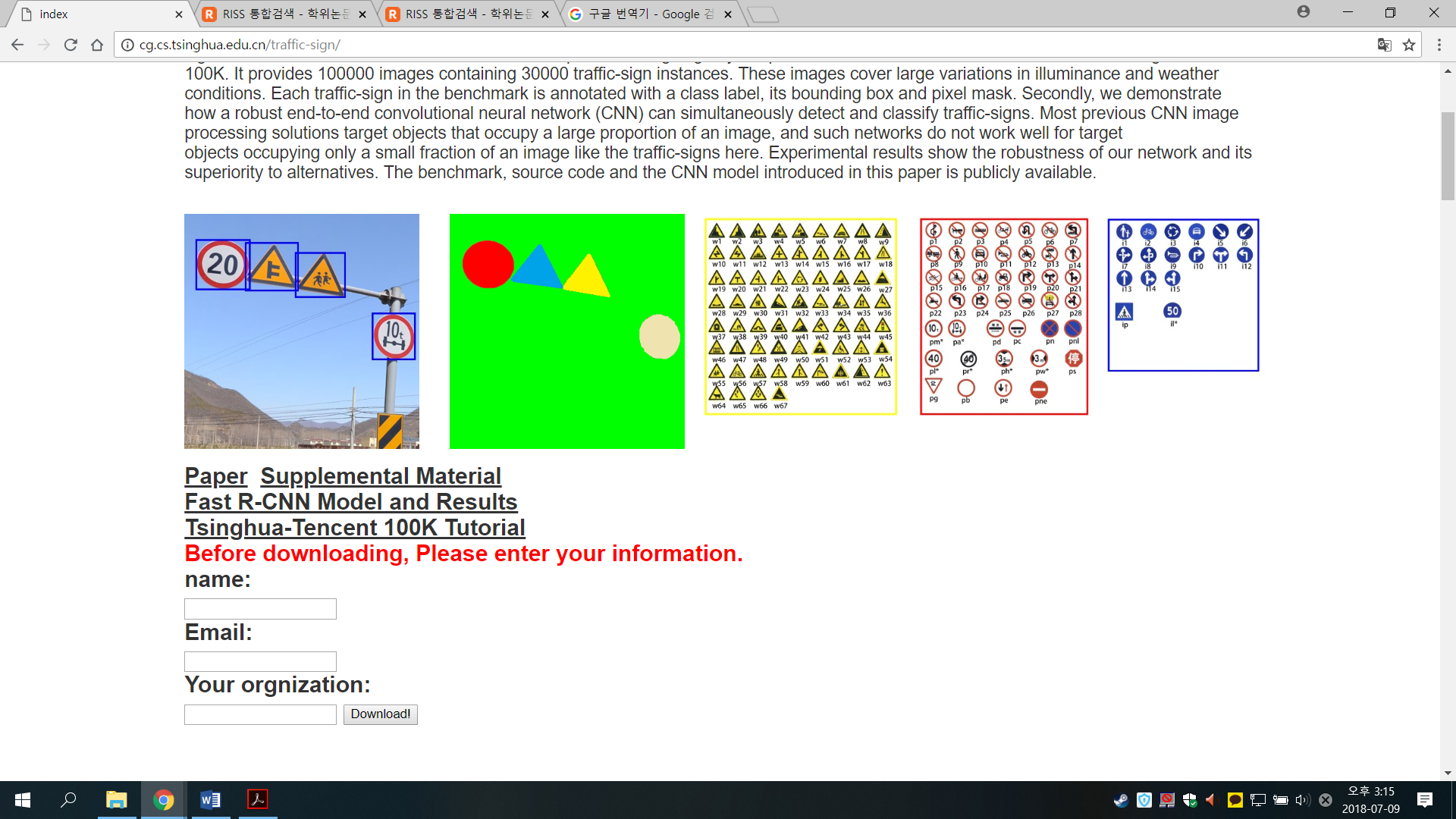


Figure . TT100K의 traffic sign 종류

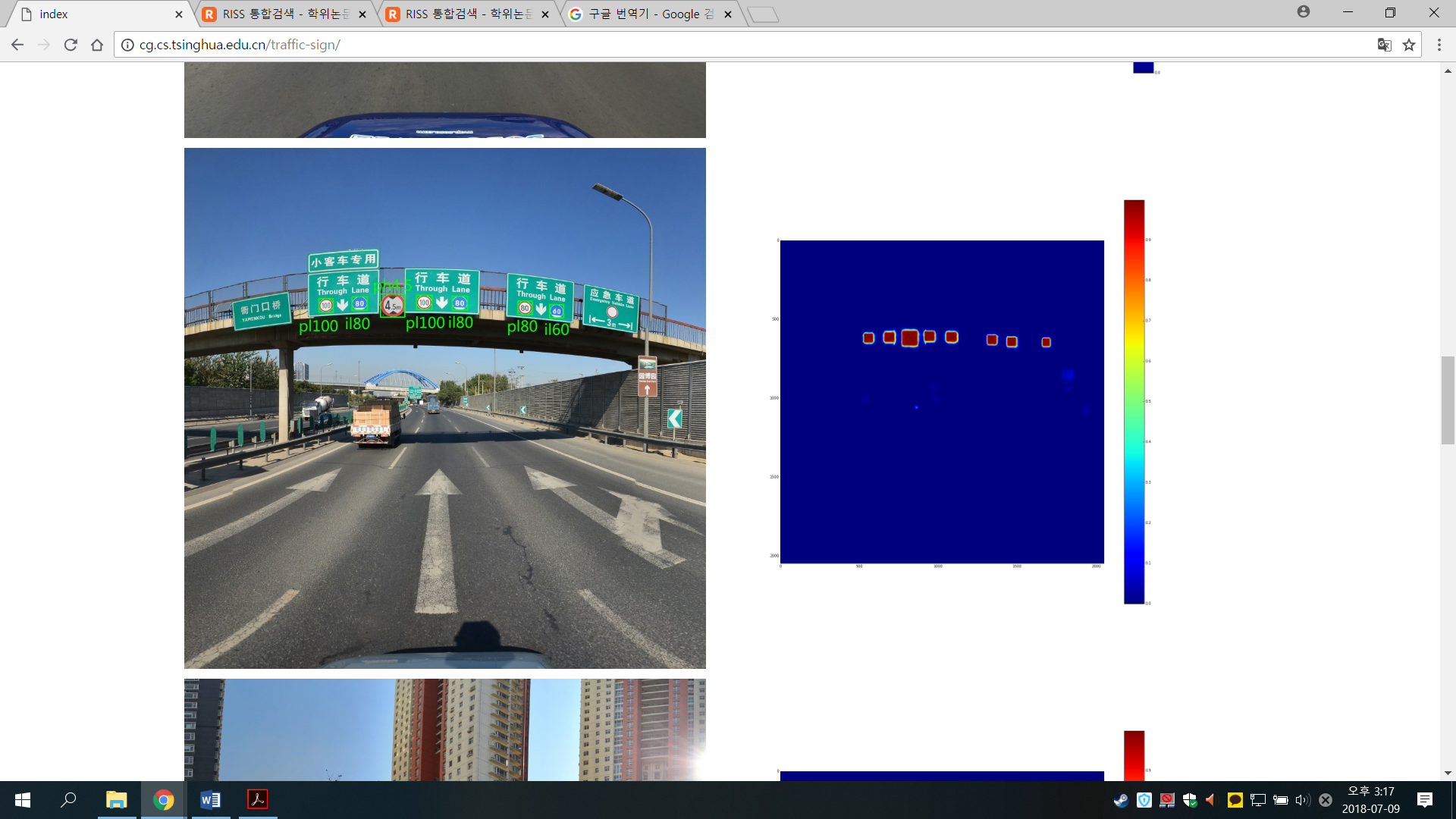


Figure . TT100K의 data 예시

2)