

# R-CNN을 이용한 음식 영역 추출하기

60150649 권 강 현

명지대학교

kgh9959@naver.com

*Abstract*— 지금까지 Object Detection 은 머신러닝에서 가장 뜨거운 주제 중 하나였다. 2013 년 R-CNN 논문발표를 시작으로 7 년 간 Fast R-CNN, Faster R-CNN, YOLO, Mask R-CNN, M2Det 등의 알고리즘이 발표되며 꾸준한 성능 향상이 이루어지고 있다. 이 프로젝트에서는 이 중 하나를 골라 어떻게든 구현해보고 가지고 있는 데이터셋에 적용해 보는 것을 목표로 한다.

*Keyword*— R-CNN

## I. INTRODUCTION

이 주제는 2019년 AI Starthon이라는 인공지능 경진 대회에서 가져온 것이다. 기존에 존재하는 Object Detection 알고리즘들을 통해 이미지 데이터에서 객체 탐지를 해보는 것이 목적이다.

## II. RELATED WORK

Object Detection에 대한 가장 전통적인 연구는 R-CNN(region based CNN)이다. 전통적인 R-CNN은 첫째 단계로 주어진 영상(이미지) 내에서 물체가 존재하는 위치를 Region of Interest(ROI)로 추출한 뒤, 둘째 단계로 그 ROI 각각에 존재하는 물체가 무엇인지 제각각 CNN으로 판별해내는 방식으로 작동했다.

하지만 2015년 Ross Girshick이 제안한 Fast R-CNN이라는 논문을 통해서 앞선 두 단계의 과정을 하나로 묶어내는 ROI pooling 방법을 제안하며 학습 속도의 큰 향상을 가져왔다. 그럼에도 불구하고 Fast R-CNN의 경우 GPU기반으로 작동하는 CNN 부분 연산은 빠르지만, 여전히 CPU기반으로 작동하게 되는 region proposal 추출 시간이 상대적으로 너무 오래 걸린다는 약점을 가지고 있었다.

그래서 2015년 같은 해에 그 문제를 개선한 Faster R-CNN이라는 논문이 발표되었다. 이 논문의 핵심은 region proposal 생성에 관련된

외부 알고리즘을 neural network로 만들어서 그걸 CNN 내부로 가져오는 것이었다. 이로써 Faster R-CNN은 R-CNN의 기본으로 자리잡으며 여타 object detection 관련 알고리즘에 비해 좋은 성능을 보여주게 되었다.

그리고 2017년 Kaiming He는 Mask R-CNN이라는 논문을 통해 주목할만한 R-CNN 계열의 성능 향상을 가져왔다. 이 논문의 핵심은 이미지에 존재하는 각각의 픽셀이 객체인지 아닌지를 masking하는 CNN을 추가하는 것이었다.

그리고 Object Detection에 관련하여 가장 최근 주목된 연구는 2019년 발표된 M2Det이라는 알고리즘이다. 이 알고리즘은 영상 속에서 실시간으로 크기가 변화하는 물체들을 탐지하는 작업에 Multi-level & Multi-scale 특징을 사용했다. 이 연구는 자신들의 방법이 기존의 유명한 YOLO 보다 좋은 성능을 냈다고 보고하고 있다.

## III. 실험방법

데이터는 AI hub에서 제공하는 ‘2019 AI Starthon X 네이버’의 데이터를 다운받아 사용할 것이다.

## IV. 예상 결과

R-CNN을 어떻게든 구현해서 데이터셋에 적용만 해내도 본인 기준에서 성공한 것으로 치겠다.