

YOLO를 활용한 여우의 분류

모두를 위한 인공지능의 활용 3조

21300816 한승수

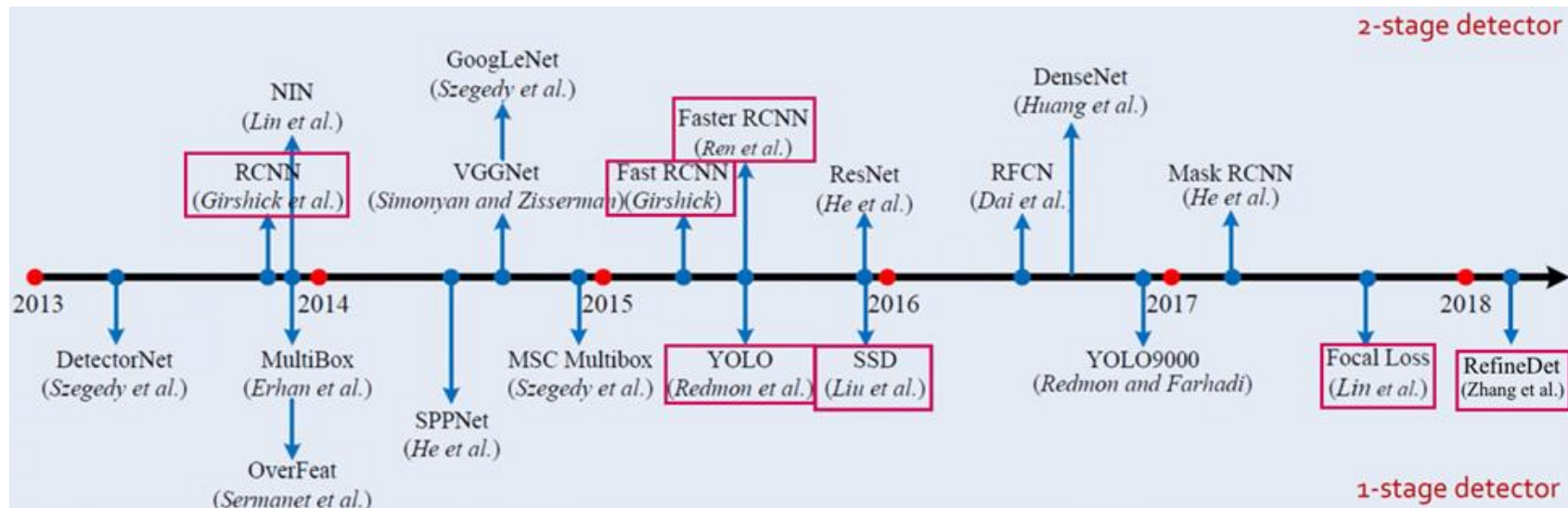
21500368 신동현

22000696 정주은

• 프로젝트 소개

- 꾸준히 증가양상을 보이는 여우에 대한 관심과 멸종위기 종에 대한 현실을 알리고자 하였습니다.
- 명강의에서 나온 YOLO를 활용하여 데이터를 학습시킨 후 도출해 낸 결과값을 통해 AI의 사물인식 기능에 대한 관심과 이해도를 향상시키고 싶었습니다.

객체 인식



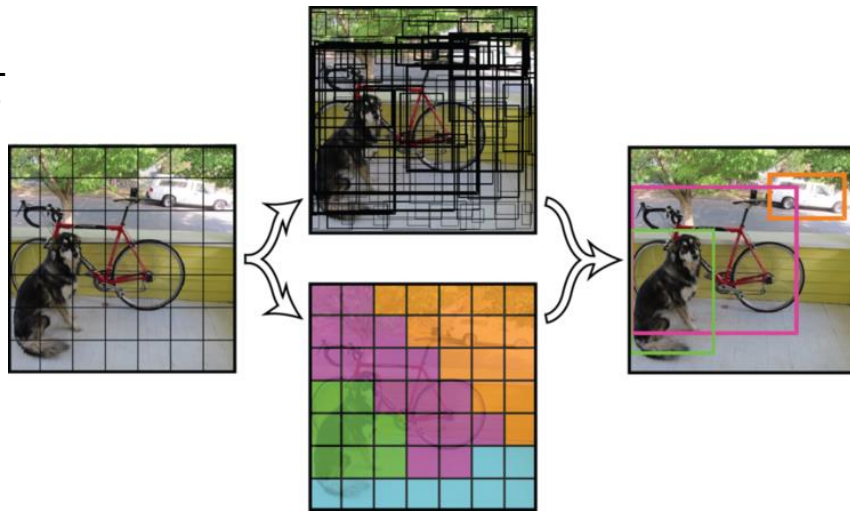
1-stage detector vs 2-stage detector

- 2-stage - localization과 classification 순차적
 - R-CNN 계열
 - 1000개 이상의 후보 제안
- 1-stage – localization과 classification 동시에
 - YOLO, SSD 계열
 - 98개의 물체 후보 제안
 - 작은 물체가 뭉쳐있을 경우 검출이 어렵다

- 객체 탐지 방법
 - CNN – sliding window 방식으로 다양한 크기의 사각형을 전체 이미지에서 쭉 이동하면서, 모든 영역을 CNN 모델로 인식 -> 속도 느림
 - R-CNN – 전체 이미지를 한 번만 CNN으로 돌려 특성 맵을 만들어 객체가 있을 만한 영역(ROI)들을 특성 맵에서 추출하여 각각을 별도의 분류기로 판단 -> 여전히 느림
 - YOLO – 실시간 객체 검출 가능

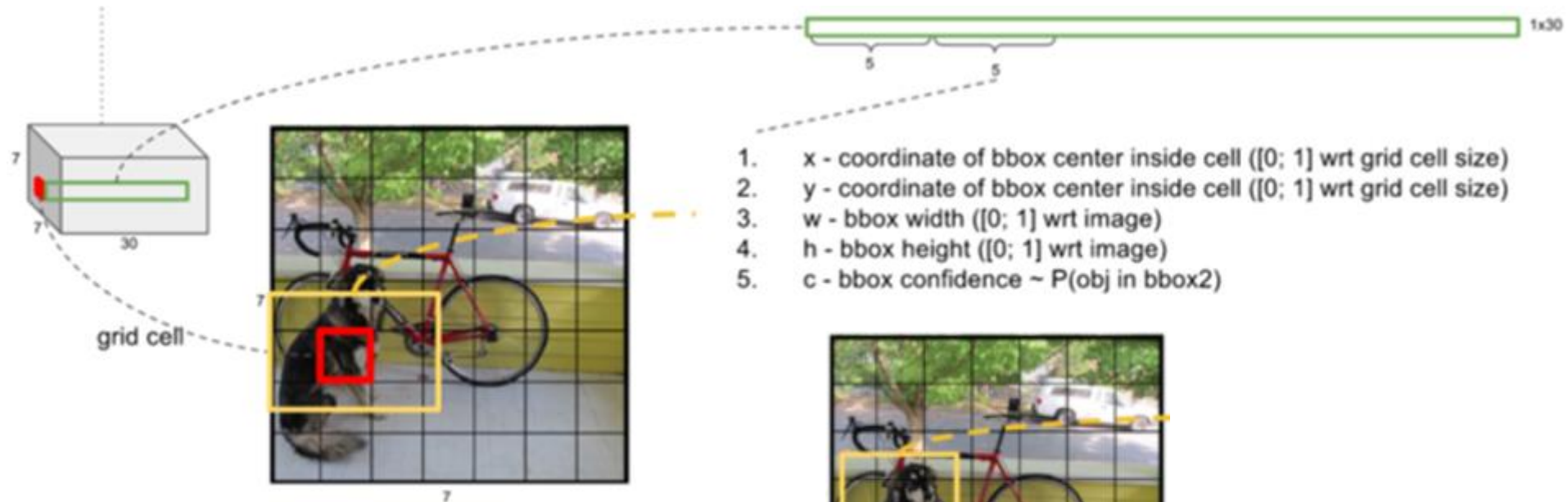
YOLO 탐지 방법

1. Input 이미지를 동일한 크기의 그리드(7*7)로 나눔
2. 각 그리드에 대해 그리드 중앙을 중심으로 predefined shape으로 지정된 경계 박스의 개수를 예측
3. Anchor boxes 이를 기반으로 신뢰도를 계산
4. 높은 객체 신뢰도를 가진 위치를 선택하여 class를 파악

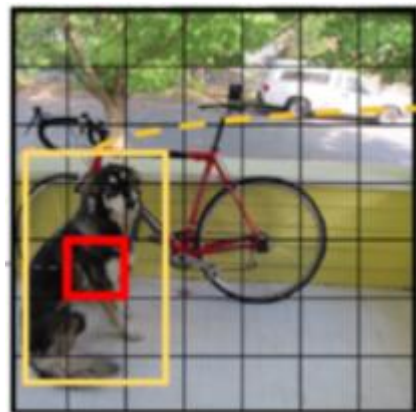


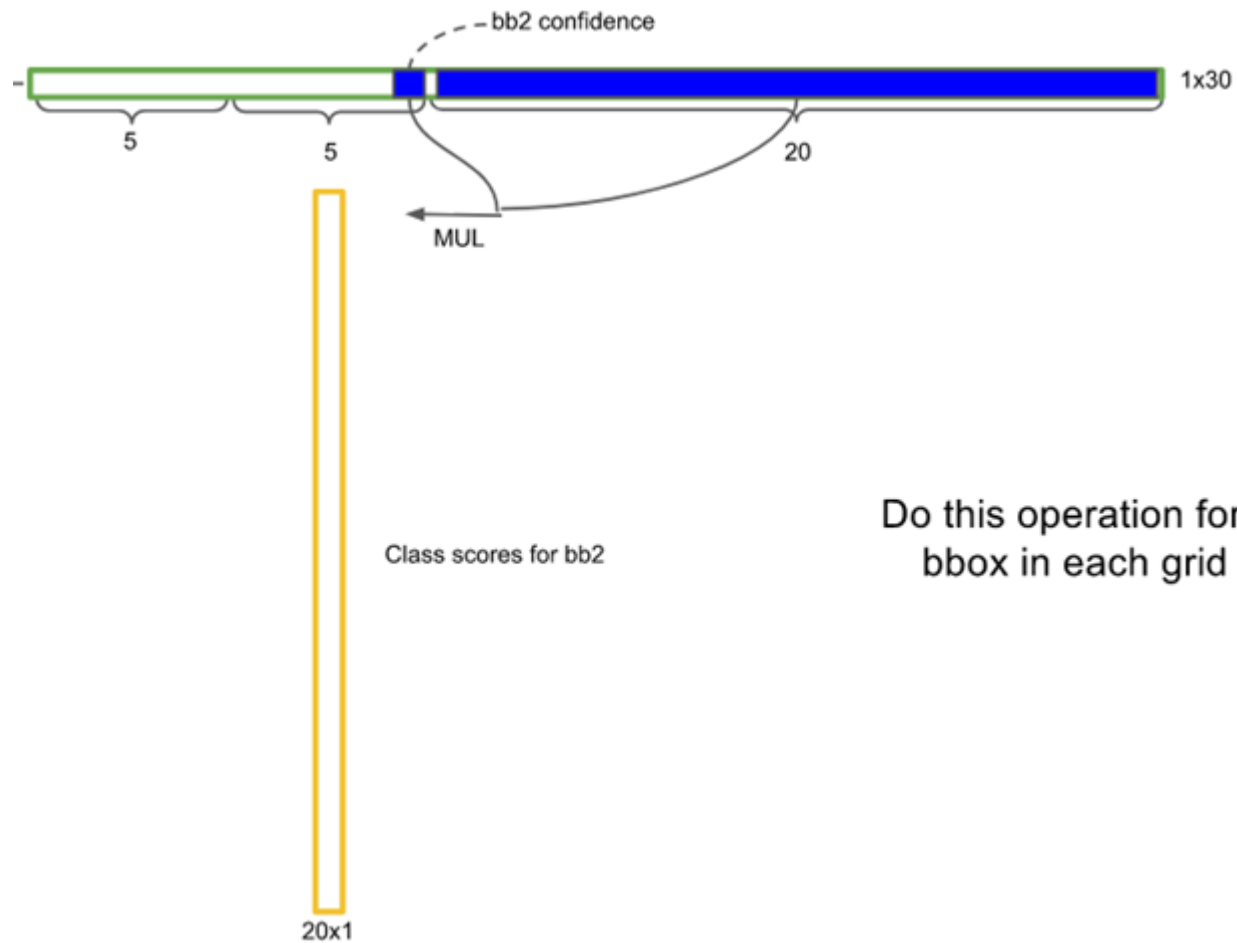
탐지 결과 - bounding box

- X, y, w, h , confidence score 로 구성됨
- X, y 는 grid cell의 경계를 기준으로 한 box의 중심좌표
- W, h 는 예측된 object와 input이미지의 width와 height의 비율
- Confidence는 해당 box에서 각각의 클래스들이 존재할 확률



1개의 Grid box에서 2개의 후보를 선택
2개의 bounding box에 대한 정보 저장

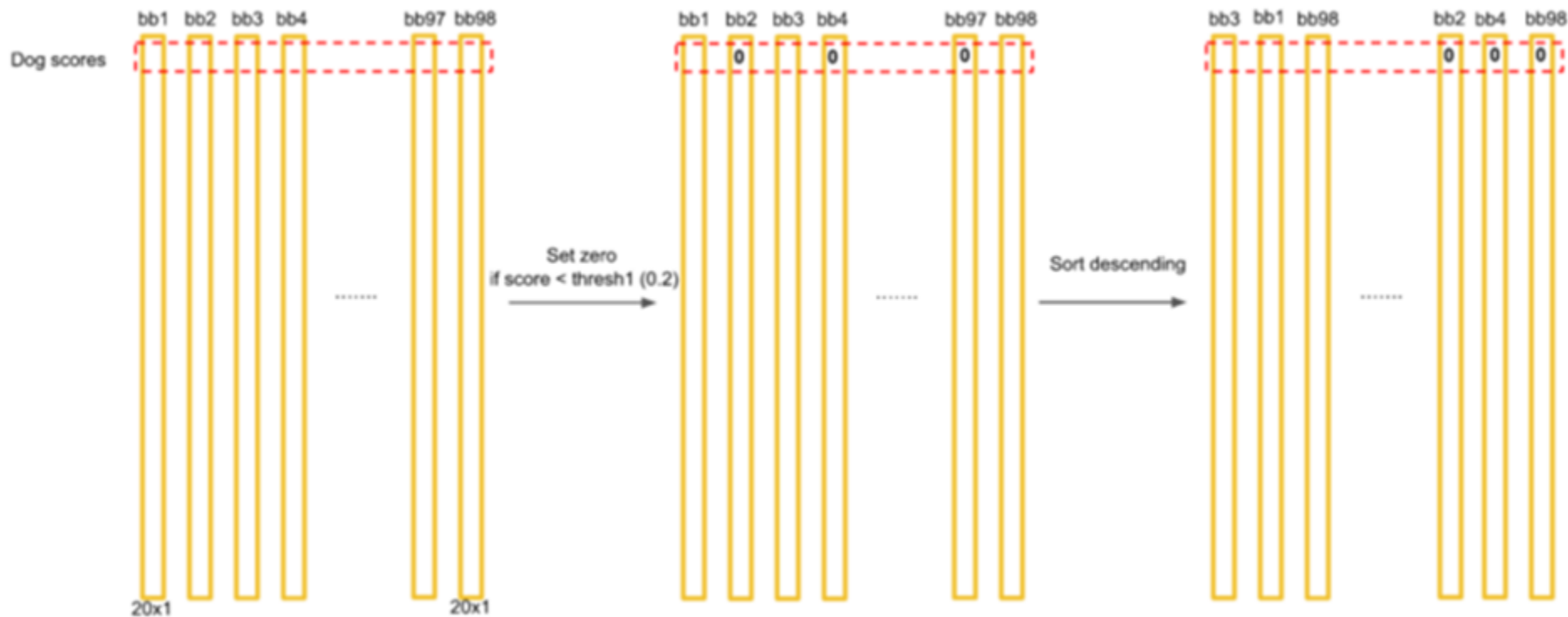


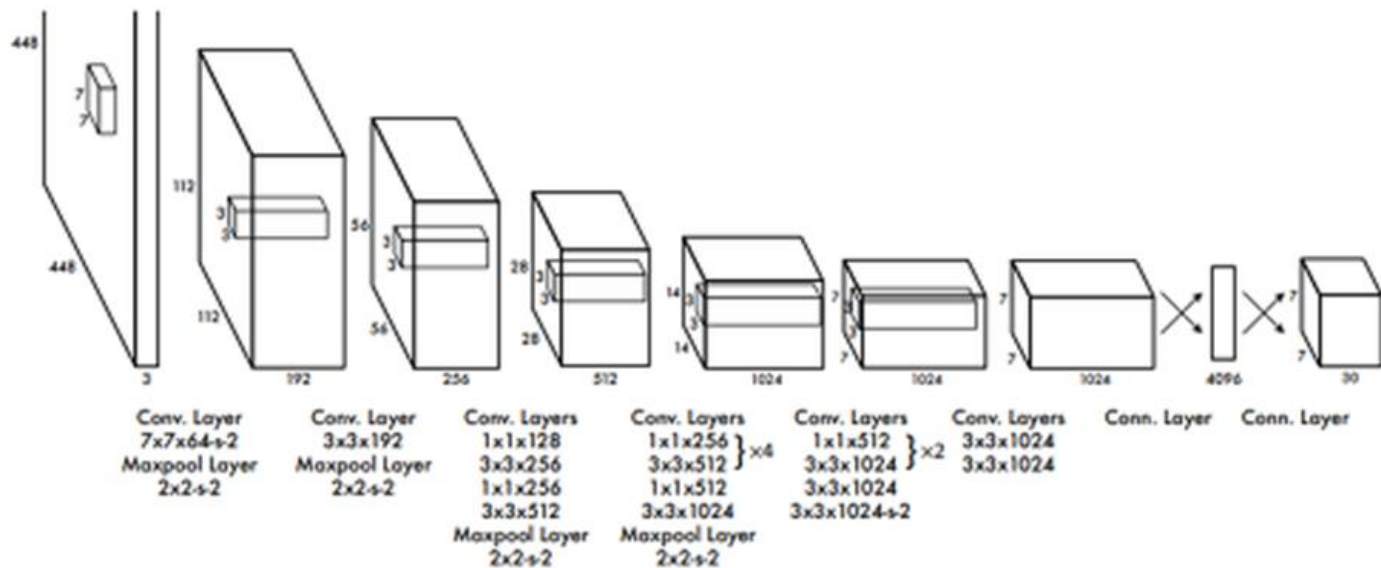


각각의 클래스들이
존재할 확률정보가
담겨있음

Do this operation for each
bbox in each grid cell

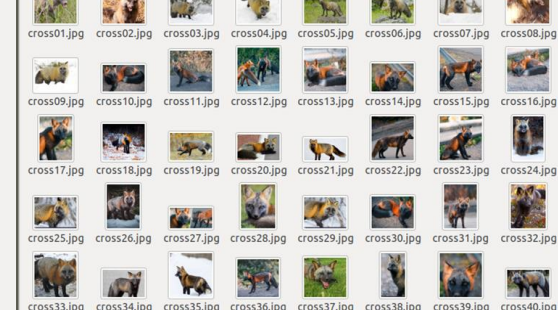
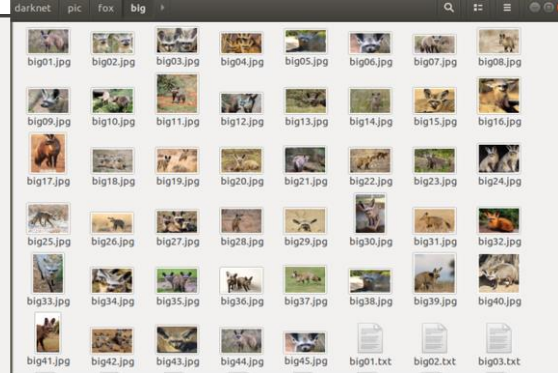
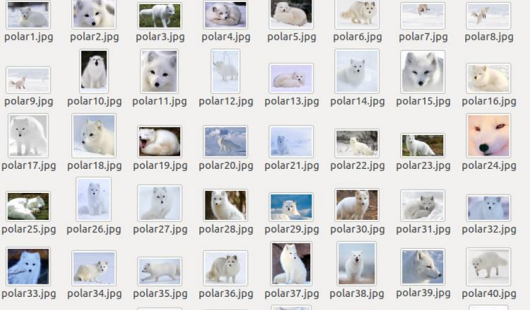
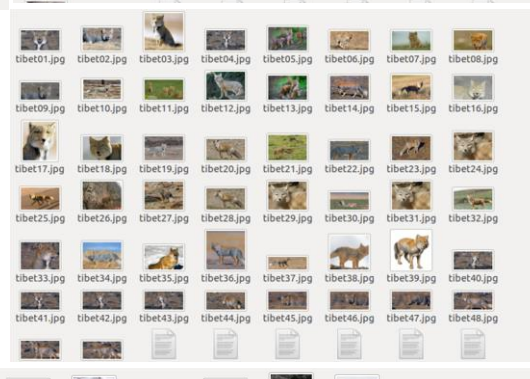
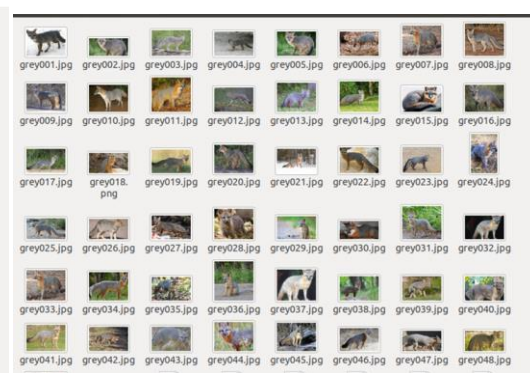
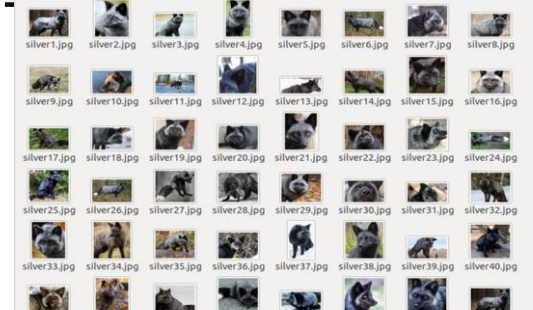
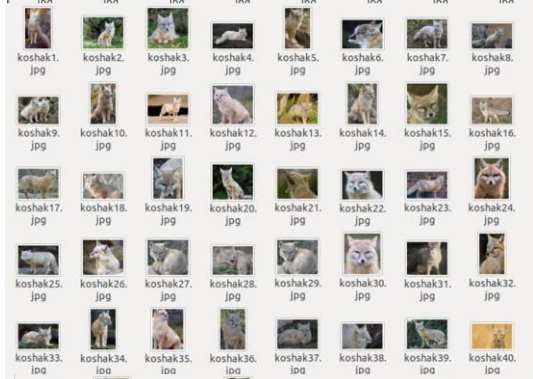
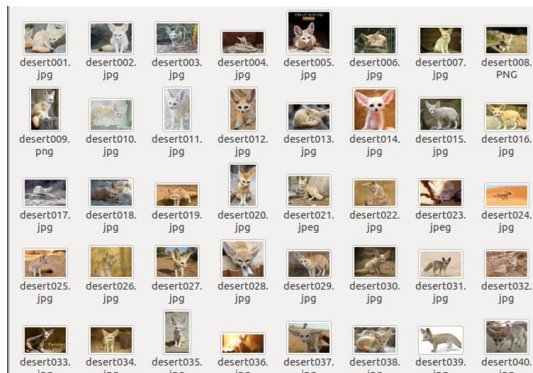
Total class classification confidence





Input 이미지 -> 24 convolution layer -> 2 full connection layer -> resize -> 특징데이터 추출

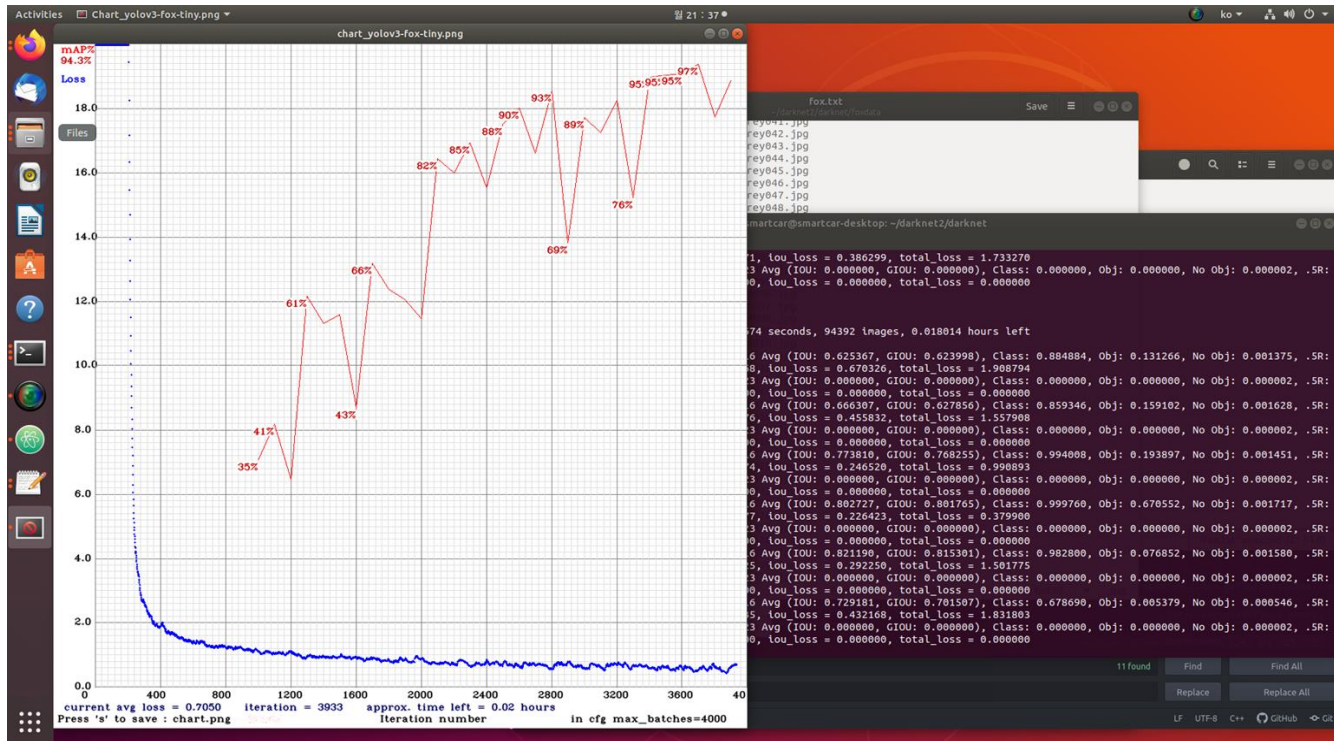
Input image를 Convolution 연산을 통해 이미지의 특징추출
Max polling 방법으로 이미지의 크기를 줄이고 feature를 강조



학습 진행



학습



학습 결과

| index | name | average precision | TP | FP |
|-------|------------|-------------------|----|----|
| 1 | red fox | 93.23% | 38 | 3 |
| 2 | desert fox | 90.60% | 29 | 2 |
| 3 | grey fox | 91.77% | 38 | 0 |
| 4 | polar | 97.56% | 45 | 3 |
| 5 | silver | 91.77% | 50 | 7 |
| 6 | koshak | 89.56% | 50 | 8 |
| 7 | tibet | 88.31% | 45 | 7 |
| 8 | cross | 99.23% | 35 | 1 |
| 9 | big | 97.54% | 55 | 7 |

mAP : 93.72 %

Thresh hold 25% 일때

TP : 385, FP : 38

FN : 93, TN : 93

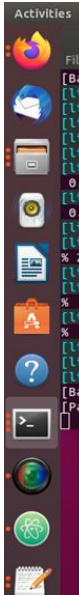
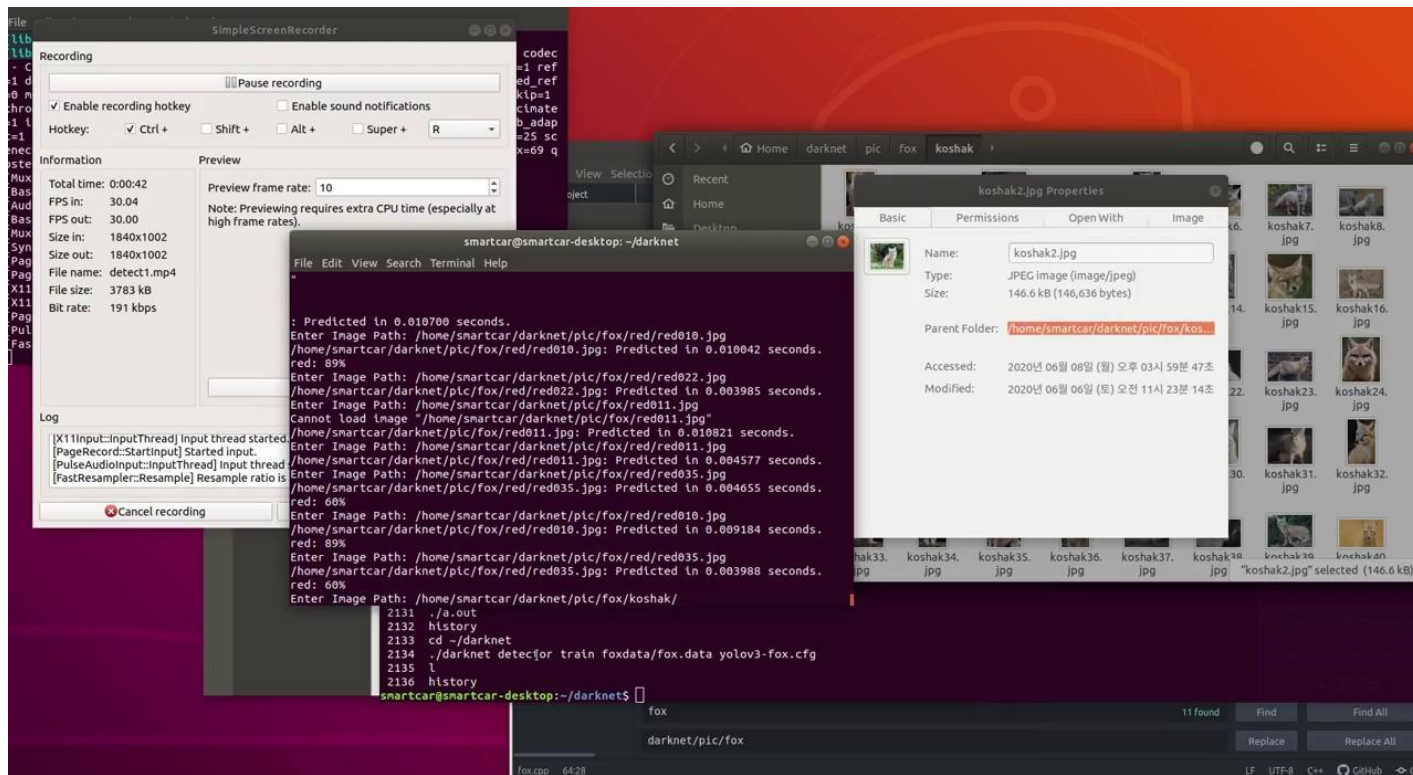
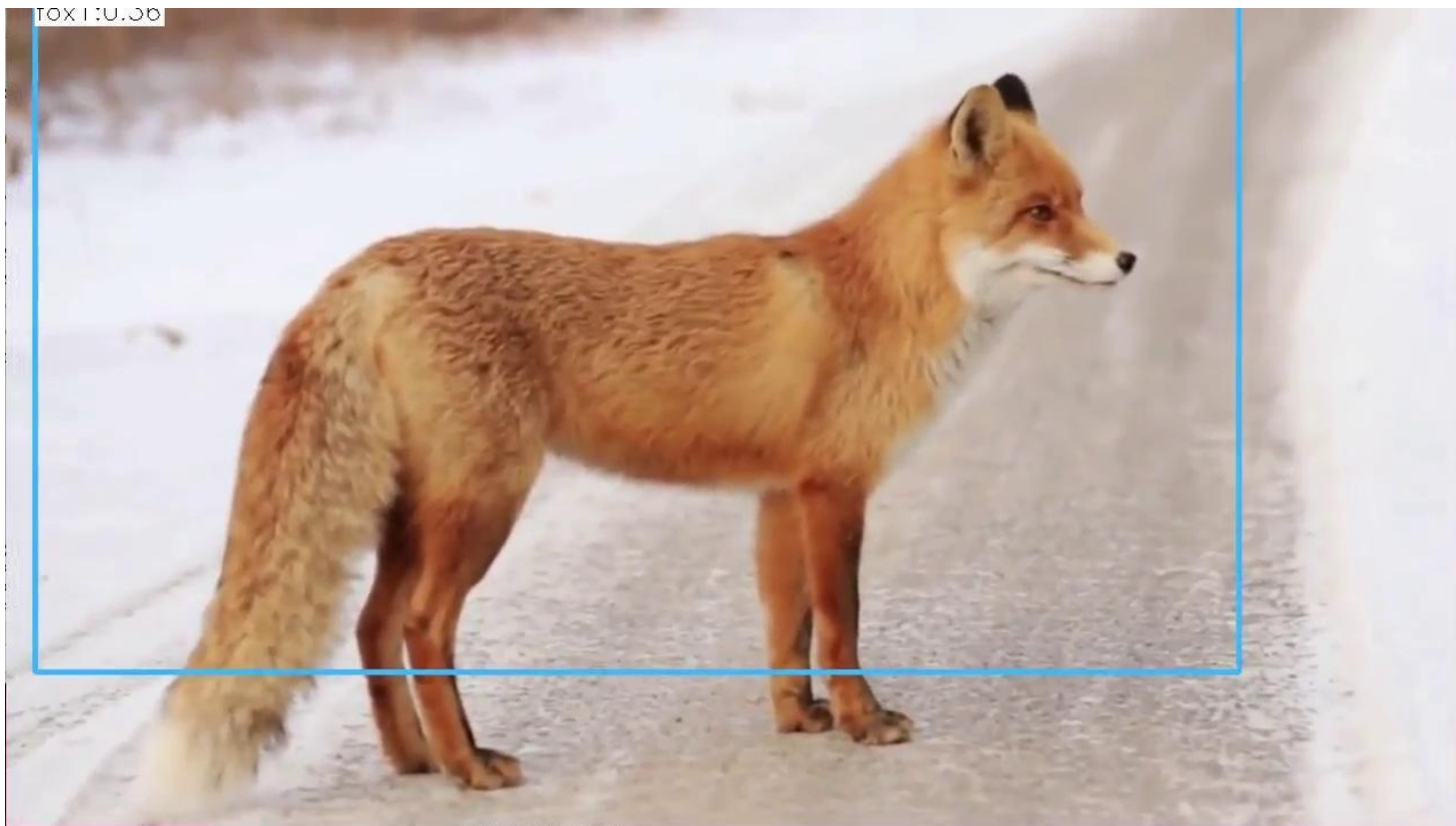


Image Detection



Object Detection



한계점

구글 이미지에서 찾은 여우 사진만으로는 데이터가 충분하지 않음

다양한 각도, 밝기, 주변환경에서 찍은 이미지가 필요한데 각 클래스당 대략 50장 정도밖에 확보하지 못해 낮은 인식률을 보였다.