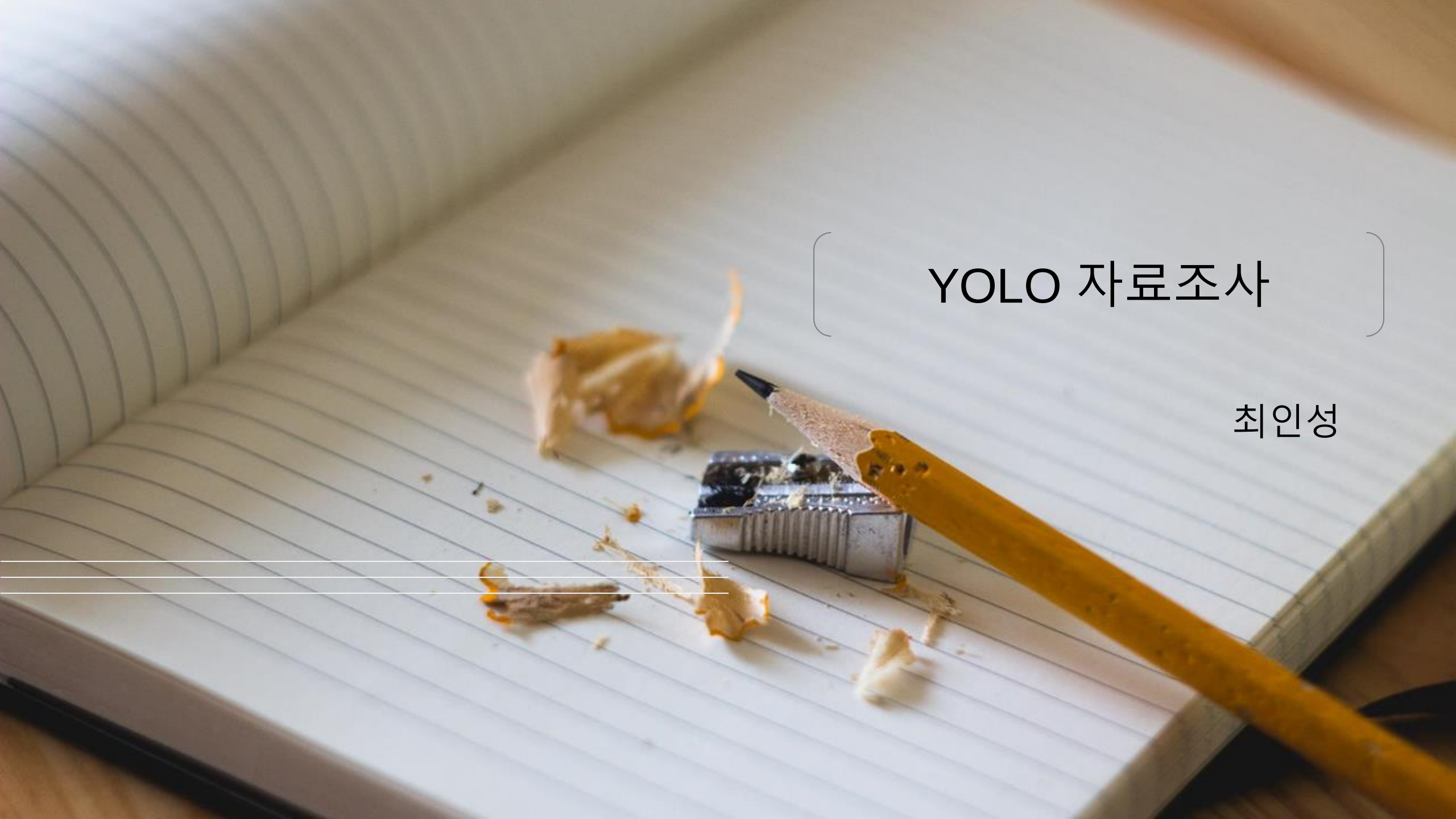


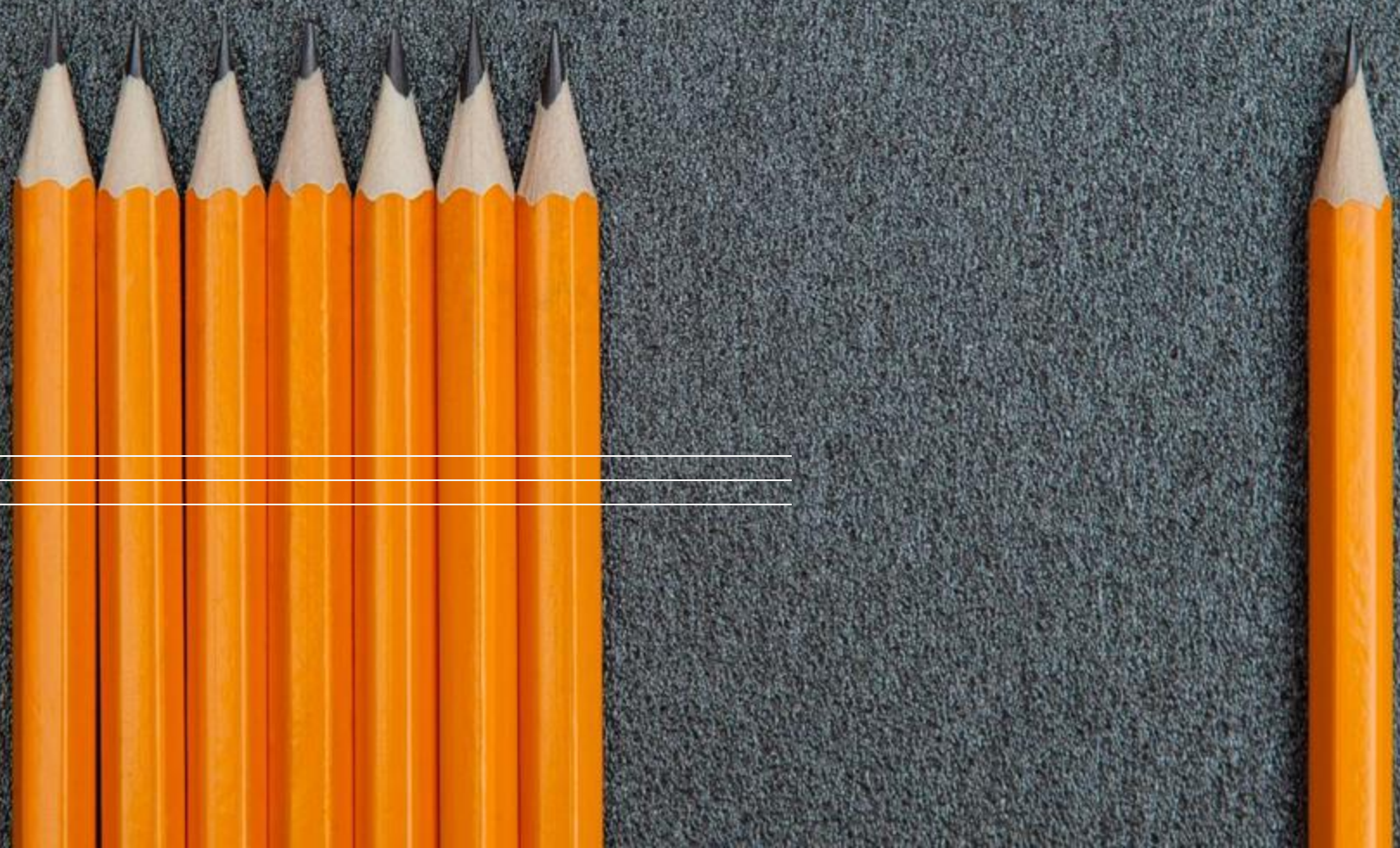
YOLO 자료조사

최인성

목차

Part 1,

정의



We introduce YOLO, a unified model for object detection.

Object Detection을 위한 통일된 모델입니다.

Title - You Only Look Once: Unified, Real-Time Object Detection

당신은 한번만 본다 : 통합된(통일된) , 실시간 객체 탐지

You only Look Once : 다른 모델들은 하나의 이미지를 여러번 본다(sliding window, region proposal 등)

Unified : 다른 모델들은 2 stage 방식(객체 크기, 객체 종류따로 예측)

Real-Time : 사람이 자연스럽게 느끼는 실시간 fps는 30 이상, yolo는 45fps 이상, fast yolo는 155 fps

Part 2,

목적

A single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance

Yolo 는 Object Detection 분야에서 성능을 최적화 하기 원했습니다. 그 결과 하나의 네트워크로 객체 크기와 객체 종류 예측을 동시에 하는 것으로 바꾸었습니다.

저자가 Yolo와 비교한 논문 Yolo가 real-time 으로 Detection이 가능하게 된 이유

2010

DPM : sliding window방식, disjoint한 pipeline

P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models

2014

R-CNN: multi pipeline 방식, selective search 사용

R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation

2014

Fast R-CNN : R-CNN을 개선하였지만 real-time에는 못미침

R. B. Girshick. Fast R-CNN

2013

OverFeat : sliding window를 효율적으로 수행하도록 바꿨지만 여전히 disjoint한 system

P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. C

2014

Deep MultiBox : Rcnn와 다르게 selective search 대신에 뉴럴 네트워크를 사용하였지만 larger 한 pipeline

D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition

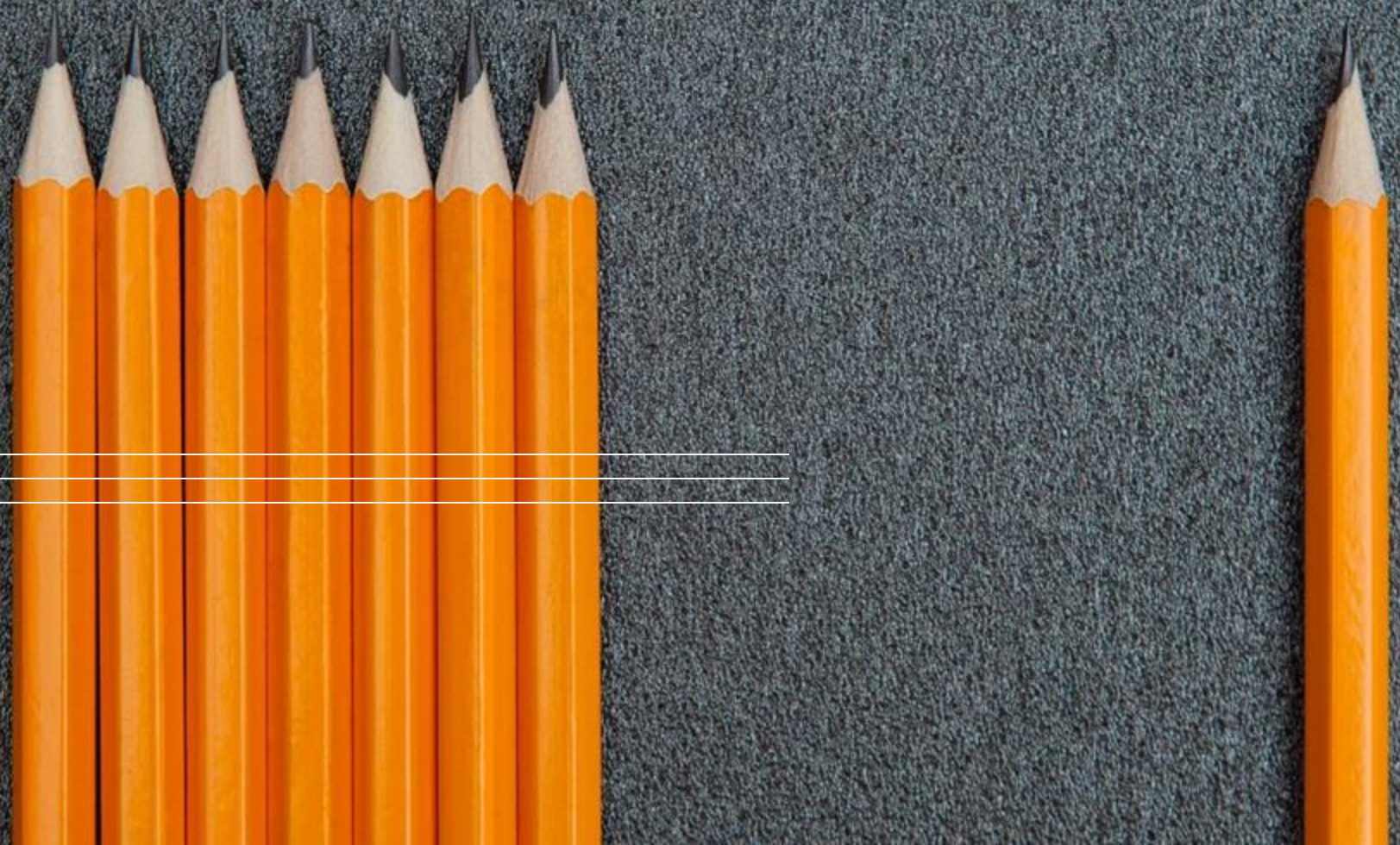
2014

MultiGrasp : yolo의 bounding box 예측 접근법에 기반이된 논문, 하지만 하나의 객체를 위한 하나의 지역만 예측할 뿐

D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition

Part 3,

방법



YOLO Detection 시스템

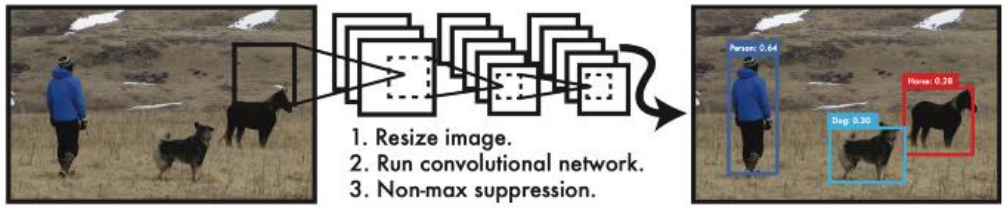


Figure 1: The YOLO Detection System. Processing images with YOLO is simple and straightforward. Our system (1) resizes the input image to 448 × 448, (2) runs a single convolutional network on the image, and (3) thresholds the resulting detections by the model's confidence.

1. Input image를 resize 한다.
2. Image가 single convolutional network를 통과한다.
3. 신경망의 결과로 객체의 bounding box와 해당 객체가 무엇인지 분류 확률을 출력한다.
4. 최종적으로 nms를 통해 Region을 결정한다.

감사합니다