Computer Vision (3주차)

Young-Gon Kim
DLI Instructor

DEEP LEARNING INSTITUTE

DLI Mission

Helping people solve challenging problems using AI and deep learning.

- Developers, data scientists and engineers
- Self-driving cars, healthcare and robotics
- Training, optimizing, and deploying deep neural networks

TOPICS

- Week 2 Review
- Object Detection

WEEK 2 REVIEW

OBJECT DETECTION

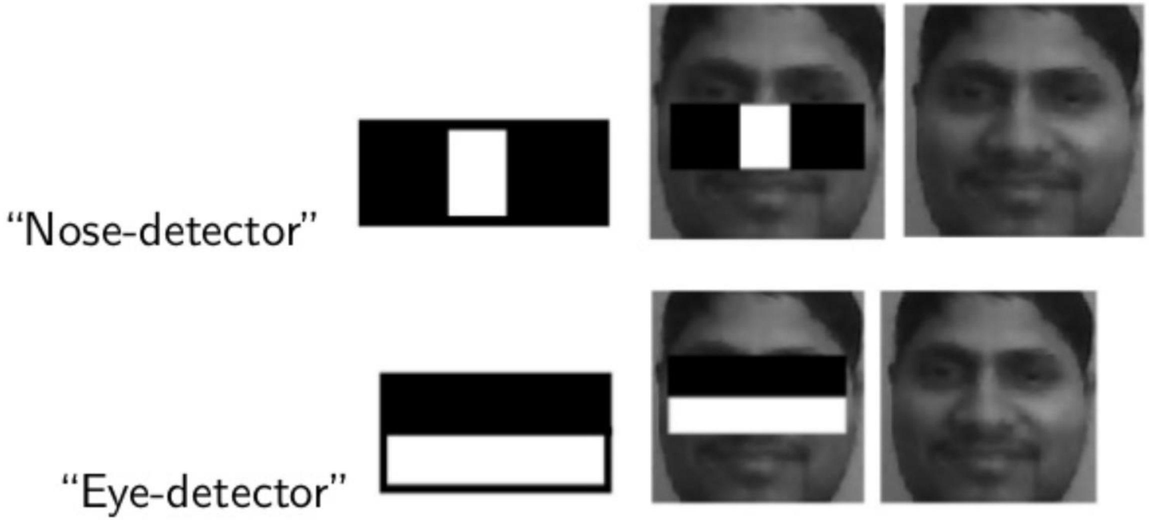
OBJECT DETECTION

- Object detection in the natural image
  - **History**
    - 초기 object detection 기법은 handcraft feature를 기준으로 함
    - **Sliding window classifier**

https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

OBJECT DETECTION

- Object detection in the natural image
  - **History**
    - Boosted classifier
    - **Haar-like features (Viola and Jones, 2001.)**

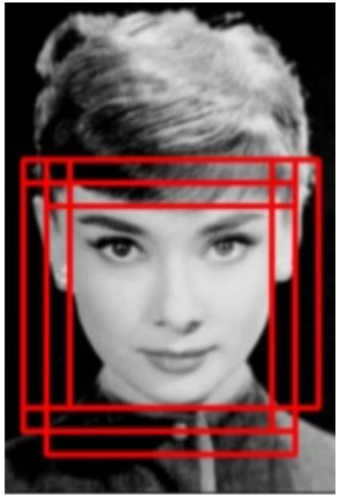


https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

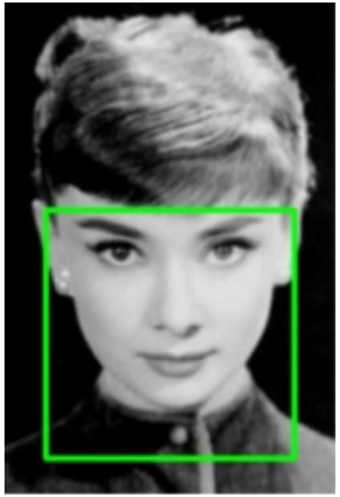
OBJECT DETECTION

- Object detection in the natural image
  - **History**
    - Sliding window classifier는 너무 많은 corrected detection을 만듬

      → **Non-Maximum Suppression (NMS)**



Before NMS

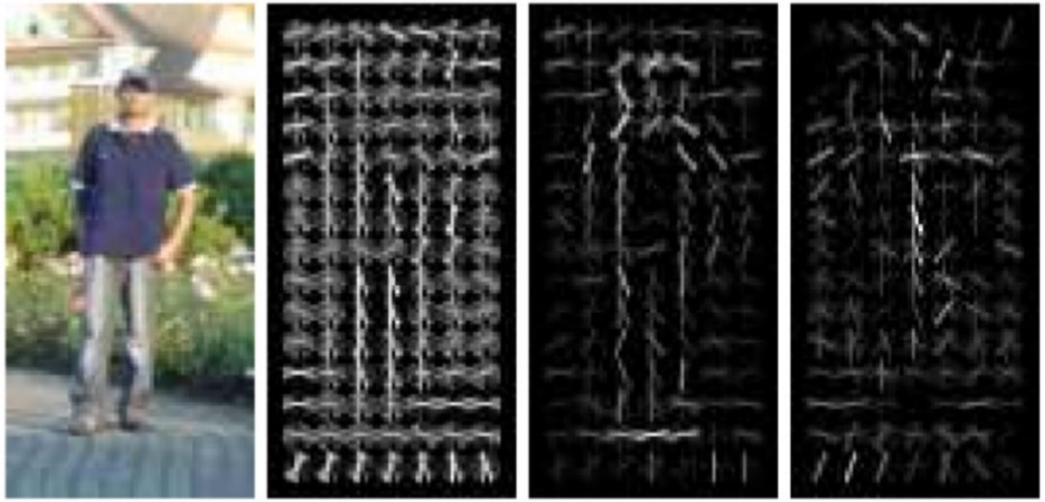


After NMS

https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

OBJECT DETECTION

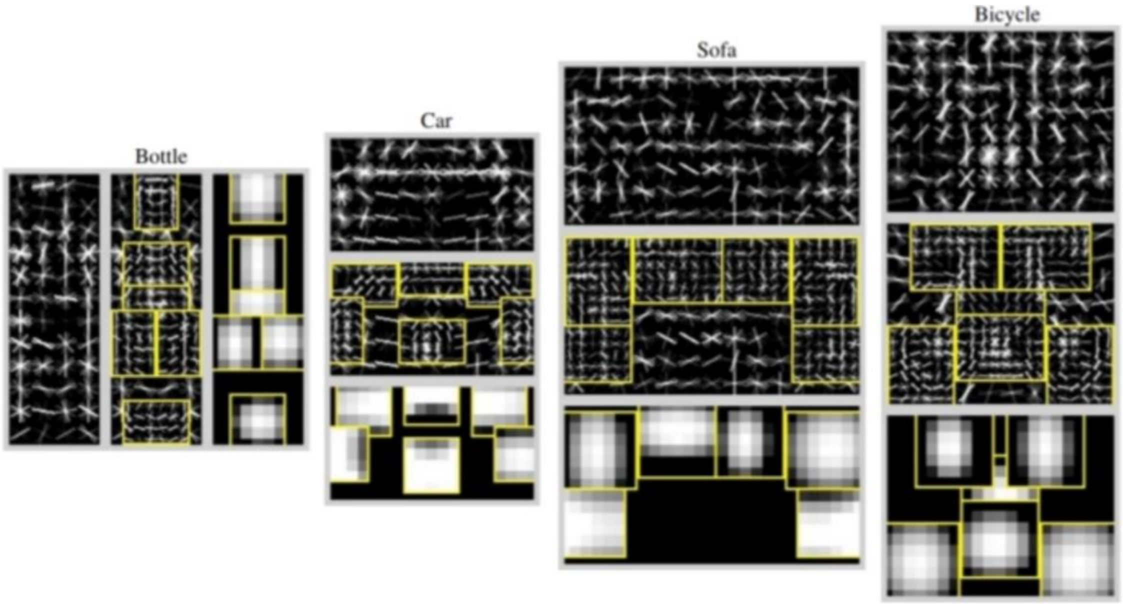
- Object detection in the natural image
  - **History**
    - Computes histogram of gradient orientation over sub-image blocks
    - **Histograms of Oriented Gradients (HOG) (Dalal and Triggs, 2005.)**



https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

OBJECT DETECTION

- Object detection in the natural image
  - **History**
    - Learn the relationships between HOG features of object parts via a latent SVM
    - **Deformable Part Models (DPM) (Felzenszwalb et al., 2008.)**

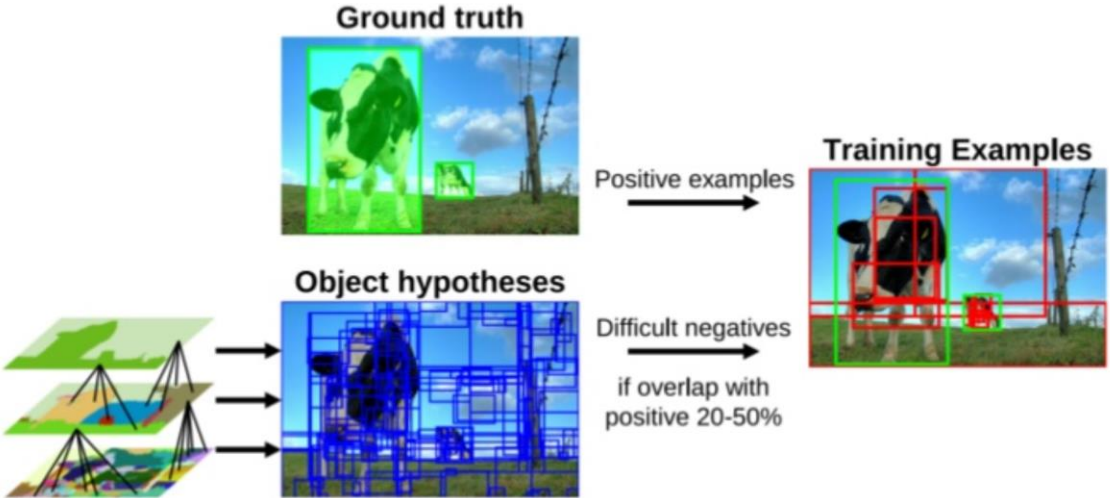


https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

NVIDIA DEEP LEARNING INSTITUTE

OBJECT DETECTION

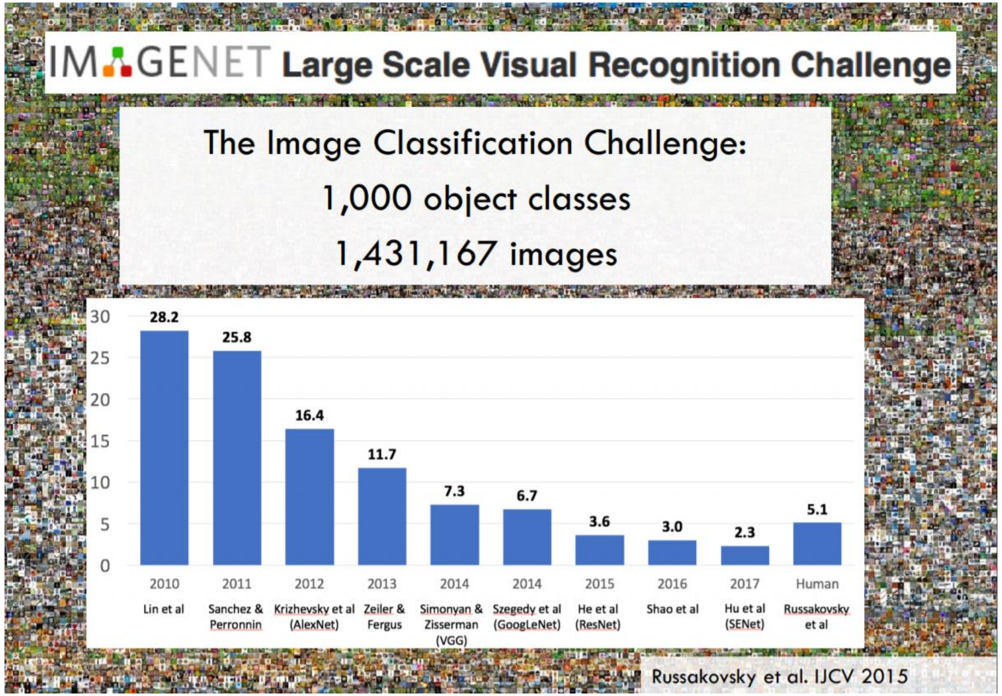
- Object detection in the natural image
  - **History**
    - Propose regions that have high “objectness”
    - Oversegment image and merge regions
    - **Selective Search (Uijlings et al., 2013.)**



https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola

OBJECT DETECTION

- Object detection in the natural image
  - **Competition**
    - **ImageNet Large Scale Visual Recognition Challenge (ILSVRC, ImageNet)**



http://image-net.org/index
http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture01.pdf

OBJECT DETECTION

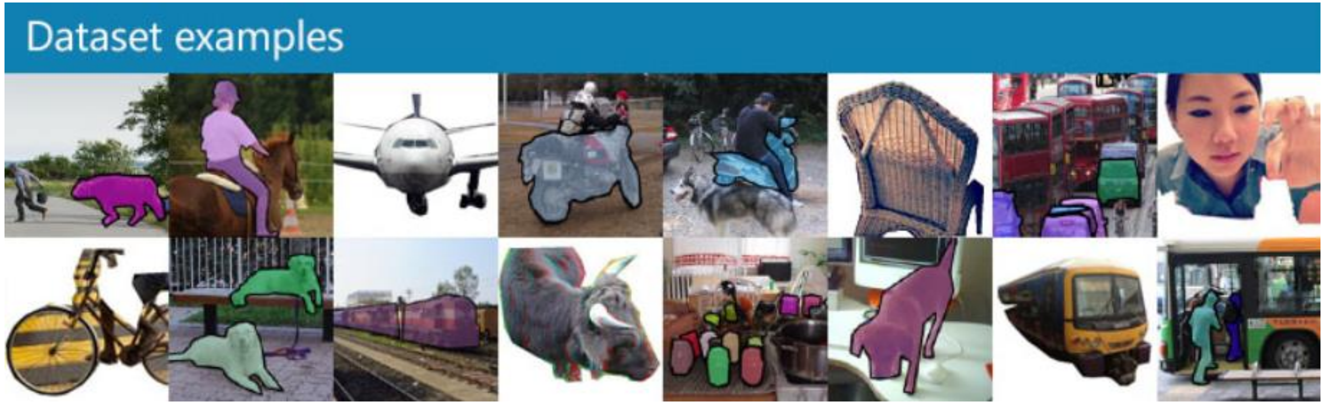
- Object detection in the natural image
    - **Competition**
        - **Visual Object Classes Challenge 2012 (PASCAL VOC 2012)**
            - Classification, Detection, Segmentation
            - # of classes : 20

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

http://host.robots.ox.ac.uk/pascal/VOC/voc2012/

OBJECT DETECTION

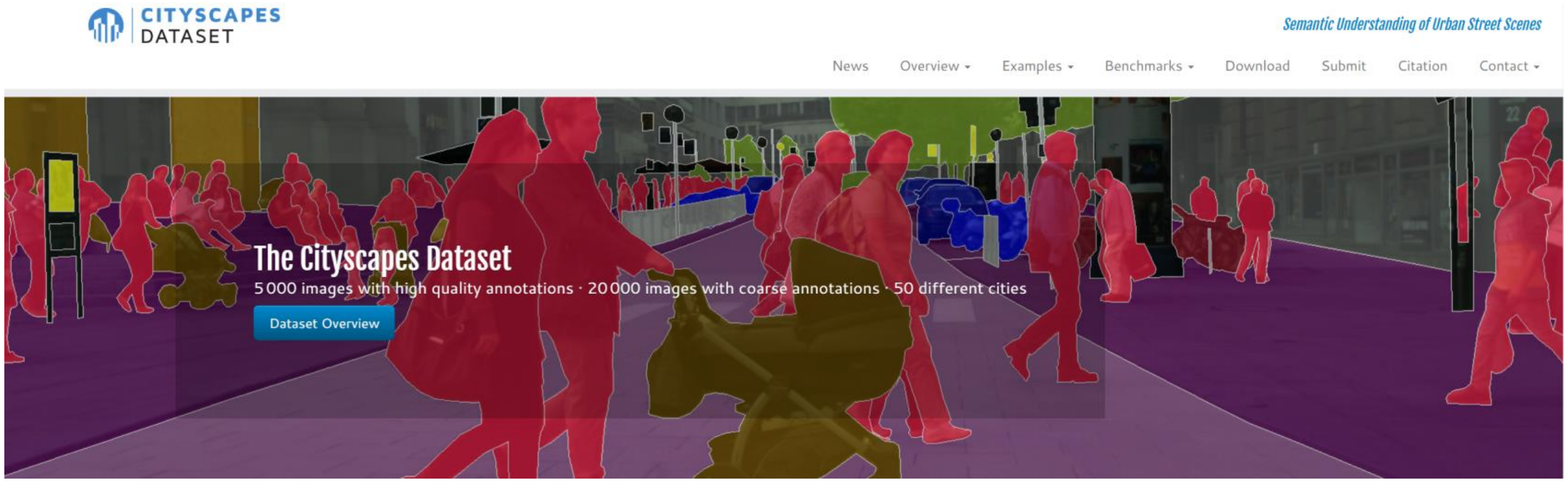
- Object detection in the natural image
  - **Competition**
    - **Microsoft Common Objects in COntest (MS COCO)**
      - Detection, Keypoint, Panoptic, DensePose
      - # of images : 330K (>200K labeled)
      - # of classes : 80



http://cocodataset.org/#home

OBJECT DETECTION

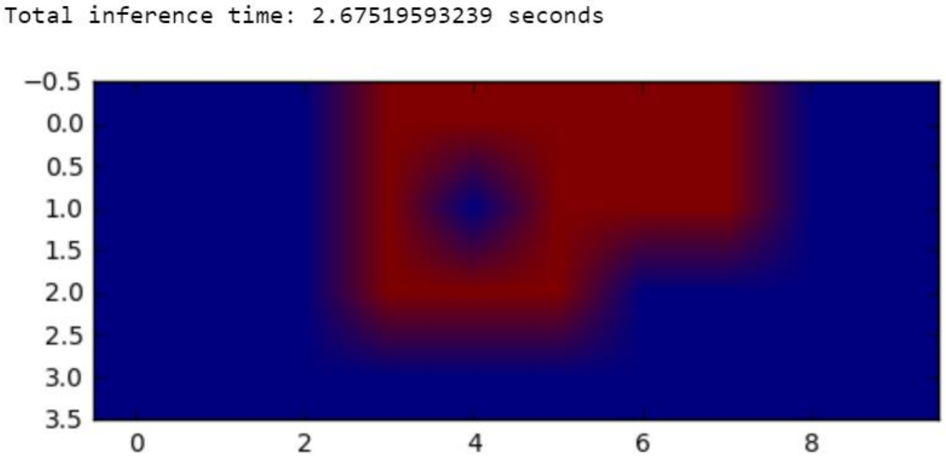
- Object detection in the natural image
  - **Competition**
    - **Cityscapes Dataset**
      - Semantic, Instance, Panoptic Segmentation
      - # of images : 5,000 (fine annotations) / 20,000 (coarse annotations)



https://www.cityscapes-dataset.com/

OBJECT DETECTION

- Object detection using **sliding window**
  - **Technique**
    - Build a dog / ‘not dog’ classifier
    - Sliding window python application runs classifier on each 256x256 segment



OBJECT DETECTION

- Architecture
  - **Two-stage detector**
    - R-CNN
    - SPPNet
    - Fast R-CNN
    - **Faster R-CNN**
    - **Feature Pyramid Networks (FPN)**
    - …
  - **One-stage detector**
    - **YOLO**
    - Single Shot MultiBox Detector (SSD)
    - **RetinaNet**
    - …

https://arxiv.org/pdf/1905.05055.pdf

OBJECT DETECTION

- Architecture
  - **Two-stage detector**
    - Faster R-CNN (S. Ren et al., TPAMI 2017.)
      - 대표적인 two-stage region proposal network

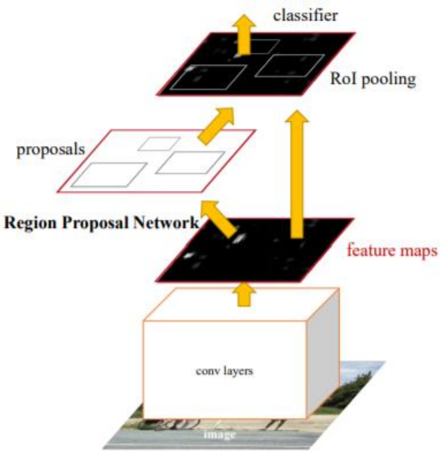


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

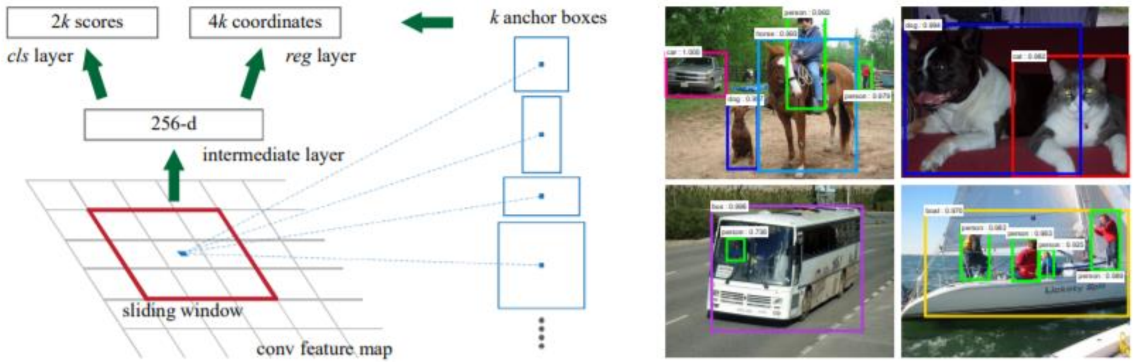


Figure 3: **Left**: Region Proposal Network (RPN). **Right**: Example detections using RPN proposals on PASCAL VOC 2007 test. Our method detects objects in a wide range of scales and aspect ratios.

https://arxiv.org/pdf/1506.01497.pdf

OBJECT DETECTION

- Architecture
  - **Two-stage detector**
    - Feature Pyramid Networks (T.-Y. Lin et al., CVPR 2017.)
      - Faster R-CNN 기반 Feature Pyramid Network 제안

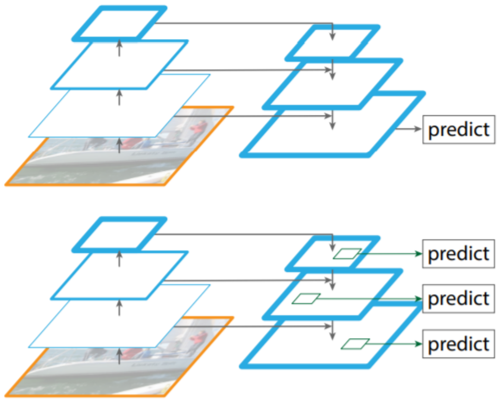
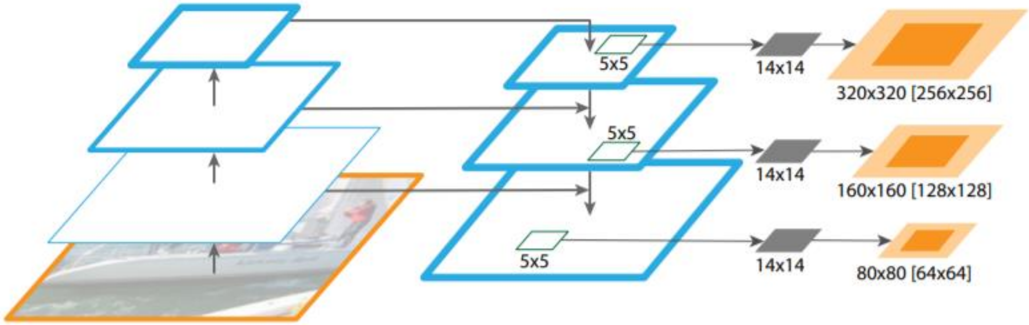


Figure 2. Top: a top-down architecture with skip connections, where predictions are made on the finest level (*e.g.*, [28]). Bottom: our model that has a similar structure but leverages it as a *feature pyramid*, with predictions made independently at all levels.



https://arxiv.org/pdf/1612.03144.pdf

OBJECT DETECTION

- Architecture
  - **One-stage detector**
    - You Only Look Once (YOLO) (J. Redmon et al., CVPR 2016.)
      - Grid + Sliding window

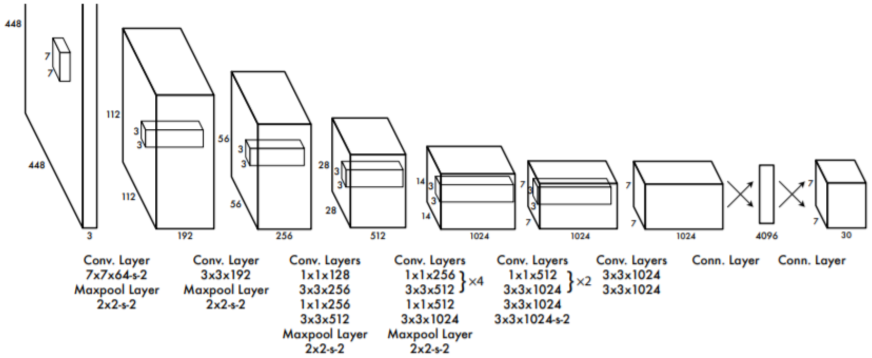


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

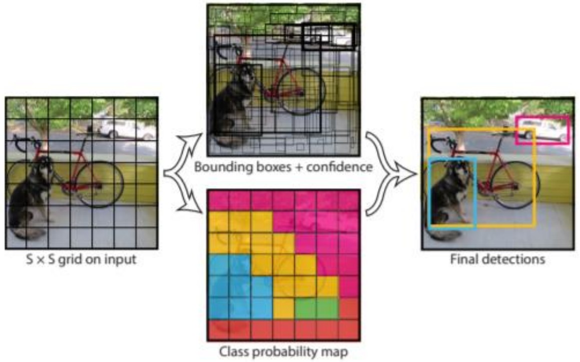
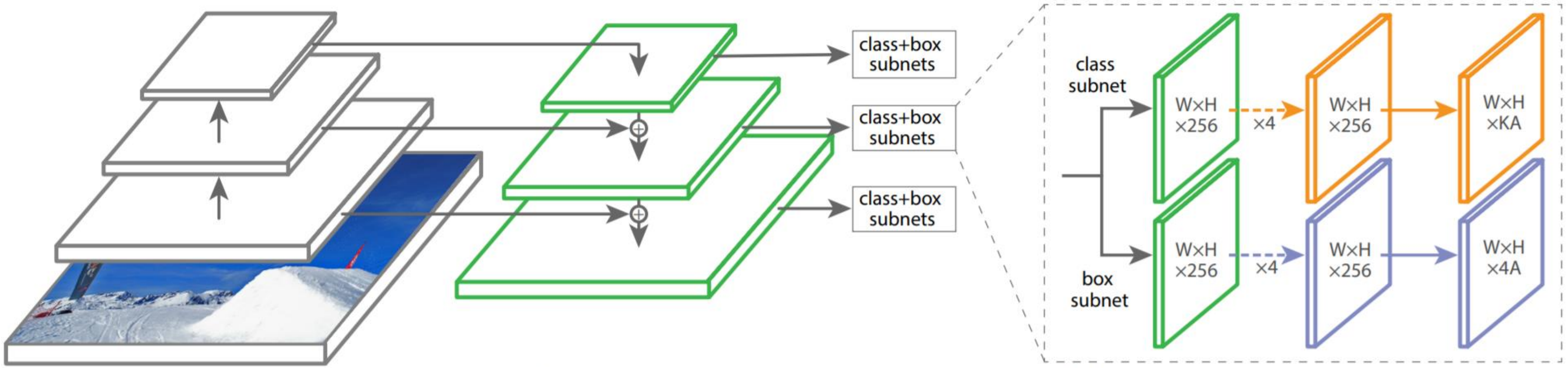


Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

https://arxiv.org/pdf/1506.02640.pdf

OBJECT DETECTION

- Architecture
  - **One-stage detector**
    - RetinaNet (T.-Yi. Lin et al., ICCV 2018.)
      - Feature Pyramid Network를 one-stage detector로 변화



(a) ResNet (b) feature pyramid net (c) class subnet (top) (d) box subnet (bottom)

https://arxiv.org/pdf/1708.02002.pdf

Reference

- https://www.slideshare.net/pfi/a-brief-history-of-object-detection-tommi-kerola
- http://image-net.org/index
- http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture01.pdf
- http://host.robots.ox.ac.uk/pascal/VOC/voc2012/
- http://cocodataset.org/#home
- https://www.cityscapes-dataset.com/
- https://arxiv.org/pdf/1905.05055.pdf
- Faster R-CNN : https://arxiv.org/pdf/1506.01497.pdf
- Feature Pyramid Networks : https://arxiv.org/pdf/1612.03144.pdf
- YOLO : https://arxiv.org/pdf/1506.02640.pdf
- RetinaNet : https://arxiv.org/pdf/1708.02002.pdf

DEEP
LEARNING
INSTITUTE

www.nvidia.com/dli