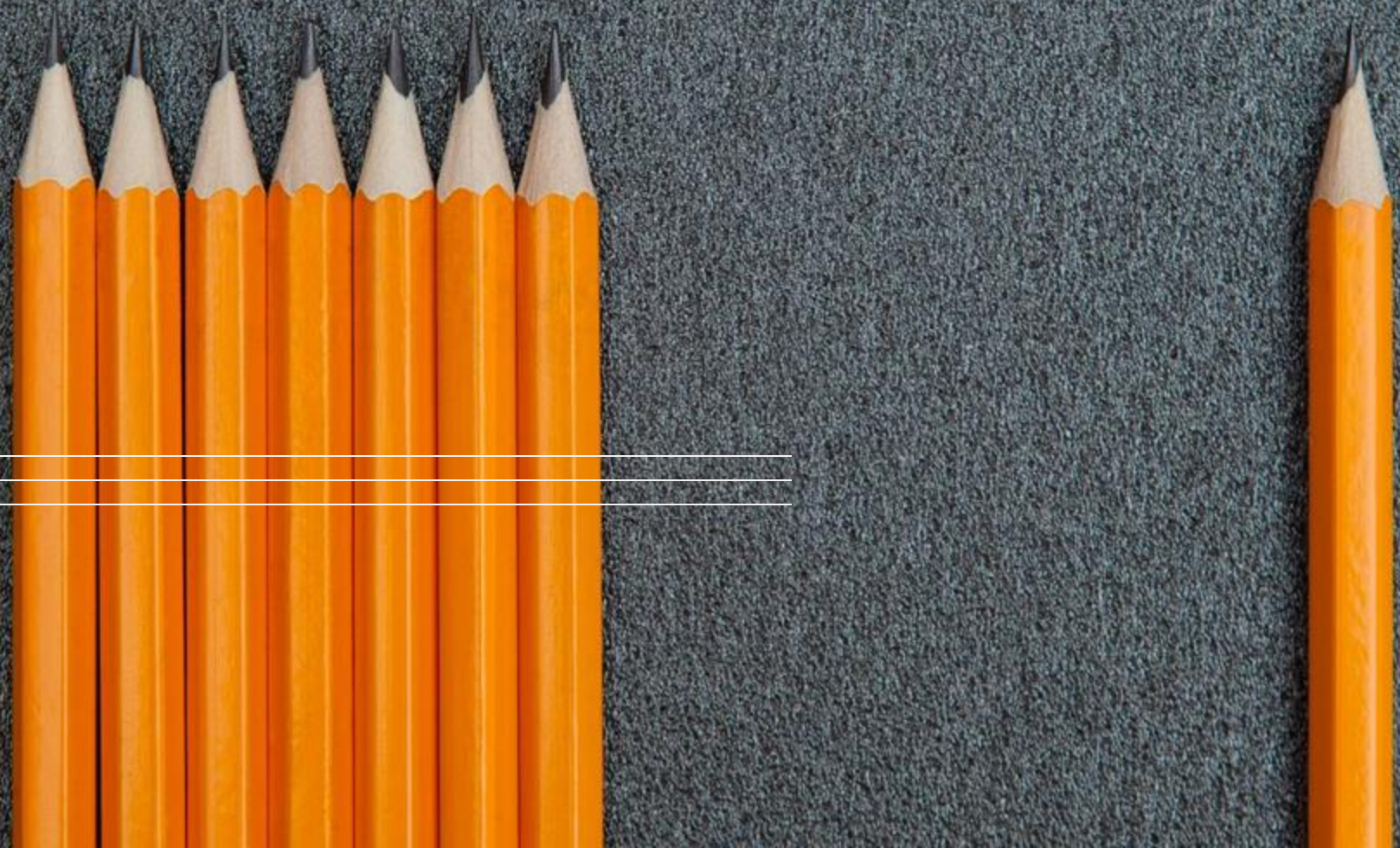
R-CNN 자료조사

최인성

목차

Part 1,

정의



Since we combine region proposals with CNNs, we call our method R-CNN: Regions with CNN features.

CNN과 Region proposal이 결합된
Object Detection 알고리즘입니다.

Object Detection 이란?

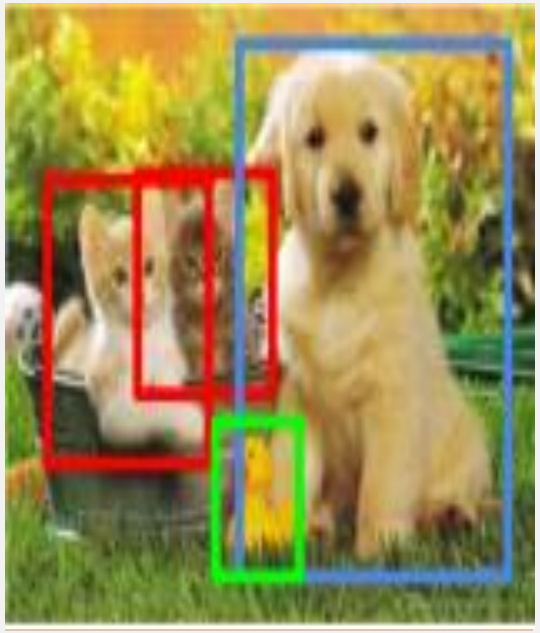
Classification

CAT

Classification + Localization

CAT

Object Detection



CAT,DOG,DUCK

Instance Segmentation

CAT,DOG,DUCK

Single object

Multiple Objects

Title - Rich feature hierarchies for accurate object detection and semantic segmentation

정확한 객체 인식과 분할을 위한 풍부한 기능을 가진 계층구조

Object detection = CNN기반 Alexnet구조 사용

Sematic segmentation = 픽셀 기준으로 region proposal 해서

hierarchies = R-CNN 구조가 단계별로 나뉘어서

Part 2,

목적

Object detection performance, as measured on the canonical PASCAL VOC dataset, has plateaued in the last few years. The best-performing methods are complex ensemble systems that typically combine multiple low-level image features with high-level context. In this paper, we propose a simple and scalable detection algorithm that improves mean average precision (mAP) by more than 30% relative to the previous best result on VOC 2012—achieving a mAP of 53.3%

2012년 전에 PASCAL VOC Object Detection분야에서 SIFT와 HOG가 많이 사용되었습니다. (비 딥러닝)
Object Detection분야에서 성능을 높이고자 딥러닝 기반의 R-CNN을 사용하였습니다.

R-CNN에 영감을 준 논문

R-CNN이 Object Detection분야에서 딥러닝이 가능하게된 이유

1989

SGD를 사용하여 CNN구조를 역전파로 학습할수 있게됨

Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel. Backpropagation applied to handwritten zip code recognition. Neural Comp.

2012

SVM의 밀려 CNN이 정체기였지만 이미지 분류 분야에서 CNN이 다시 상기됨

J. Deng, A. Berg, S. Satheesh, H. Su, A. Khosla, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Competition 2012 (ILSVRC2012).

2013

T. Dean, M. A. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, and J. Yagnik. Fast, accurate detection of 100,000 object classes on a single machine. In CVPR, 2013.

2009

R-CNN에 localization을 당담하는 region proposal을 사용하게된 논문

C. Gu, J. J. Lim, P. Arbelaez, and J. Malik. Recognition using re- ´ gions. In CVPR, 2009.

2013

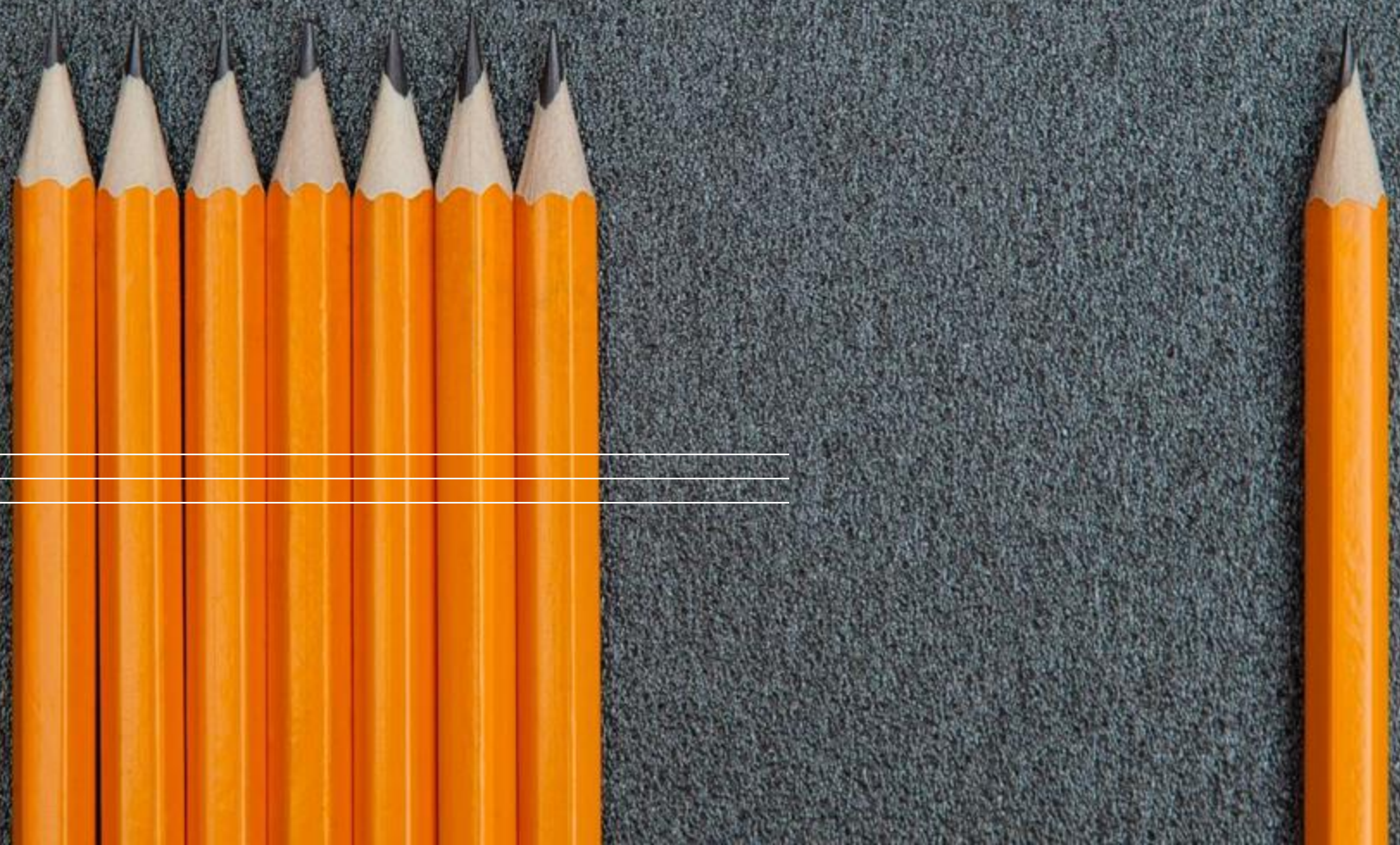
object detection분야에서 딥러닝구조의 시도가있었으나 localization 에서 sliding window 방식을 써서 mAP가 낮았음

C. Szegedy, A. Toshev, and D. Erhan. Deep neural networks for object detection. In NIPS, 2013.

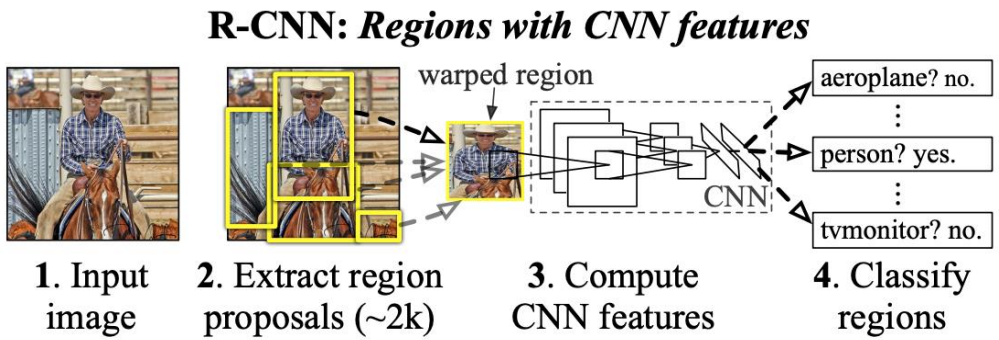
ⓒSaebyeol Yu. Saebyeol's PowerPoint

Part 3,

방법



R-CNN 전체적인 프로세스



*Our object detection system consists of three modules. The first generates category-independent **region proposals**. These proposals define the set of candidate detections available to our detector. The second module is a large convolutional neural network that **extracts a fixed-length feature vector** from each region. The third module is a set of classspecific **linear SVMs**.*

1. Input 이미지로부터 2,000개의 독립적인 region proposal을 생성
2. CNN을 통해 각 proposal 마다 고정된 길이의 feature vector를 추출
3. 이후, 각 region 마다 category-specific linear SVM을 적용하여 classification을 수행

감사합니다

참고 - https://medium.com/hyunjulie/1%ED%8F%B8-semantic-segmentation-%EC%B2%AB%EA%B1%B8%EC%9D%8C-4180367ec9cb