

이름: 문준원

Target Paper	<p>논문제목: Region-based Convolutional Networks for Accurate Object Detection and Segmentation</p> <p>게재 저널명, 권/호/페이지 정보: IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 38, Issue: 1, Pages: 142-158</p> <p>게재년월일: 25 May 2015</p> <p>저자(들): Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik</p>
논문 요약	<p>1. What: 무엇이 주된 이슈인가?</p> <p>이 논문의 주요 이슈는 객체 검출 성능이 PASCAL VOC 챌린지에서 정해진 상황에서 기존 방법들이 복잡한 양상블 시스템에 의존하고 있다는 점입니다. 특히, 저수준 이미지 특징과 고수준 문맥을 조합한 복잡한 시스템들이 주로 사용되고 있었지만, 더 나은 검출 성능을 확보하는 데 한계가 있었습니다. 연구 초반에는 두 가지 고민이 있었습니다: (1) 딥 네트워크를 이용하여 객체를 이미지 내에서 어떻게 정확히 국지화할 수 있을지, 즉 바운딩 박스를 어떻게 그릴지에 대한 문제와 (2) 적은 양의 라벨된 데이터로 고용량 모델을 효과적으로 학습하는 방법이었습니다. 기존에는 비지도 사전 학습과 지도 기반 미세 조정 방법이 사용되었습니다.</p>
	<p>2. How: 어떻게 문제에 접근했나?</p> <p>논문은 간단하면서도 확장 가능한 R-CNN(Region-based Convolutional Neural Network) 알고리즘을 제안하여, 평균 정밀도(mAP)를 기존 최고 성과보다 50% 이상 향상시켰습니다. R-CNN 은 하위 지역 제안을 CNN 에 적용하여 객체를 국지화하고 분할하며, 특히 (1) 객체를 국지화하는 문제를 해결하기 위해 Recognition Using Regions 기법을 사용하여 Detection 과 Segmentation 모두에서 뛰어난 성능을 보입니다. (2) 적은 양의 레이블된 데이터 문제를 해결하기 위해, ILSVRC 대량 데이터셋에서 Supervised Pre-training 후, PASCAL 과 같은 소규모 데이터셋에서 도메인별 미세 조정을 수행하여 효과적인 성능을 달성했습니다. 이러한 접근법은 데이터가 적을 때도 높은 성능을 발휘할 수 있는 중요한 문제 해결법으로 작용했습니다.</p>
	<p>3. Cons & Pros: 문제점과 개선 방안은 무엇인가?</p> <p>R-CNN 의 성능 개선을 위한 접근에는 두 가지 중요한 논점이 있습니다. 첫째, CNN 의 fine-tuning 과정에서 positive 와 negative 예시가 다르게 정의되는 이유는 데이터 부족에서 기인하며, 이 정의는 정확한 localization 을 강조하지 않습니다. 둘째, fine-tuning 후 SVM 을 학습하는 이유는 softmax 분류기를 사용하는 것보다 성능이 높기 때문입니다. 이 성능 차이는 SVM 이 “hard negatives”에 대해 학습하는 반면, softmax 는 무작위 negative 예시로 학습하기 때문입니다. 따라서 추가적인 fine-tuning 수정으로 성능 격차를 줄일 수 있으며, 이는 R-CNN 의 학습을 단순화하고 가속화할 가능성이 있습니다.</p>