

제목 : Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition

1. 목적

이미지 인식 분야에서 이미지의 크기나 스케일에 관계없이 고정 길이의 표현을 생성하는 것은 중요한 것. 이때 풀링이 사용됨.

공간 피라미드 풀링이라는 새로운 풀링 전략 제안.

공간 피라미드 풀링을 사용하면 객체 변형에 강한 표현을 얻을 수 있음.

2. 핵심 기술

- SPP-net은 고정 크기의 입력 이미지를 요구하지 않고, 공간 피라미드 풀링을 통해 이미지의 크기나 스케일에 관계없이 고정 길이의 표현을 생성할 수 있음.

객체 변형에 강함.

- SPP-net은 ImageNet 2012 데이터셋에서 다양한 CNN 아키텍처의 정확도를 향상시키며, Pascal VOC 2007 및 Caltech101 데이터셋에서 최고의 분류 결과가 나옴.

또한 객체 탐지도 잘함.

-SPP-net을 사용하면 전체 이미지에서 특징 맵을 한 번만 계산한 다음 임의의 영역에서 풀링하여 검출기를 학습하는 데 사용할 수 있음.

따라서, 맵 계산 비용을 줄이면서도 뛰어난 성능을 얻을 수 있음.

- 테스트 이미지 처리 시 R-CNN 방법보다 빠르면서도 더 나은 또는 비슷한 정확도가 나옴.

- SPP-net은 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014에서는 객체 탐지에서 2위, 이미지 분류에서 3위를 차지함.