



6-2,3장



6-2장. 객체 인식을 위한 신경망

객체 인식: 이미지나 영상 내에 있는 객체를 식별하는 컴퓨터 비전 기술

* 어떤 객체인지(분류)+어디에 위치하는지(위치 검출-localization)

1 단계 객체 인식: 분류와 위치 검출을 동시에 실행

2 단계 객체 인식: CNN을 처음으로 적용시킨 R-CNN계열

6-2-1. R-CNN

이미지에서 객체가 있을 만한 영역을 제안해 주는 후보 영역 알고리즘을 결합한 것.

이미지 삽입 → 후보 영역 추출(cropping, warping) → CNN 특성 계산 → 영역 분류

* 선택적 탐색

object detect을 위해 가능한 후보 영역을 알아내는 것. (bounding box)

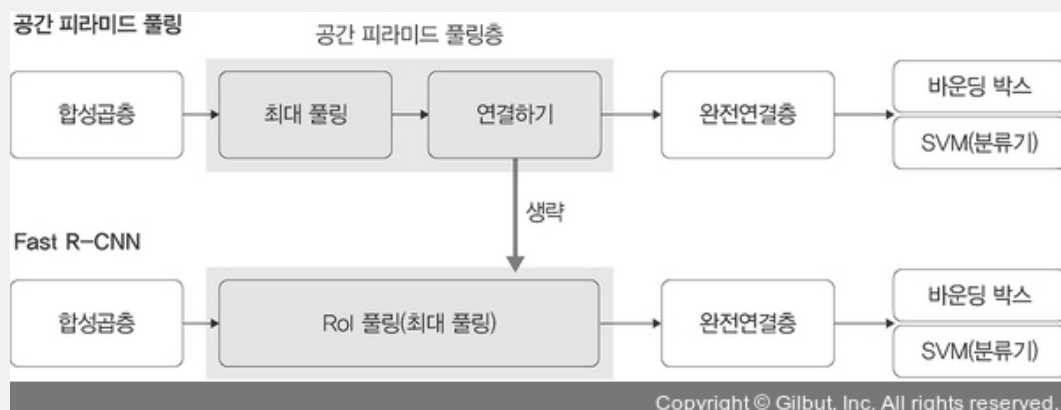
6-2-2. 공간 피라미드 풀링

기존 CNN 구조: 모두 FC layer로 입력 이미지를 고정해야 했음. 신경망을 통과 시키기 위해 crop와 warp과정이 필수적이었다. → 피라미드 풀링은 입력 이미지의 크기에 관계없이 conv layer를 통과, fc layer에 전달하기 전에 특성 맵을 동일한 크기로 조절하여 풀링층을 적용. (원본 이미지의 특징이 훼손되지 않음)

6-2-3. Fast R-CNN

R-CNN의 속도 문제를 개선(RoI 풀링을 도입)

RoI 풀링: 바운딩 박스가 CNN을 통과하면서 정보를 유지하도록 하고 최종 특성 맵은 풀링을 적용하여 시간 단축.



*RoI 풀링

→ 크기가 다른 특성 맵의 영역마다 스트라이드를 다르게 하여 maxpooling을 적용하여 결과적으로 크기가 동일하게 되도록 맞춘다.

6-2-4. Faster R-CNN

후보 영역 생성을 CNN 내부 네트워크에서 진행할 수 있도록 설계함. 기존의 Fast R-CNN에 후보 영역 추출 네트워크를 추가한 것이 핵심.



6-3장. 이미지 분할을 위한 신경망

image segmentation: 신경망을 훈련시켜 픽셀 단위로 분할하는 것.

6-3-1. 완전 합성곱 네트워크

fc layer의 한계: 고정된 크기의 입력만 받아들여, 완전 연결층을 지나면 위치 정보가 사라진다. → fc layer를 1x1 합성곱으로 대체함.

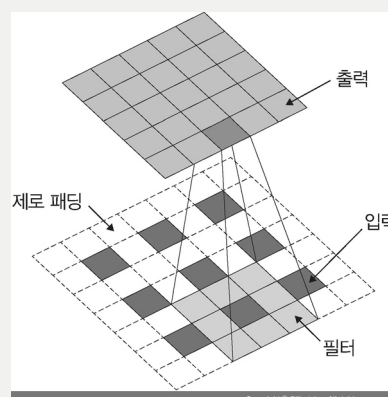
→ 합성곱 층으로 사용되기 때문에 입력 이미지에 대한 크기 제약이 사라진다.

6-3-2. 합성곱, 역합성곱 네트워크

완전 합성곱 네트워크는 위치 정보는 보존되나 해상도가 낮아지고 이런 해상도를 복구하기 위해 업 샘플링 방식을 사용하여 이미지의 세부 정보를 잃어버리는 문제가 발생한다.

업 샘플링(역합성곱): 특성 맵 크기를 증가시키는 방식으로 동작함.

→ 각 픽셀 주위에 제로 패딩을 적용하여 합성곱을 수행.



6-3-3. U-Net

바이오 이미지 분할에 많이 사용함.

이미 검증이 끝난 패치는 건너뛰어 속도가 빠르다.

트레이드 오프 해결: 패치 크기가 커지면 context인식은 좋으나 localization에는 한계가 있다. 하지만 U-net은 두마리 토끼 다 잡음.

FCN기반(contracting path+expansive path)

수축 경로: 컨텍스트 정보, 3x3 합성곱 두 개로 구성된 것이 네 개가 있는 형태 / 최대 풀링을 이용하여 다음 블록으로 넘어감.

확장 경로: 지역화 정보, 합성곱 앞에 up-conv붙여 수축 과정에서 줄어든 것을 다시 확장함

