

6.2 / 6.3 예습과제

▼ 6.2 객체인식을 위한 신경망

- 객체 인식 : 이미지나 영상 내에 있는 객체를 식별하는 컴퓨터 비전 기술
- 객체가 무엇인지 분류하는 문제 / 객체 위치가 어디인지 박스로 (bounding box)로 나타내는 위치검출(localization)문제를 다루는 분야
- 1단계 객체 인식 : 두문제를 동시에 행하는 방법
 - 비교적 빠르지만, 정확도가 낮음
 - YOLO 계열과 SSD 계열등이 포함됨
- 2단계 객체 인식 : 순차적으로 행하는 방법
 - 비교적 느리지만 정확도가 높음
 - R-CNN 계열이 대표적

▼ 6.2.1 R-CNN

- 예전의 알고리즘들은 윈도운 방식, 일정한 크기를 가지는 윈도우를 가지고 모든 영역을 탐색하면서 객체를 검출해 내는 방식 → 비효율성 때문에 많이 사용하지 않았음 → 현재는 **선택적 탐색 알고리즘**을 이용한 **후보영역**을 많이 사용
- R-CNN은 이미지 분류를 수행하는 CNN과 이미지에서 객체가 있을 만한 영역을 제안해주는 후보 영역 알고리즘을 결합한 알고리즘
 - 이미지를 입력을 받는다
 - 2000개의 바운딩박스를 선택적 탐색 알고리즘으로 추출 → 잘라내고 → 같은 크기로 통일
 - 2000개에 각각 CNN모델 적용
 - 각각 분류를 진행
- 선택적 탐색 알고리즘이란?
 - 후보영역을 알아내는 방법. 분할 방식을 이용하여 시드를 정하고, 그 시드에 대한 완전 탐색
 - 1단계 : 초기 영역 생성(sub-segmentation)
 - 입력된 이미지를 영역 다수 개로 분할하는 과정
 - 2단계 : 작은 영역의 통합

- 여러개로 나눈 것들을 비슷한 영역으로 통합. 탐욕알고리즘을 사용하여 비슷한 영역이 하나로 통합될 때까지 반복

○ 3단계 : 후보 영역 생성

- 후보영역(바운딩박스)를 추출

- 하지만 복잡한 학습과정/ 긴학습시간, 대용량 저장공간/객체검출 속도문제 의 단점이 있음

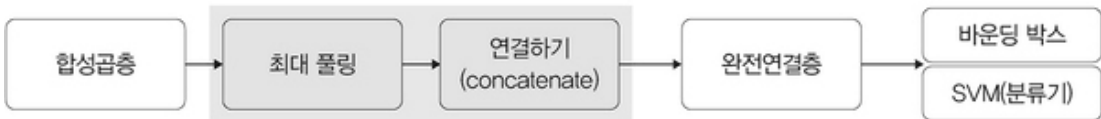
▼ 6.2.2 공간피라미드 풀링

- 신경망을 통과시키려면 이미지를 고정된 크기로 자르거나, 비율을 조정해야함. 이렇게 하면 물체의 일부분이 잘리거나 본래의 생김새와 달라지는 문제점 발생
→ 공간피라미드 풀링
- 이미지의 크기에 관계없이 합성곱층을 통과시키고, 완전연결층 전에 동일한 크기로 통일. 원본이미지의 특징이 훼손되지 않음

R-CNN



공간 피라미드 풀링



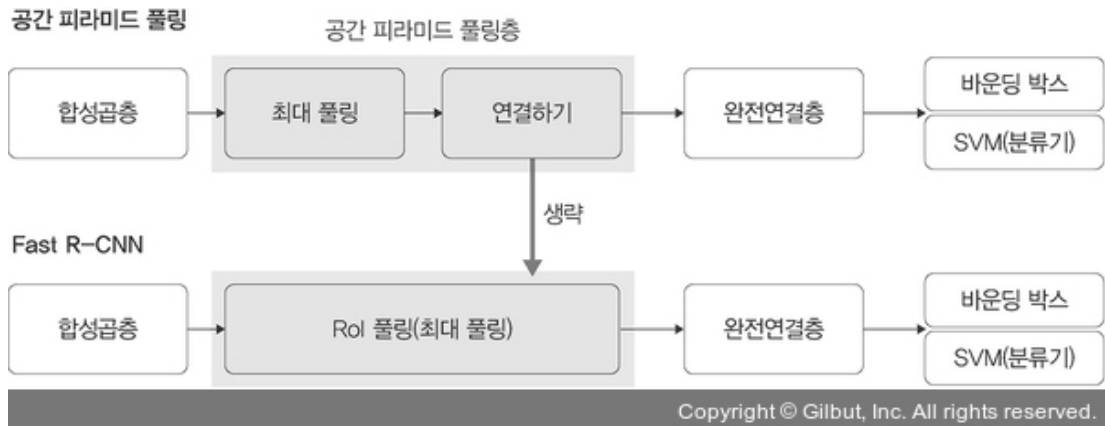
공간 피라미드 풀링층

Copyright © Gilbut, Inc. All rights reserved.

▼ 6.2.3 Fast R-CNN

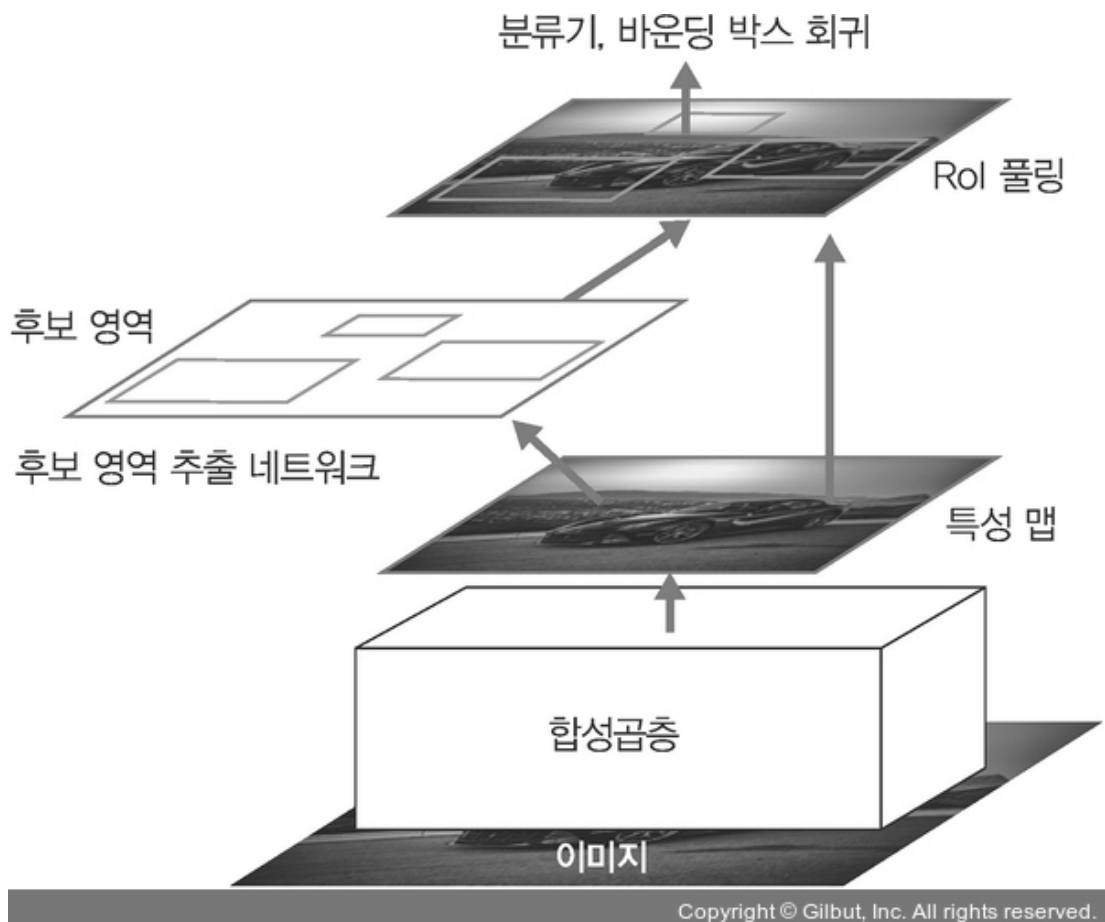
- R-CNN은 분류를 위한 긴 학습시간이 문제였음. 속도문제를 개선하려고 RoI풀링 도입. (영역마다 스트라이드를 다르게 최대풀링을 적용하여 결괏값 크기를 동일하게 맞추는 방법)

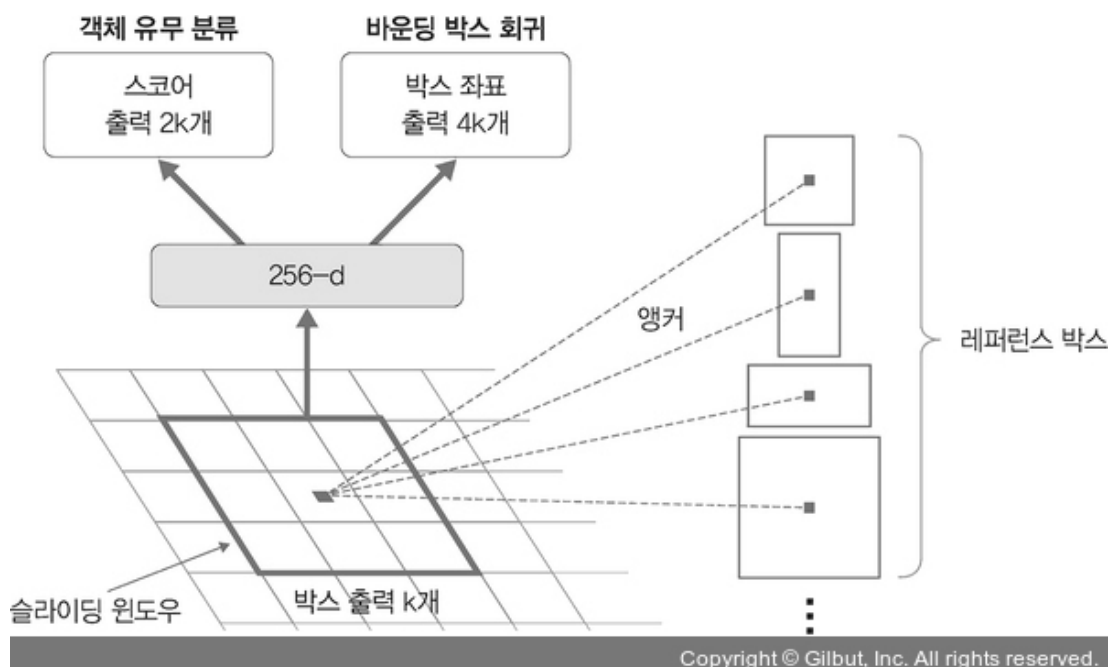
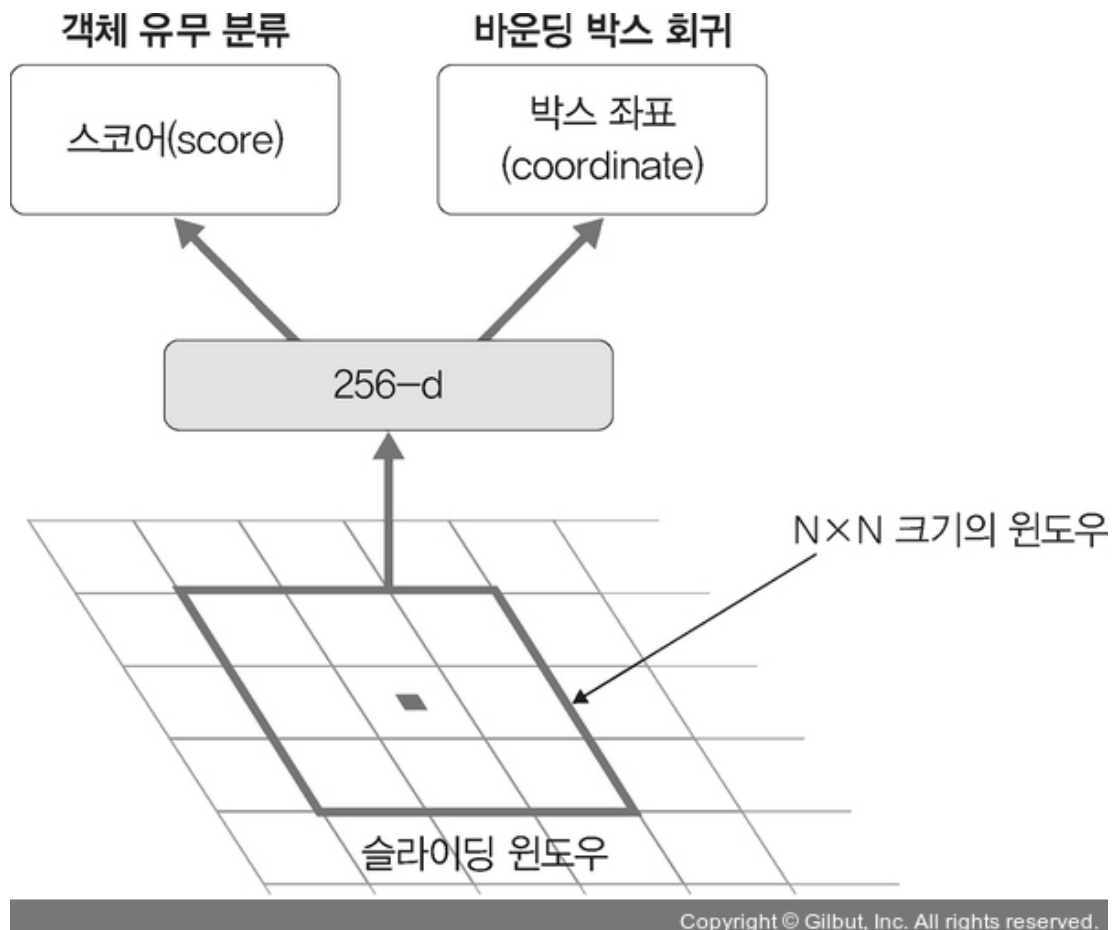
•



▼ 6.2.4 Faster R-CNN

- 후보 영역 생성을 cnn 내부 네트워크에서 진행할 수 있도록 설계. 기존의 Fast R-CNN에 후보 영역 추출 네트워크를 추가!!



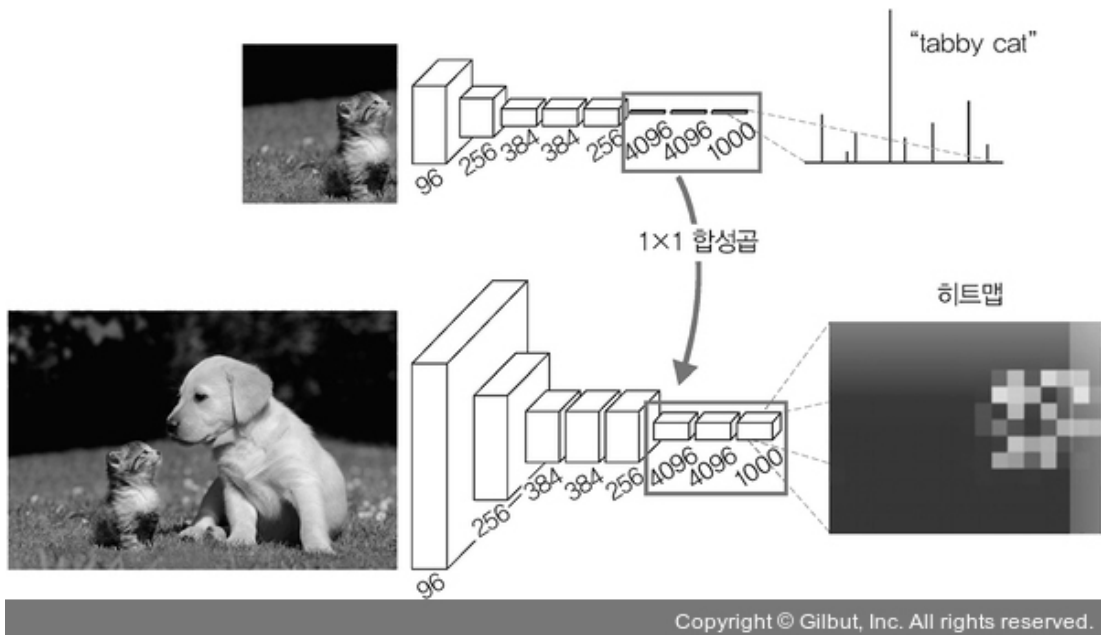


▼ 6.3 이미지 분할을 위한 신경망

이미지 분할 : 신경망을 훈련시켜 이미지를 픽셀 단위로 분할하여 이미지에 포함된 객체를 추출하는 것

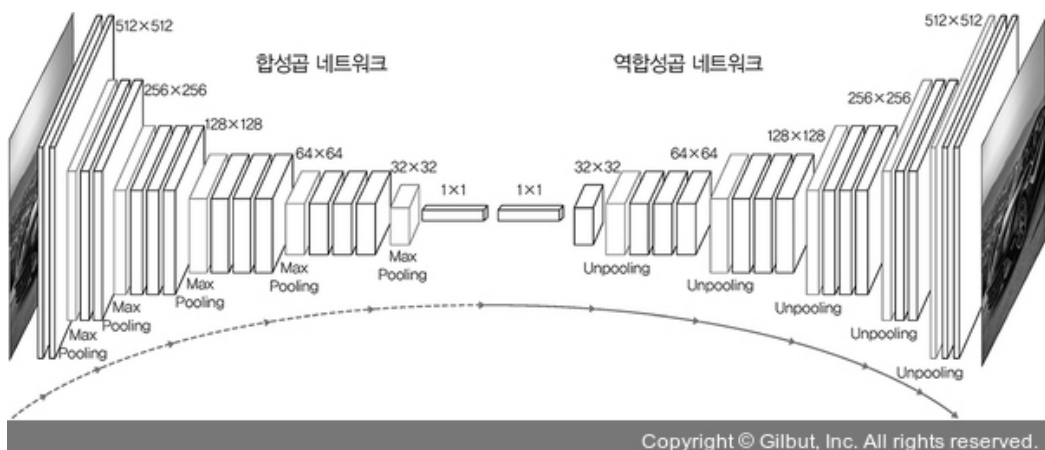
▼ 6.3.1 완전 합성곱 네트워크

- 완전 연결층의 한계 : 고정된 크기의 입력만 받아들이며, 완전 연결층을 거친 후에는 위치정보가 사라진다는 것
- → 1x1 합성곱으로 대체
-

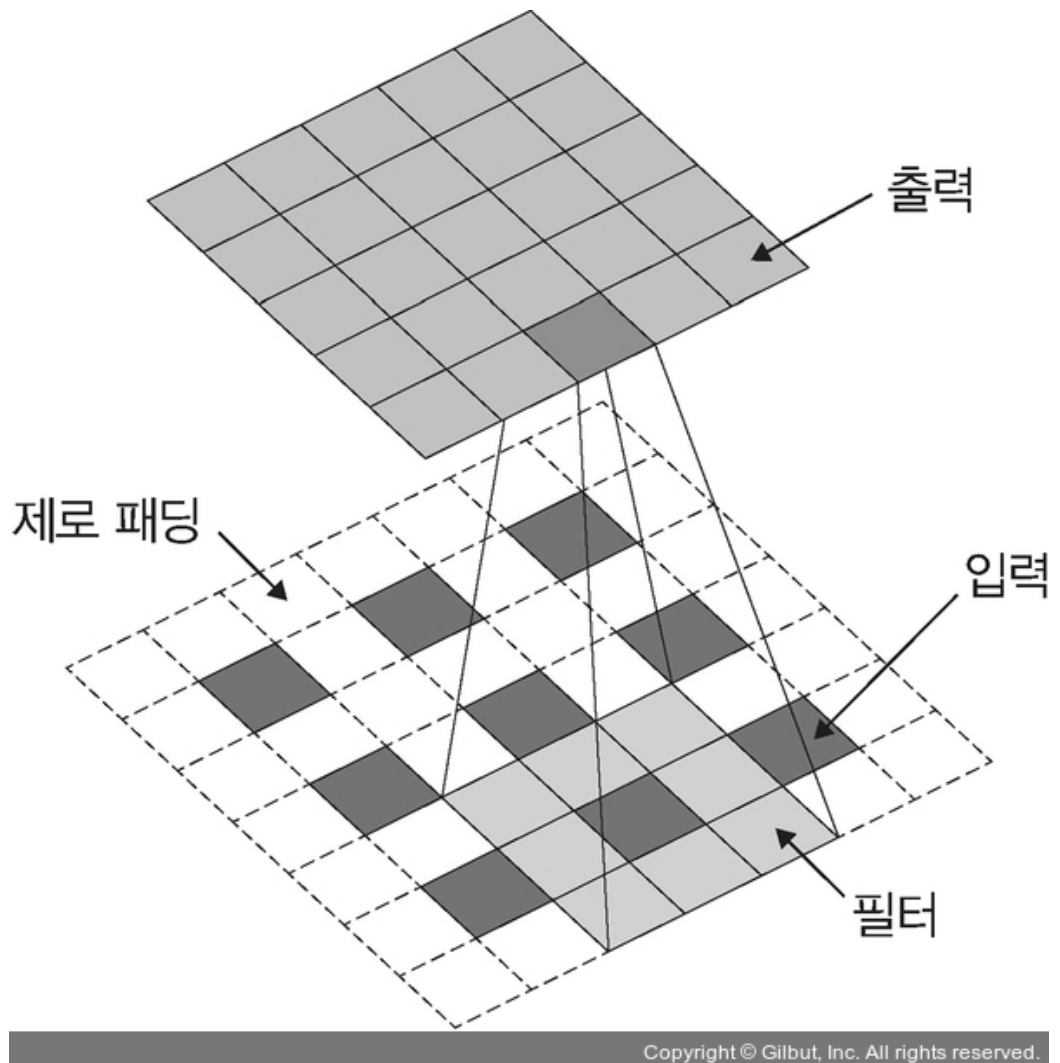


▼ 6.3.2 합성곱 & 역합성곱 네트워크

- 완전 합성곱 네트워크의 한계
 - 해상도가 낮아짐 → 복원하기 위해 업 샘플링 방식을 사용하기 때문에 이미지의 세부정보들을 잃어버리는 문제 발생
-

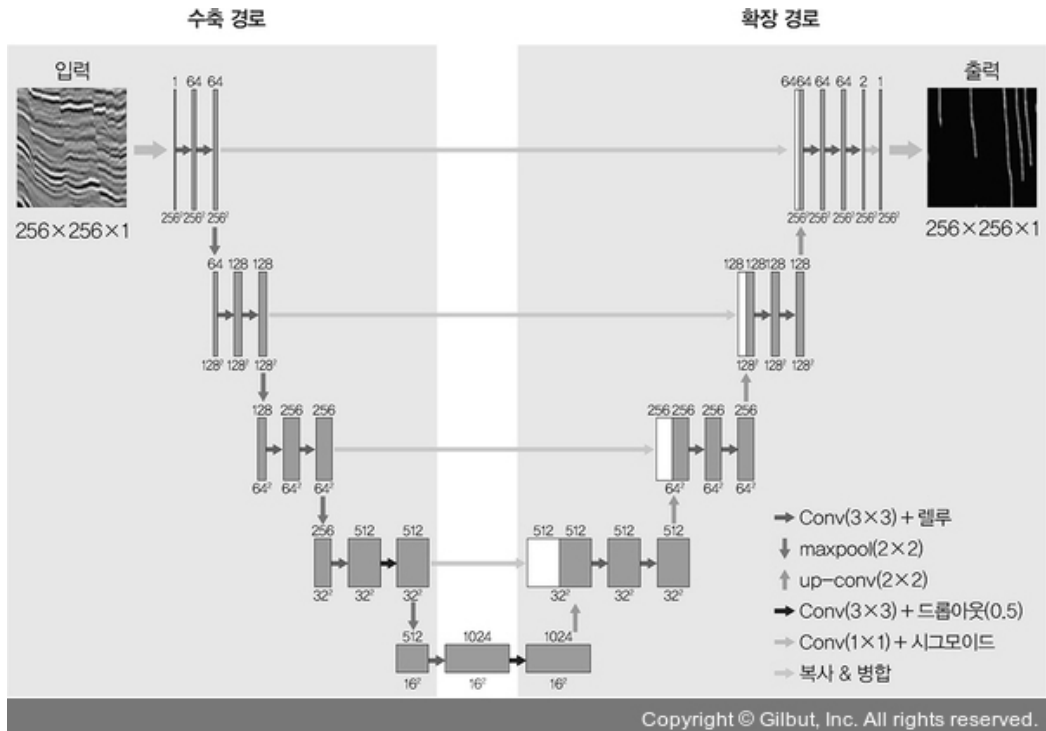


- 역합성곱(=업샘플링) : 시멘틱분할등에 활용 가능
 - 각각의 픽셀 주위에 제로 패딩 추가
 - 패딩된 것에 합성곱 연산 수행
 -



▼ 6.3.4 U-NET

- 바이오 메디컬 이미지 분할을 위한 합성곱 신경망
- 속도가 빠름
- 트레이드 오프에 빠지지 않는다.
- FCN기반으로 구축
- 수축경로(컨텍스트 포착)와 확장경로(업샘플링, 컨텍스트와 결합하여 정확한 지역화)로 구성



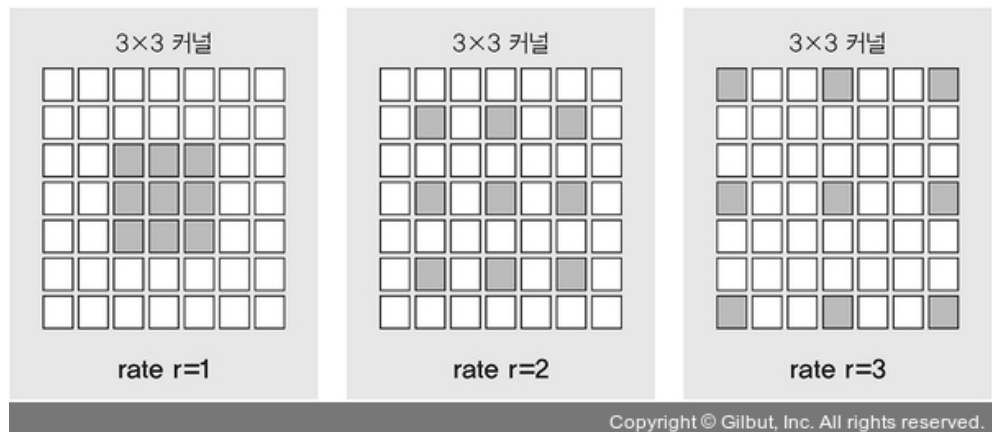
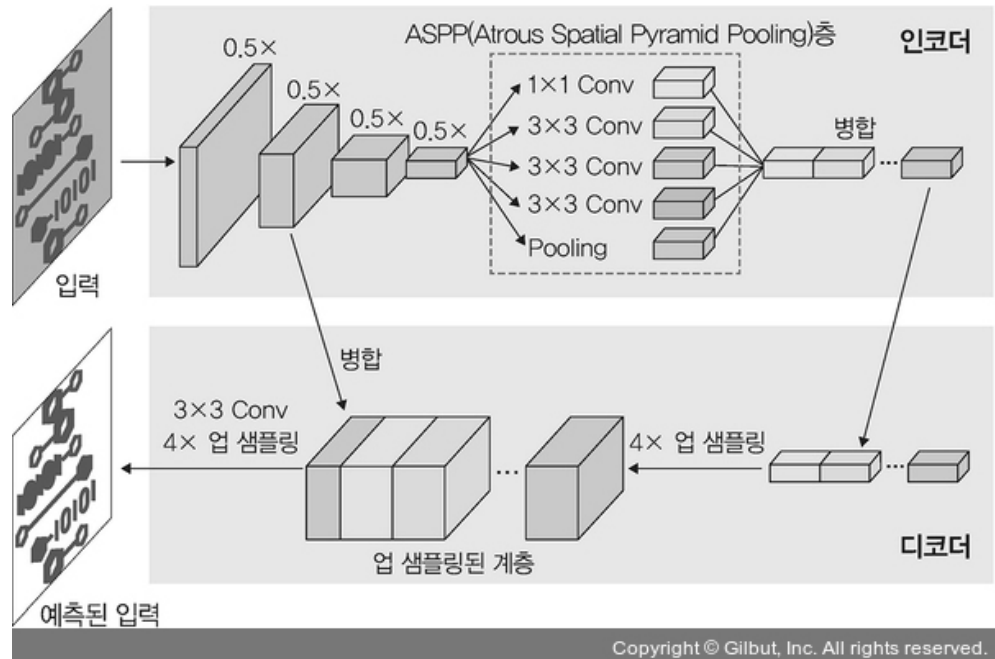
▼ 6.3.4 PSPNet

- 완전 연결층의 한계를 극복하기 위해 **피라미드 풀링 모듈**을 추가
1. 이미지 출력이 서로 다른 크기가 되도록 여러 차례 풀링. 즉, 1×1, 2×2, 3×3, 6×6 크기로 풀링을 수행하는데, 이때 1×1 크기의 특성 맵은 가장 광범위한 정보를 담음
 2. 이후 1×1 합성곱을 사용하여 채널 수를 조정. 풀링층 개수를 N이라고 할 때 출력 채널 수=입력 채널 수/N
 3. 이후 모듈의 입력 크기에 맞게 특성 맵을 업 샘플링. 이 과정에서 양선형 보간법(bilinear interpolation)이 사용.
 4. 원래의 특성 맵과 1~3 과정에서 생성한 새로운 특성 맵들을 병합.

▼ 6.3.5 DeepLabv3/DeepLabv3+

- 완전 연결층의 한계를 극복하기 위해 **Atrous 합성곱** 사용
- 인코더, 디코더 구조를 가지며 해상도를 **Atrous 합성곱**을 도입하여 제어
 - **Atrous 합성곱**은 필터내부에 빈 공간을 둔 채로 작동
 - 얼마나 빈 데이터를 결정하는 파라미터 : rate
 - r이 커질수록 빈공간이 많아짐

○



- 보통 이미지 분할에서 높은 성능을 내려면 수용 영역(외부자극이 전체에 영향을 주는 것이 아니라 특정 영역에만 영향을 준다는 의미)의 크기가 중요한데, 수용 영역을 확대하여 특성을 찾는 범위를 넓게 해줌. 이 합성곱을 활용하면 파라미터 수를 늘리지 않으면서도 수용영역을 크게 키울 수 있기 때문에 이미지 분할 분야에서 많이 사용

