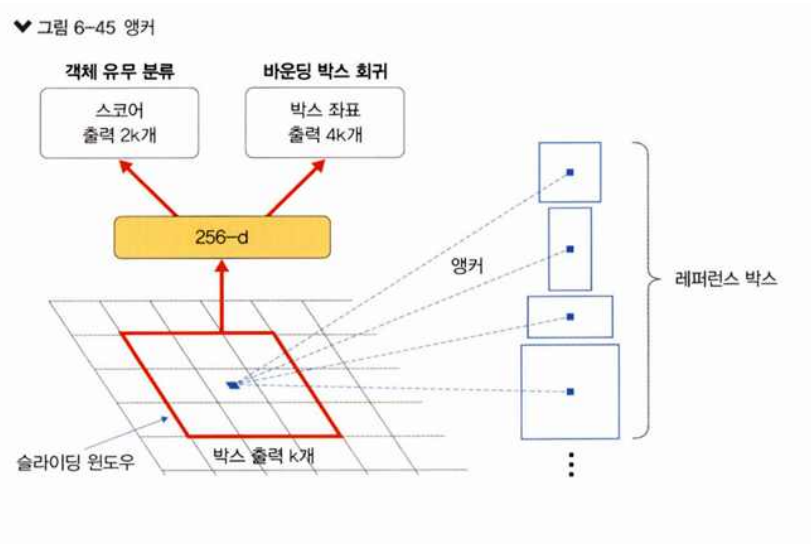


- 후보 영역 추출 네트워크는 feature map $n \times n$ 크기의 작은 윈도우 영역을 입력으로 받고 해당 영역의 객체 존재 유무 판단을 위해 binary classification을 수행하는 작은 네트워크를 생성함
- 하나의 feature map에 대해 모든 영역에 대한 객체 존재 여부를 판단하기 위해 슬라이딩 윈도우 방식으로 탐색함
- 이미지에 존재하는 객체의 크기와 비율이 다양하다는 점을 고려하여 여러 크기와 비율의 reference box를 k개 미리 정의하고 각각의 슬라이딩 윈도우 위치마다 k개를 출력하도록 설계함 → aka anchor 방식



▼ 6장 Part 2

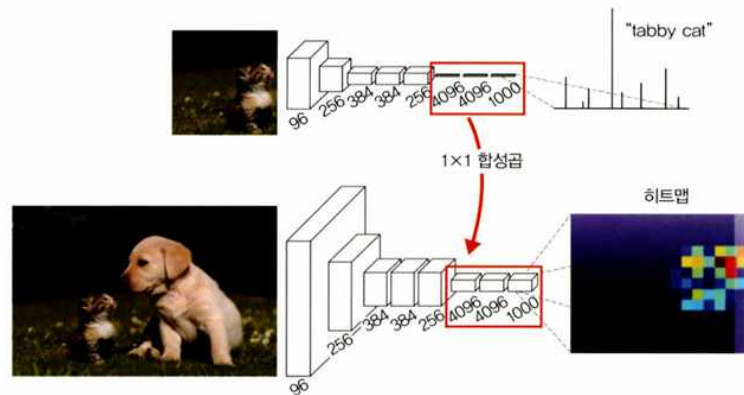
6.3 이미지 분류를 위한 신경망

- 이미지 분할은 신경망을 훈련시켜 이미지를 픽셀단위로 분할하여 이미지에 포함된 개체를 추출

6.3.1 완전 합성곱 네트워크 Fully Convolutional Network

- 완전연결층의 한계는 고정된 크기의 입력만 받아들이고 완전연결층을 거친 후에는 (flatten을 사용하므로) 위치 정보가 사라진다는 것
- FCN은 이미지 분류에서 우수한 성능을 보인 CNN 기반 모델을 변형시켜 만든 이미지 분할 맞춤 네트워크

▼ 그림 6-46 완전 합성곱 네트워크



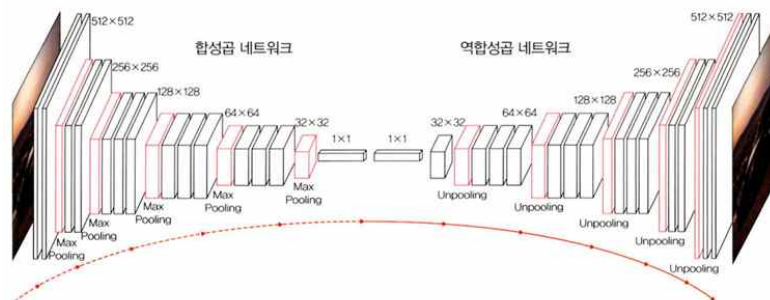
AlexNet에서의 FC 3layers를 1x1 CNN으로 변환하여 위치 정보 저장

- 기존의 fc layer 대신 합성곱층이 사용되기 때문에 입력이미지에 대한 크기 제약이 사라지는 장점

6.3.2 합성곱&역합성곱 네트워크

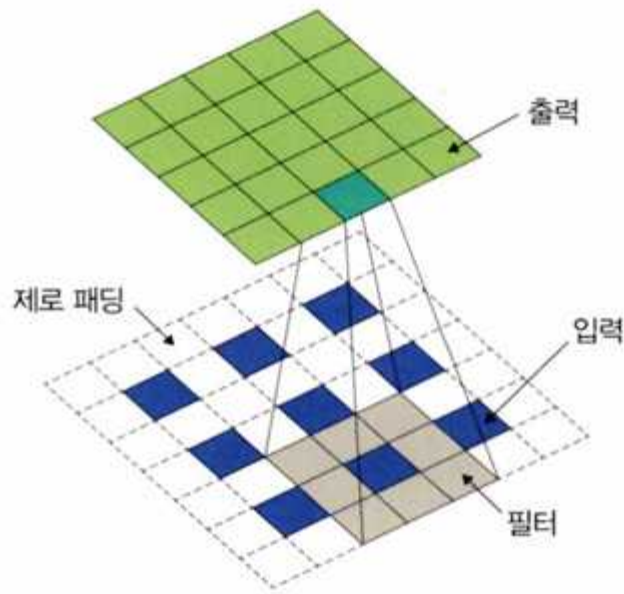
- FCN의 단점: 1) 여러 단계의 CNN과 Pooling층으로 인해 해상도가 낮아짐, 2) 이를 위해 upsampling방식을 사용하기 때문에 이미지의 세부 정보를 잃어버리는 문제 → 복잡한 객체의 세밀한 경계나 작은 객체는 업샘플링 과정에서 완벽하게 복원되지 않기 때문
- 역합성곱을 통해 CNN 최종 결과를 원래의 이미지 입력과 같은 크기로 만들 → upsampling

▼ 그림 6-47 합성곱 & 역합성곱 네트워크



- 역합성곱의 작동 방식: 1) 각 픽셀 주위에 zero padding, 2) 이 결과에 CNN 연산 수행

▼ 그림 6-48 역합성곱 진행 방식

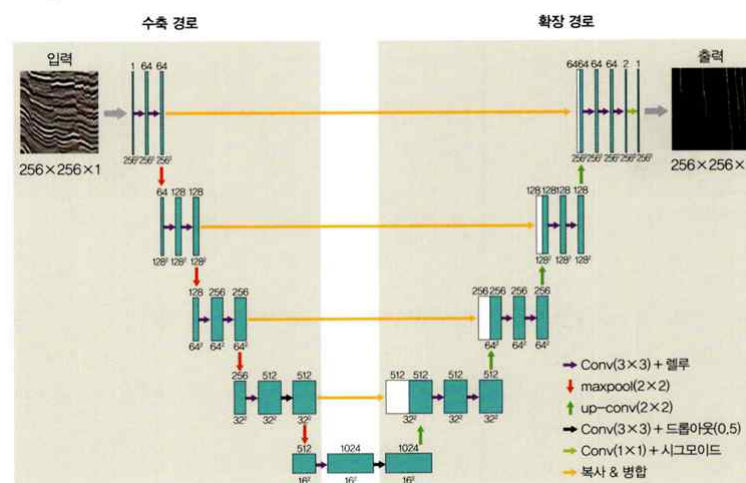


아래의 CNN 결과에 제로 패딩을 한후 CNN을 거치면 위와 같은 결과가 나옴

6.3.3 U-Net

- 바이오 메디컬 분야에서 이미지 분할을 위해 주로 사용됨
- 이전 패치에서 검증이 끝난 부분을 재사용하는 기존 슬라이딩 윈도우 방식과 달리 이미 검증이 끝난 패치는 건너뛰기 때문에 속도가 빠름
- 패치의 크기가 크다면 context능력에는 탁월하나 localization(bounding box를 사용해서 이미지 안에 객체의 위치 정보를 알려주는 것)에 한계가 있음

▼ 그림 6-51 U-Net



- FCN기반 수축경로로 컨텍스트 포착, 확장 경로로 featuremap을 upsampling하여 localization 수행

- 각 합성곱 블록은 2개의 3x3 합성곱과 dropout으로 구성되어 있음. 블록은 수축경로에서 4개가 있음. 각 블록은 maxpool로 크기를 줄여 다음 블록으로 넘어감
- 확장 경로에는 up-conv로 줄어든 크기를 복원하며 합성곱 블록을 이용
- 따라서 downsampling과 upsampling을 통해 다양한 featuremap을 병합할 수 있어 localization과 context 능력 둘다 증가(trade-off에 빠지지 않음)

6.3.4 Pyramid Scene Parsing Network PSPNet

- 위와 같은 FCN의 단점(해상도가 낮아지고 복원이 힘들)을 극복하기 위해 피라미드 풀링 모듈을 추가

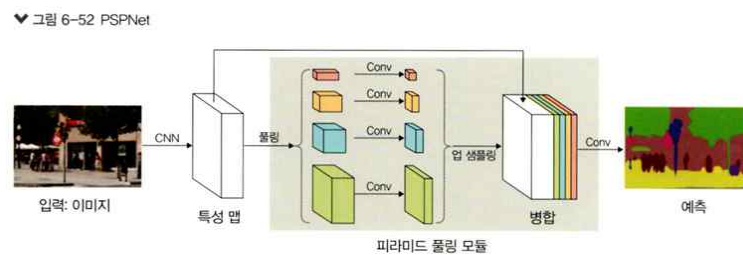


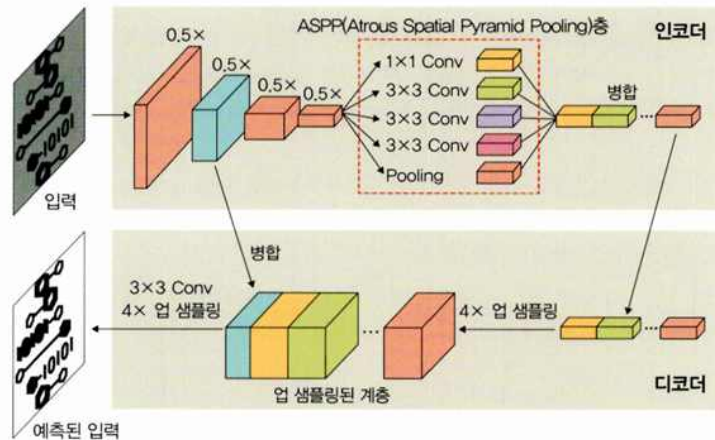
그림 6-52는 풀링을 네 개 사용했지만, 구현에 따라서 다르게 설정할 수 있습니다.

- 1x1 ~ 6x6 풀링을 통해 서로 다른 크기의 출력을 생성. (1x1 featuremap은 가장 context적. 광범위한 정보를 담음)
- 1x1 CNN을 통해 채널수(입력채널수/풀링층개수)를 조정
- bilinear interpolation을 통해 featuremap의 upsampling 진행. 이 featuremap을 기존의 featuremap과 병합

6.3.5 DeepLabv3/DeepLabv3+

- FCN의 단점을 보완하기 위해 Atrous CNN을 사용

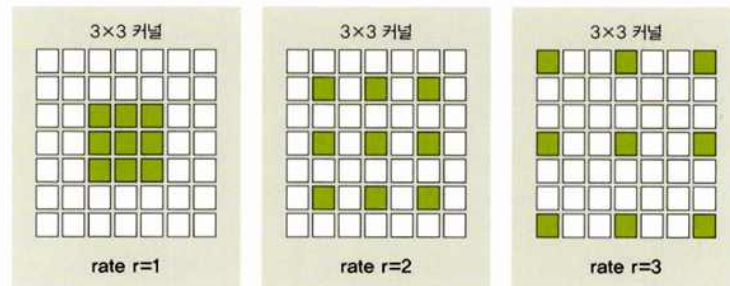
▼ 그림 6-54 DeepLab의 인코더-디코더 구조



인코더-디코더 구조 중 인코더에 Atrous 합성곱을 도입

- Atrous 합성곱이란?

▼ 그림 6-55 Atrous 합성곱



이때 rate: 얼마나 많은 빈공간을 가질 지 결정하는 파라미터

- 이미지 분할에서 중요한 \의 여역을 확대하여 특성을 찾는 범위를 넓게 만듦
- 일반적인 CNN의 경우 출력은 입력의 1/32배로 줄어들지만 Atrous 합성곱을 적용하면 featuremap이 기존보다 4배 보존되어 출력이 입력의 1/8배가 됨