

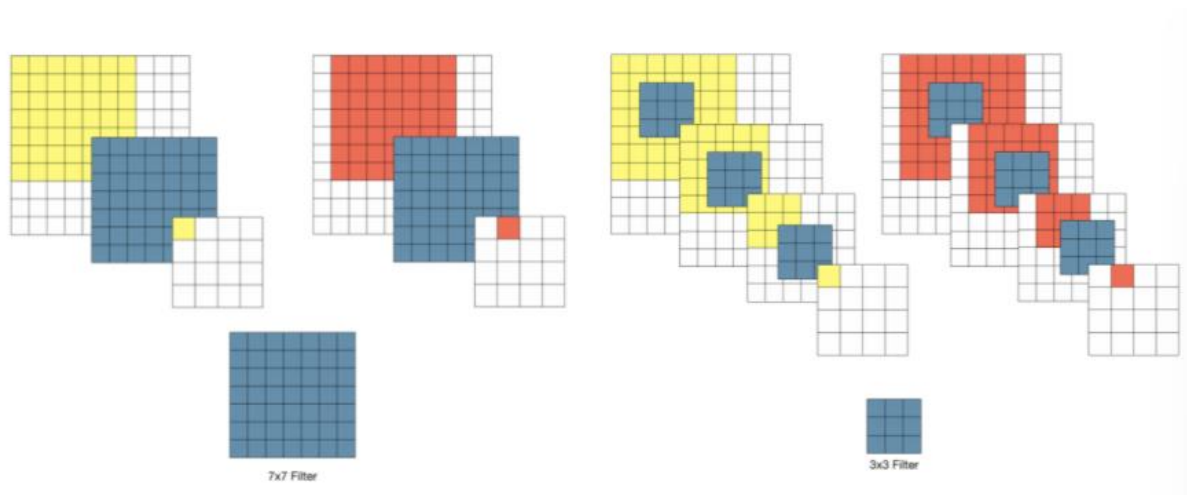
VGGNet Summary

옥스포드 대학의 연구팀 VGG에 의해 개발된 모델로써 2014년 이미지넷 인식 대회에서 준우승을 한 모델이다.

키워드 – Depth, Small convolution filters

VGGNet이 모든 합성곱 레이어에서 3X3 필터를 사용하여 16-19 레이의 깊은 신경망을 학습하는 과정이 이 논문의 핵심이다.

Q. 왜 Conv 필터의 사이즈가 3x3으로 고정되었을까?



1. 결정함수의 비선형성이 증가한다.

각 Convolution 연산에는 ReLU 함수를 포함한다.

7X7 필터는 비선형 함수를 1번 적용하지만 ,3X3 필터는 3번 적용한다.

따라서, 레이어가 증가함에 따라 비선형성이 증가하게 되고 이것은 모델의 특징 식별성 증가로 이어진다.

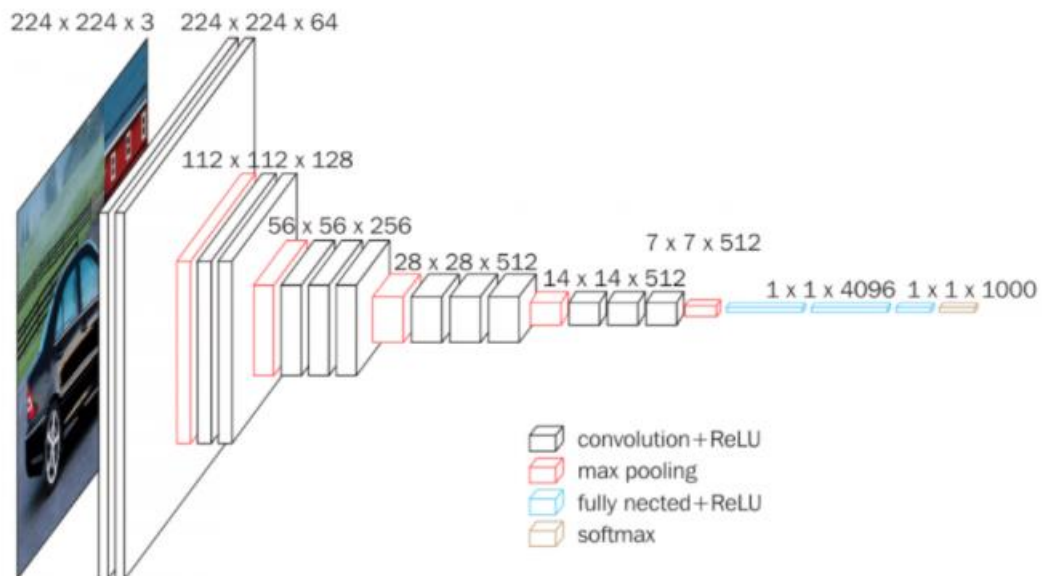
2. 학습 파라미터 수가 줄어든다

10X10 이미지에 7X7 필터 1번 -> 학습 파라미터 수 49

10X10 이미지에 3X3 필터 3번 적용 -> 학습 파라미터 수 27

파라미터의 수가 줄어들수록 정규화할 때 이점을 얻을 수 있다.

-학습



Batch Size: 256

Momenum: 0.9

L2 Regularization: 5×10^{-4}

Learning rate: 10^{-2}

Input Image: 224 x 224 x 3

-학습 이미지 크기



학습 데이터를 다양한 크기로 변환하고 샘플링하였다

- ➔ 데이터의 수를 늘릴 수 있다.
- ➔ 하나의 이미지를 다양한 측면으로 학습할 수 있다.

-실행

더 깊은 Depth로 학습 시 성능 향상이 없을 것이라고 추측할 수 있다.

-결론

장점

- > 작은 Filter Size Conv의 중첩으로 좋은 효과를 낼 수 있다.
- > 기존 AlexNet 대비 2배 이상 깊은 네트워크이다.

단점

- ➔ 마지막 FC Layers에 많은 파라미터가 몰려있다.