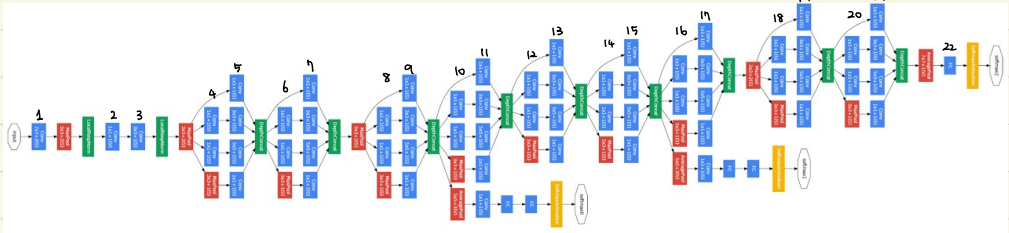


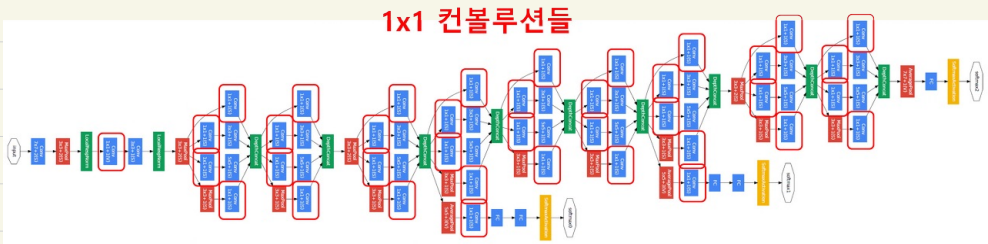
GoogleNet

22층 계층



파라미터 22

1) 1x1 Convolution



특성 맵의 갯수를 줄이는 목적 \Rightarrow 연산량 \downarrow

ex) 480장의 14×14 size 특징맵 ($14 \times 14 \times 480$)

\downarrow

48개의 $5 \times 5 \times 480$ 의 필터 커널을 Convolution \Rightarrow 16장의 14×14 의 특징맵 ($14 \times 14 \times 48$) 생성 (Zero padding: 2, Convolution 변 · 1)

\therefore 연산량 : $(14 \times 14 \times 48) \times (5 \times 5 \times 480) = 112 \text{ G}$

480장의 14×14 size 특징맵 ($14 \times 14 \times 480$)

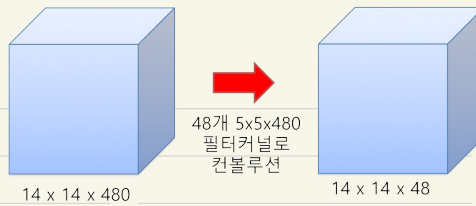
\downarrow

16개의 $1 \times 1 \times 480$ 의 필터 커널 Convolution \Rightarrow 16장의 14×14 의 특징맵 ($14 \times 14 \times 16$) 생성

\downarrow

48개의 $5 \times 5 \times 16$ 의 필터 커널 Convolution \Rightarrow 48장의 14×14 의 특징맵 ($14 \times 14 \times 48$) 생성

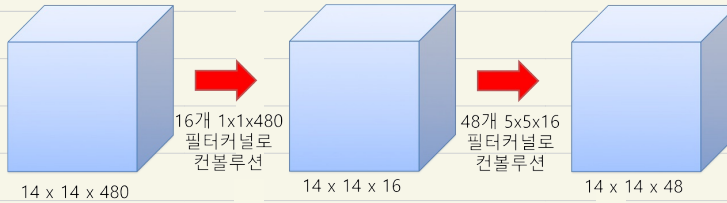
\therefore 연산량 : $(14 \times 14 \times 16) * (1 \times 1 \times 480) + (14 \times 14 \times 48) * (5 \times 5 \times 16) = 5.3 \text{ M}$



112.9M

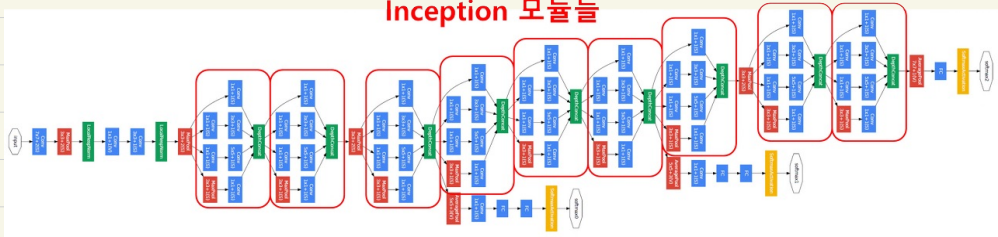


5.3M

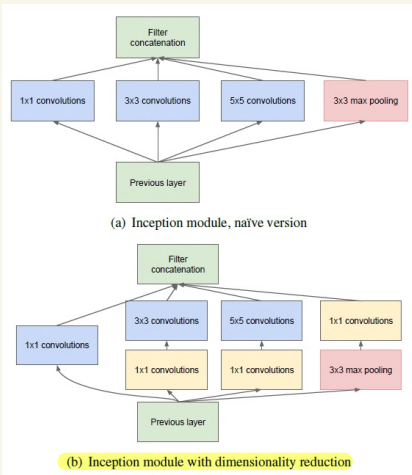


2) Inception 모듈

Inception 모듈들



종 여러
모듈



다양한 feature를 뽑기 위해 여러 종류의 convolution 커널을 병렬로 사용

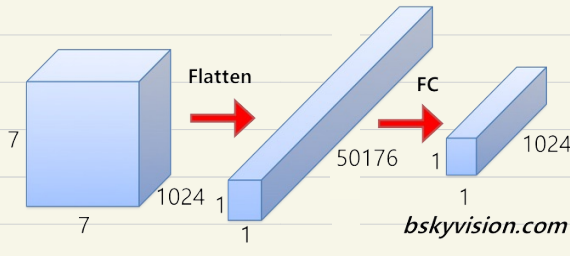
input feature에서 의미있는 feature를 뽑아내기 위해서는 다양한 representation을 받아들일 수 있는 필터들이 필요하기 때문입니다.

3) global average pooling

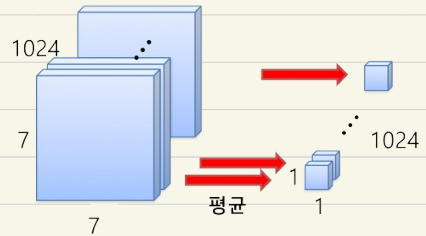
전 층에서 산출된 특성맵들을 각각 평균낸 것을 이어서 1차원 벡터를 만들어주는 것이다.

1차원 벡터를 만들어줘야 최종적으로 이미지 분류를 위한 softmax 층을 연결해줄 수 있기 때문이다.

FC 방식



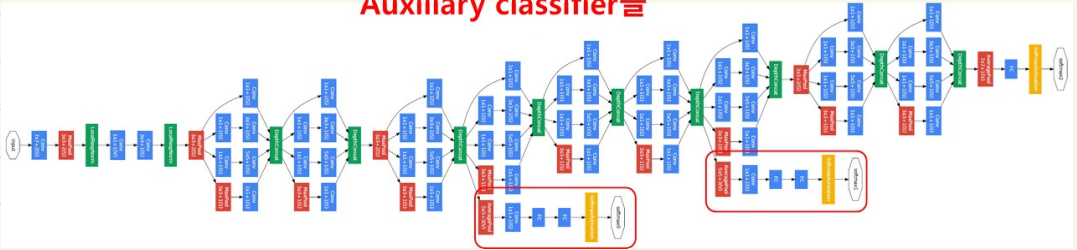
Global average pooling



⇒ 가중치의 갯수를 상당히 많이 없애줌

4) auxiliary classifier

Auxiliary classifier들



네트워크의 깊이가 깊어지면 깊어질수록 vanishing gradient 문제를 피하기 어려워진다. 그러니까 가중치를 훈련하는 과정에 역전파(back propagation)를 주로 활용하는데, 역전파과정에서 가중치를 업데이트하는데 사용되는 gradient가 점점 작아져서 0이 되어버리는 것이다. 따라서 네트워크 내의 가중치들이 제대로 훈련되지 않는다. 이 문제를 극복하기 위해서 GoogLeNet에서는 네트워크 중간에 두 개의 보조 분류기(auxiliary classifier)를 달아주었다.

훈련 시에만 활용되고 사용할 때는 제거

