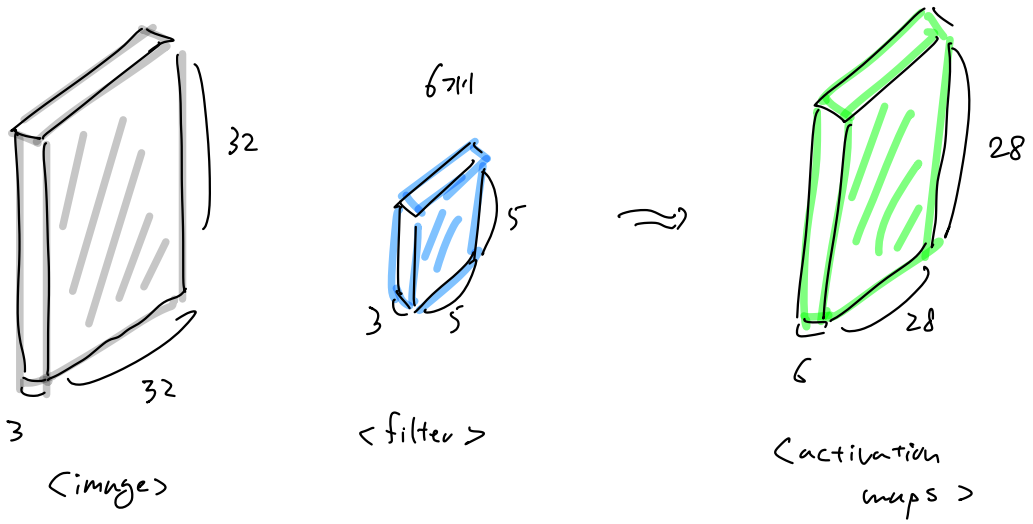
< Convolution Layer >



<image> <filter> <activation maps>

Naive Conv

+ { padding, stride }

input : $N \times N \times C$

filter : $K \times K \times C$ × D개

⇒ output $N' \times N' \times C'$

$O(N'-1) = N - K + 2P$

$C' = D$

필터가 $(N'-1)$번을 S만큼 이동한 거리

= 필터가 움직일 수 있는 공간의 크기

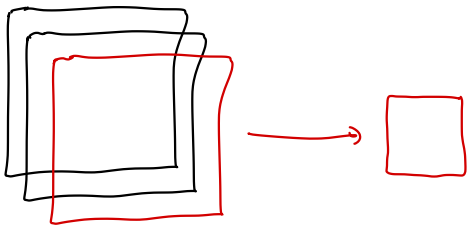
1×1 Conv Layers

$56 \times 56 \times 64$

$1 \times 1 \times 64$ 32개

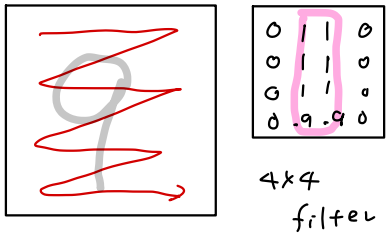
⇒ $56 \times 56 \times 32$

Pooling layer



- max sampling
- average sampling

핵심: filter

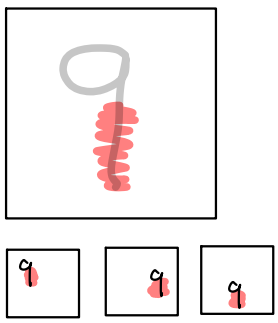


4×4 filter

"element-wise multiplications & average"

∀ locations

feature bitmap



translation invariance (이동 불변성)

Convolution은 translation에 대해 equivariance를 가진다.

$f \circ g(x) = g \circ f(x)$

↑ Conv ↑ trans ↑ trans ↑ Conv

AlexNet

227×227×3	Input
55×55×96	Conv 1 (kernel = 11×11, stride = 4, pad = 0) + ReLU
27×27×96	Max Pool 1 (kernel = 3×3, stride = 2)
27×27×96	Norm 1
27×27×256	Conv 2 (kernel = 5×5, stride = 1, pad = 2) + ReLU
13×13×256	Max Pool 2 (kernel = 3×3, stride = 2)
13×13×256	Norm 2
13×13×384	Conv 3 (kernel = 3×3, stride = 1, pad = 1) + ReLU
13×13×384	Conv 4 (kernel = 3×3, stride = 1, pad = 1) + ReLU
13×13×256	Conv 5 (kernel = 3×3, stride = 1, pad = 1) + ReLU
6×6×256	Max Pool 3 (kernel = 3×3, stride = 2)
[4096]	FC 6
[4096]	FC 7
[1000]	FC 8

① ReLU ⇒ saturating gradient ⇓

② first stride ⇑ ⇒ feature size & computation ⇓

③ Normalization is not used in recent CNNs

④ FC dropout 50% ⇒ overfitting ⇓

cf. dropout

신경망의 일부를 무작위로 사용하지 않음

- co-adaptation ⇓ "특정 feature가 뛰어나다."
- ensemble ⇑

ZFNet

- Conv1 (~~11×11~~ 7×7 stride ~~4~~ 2)

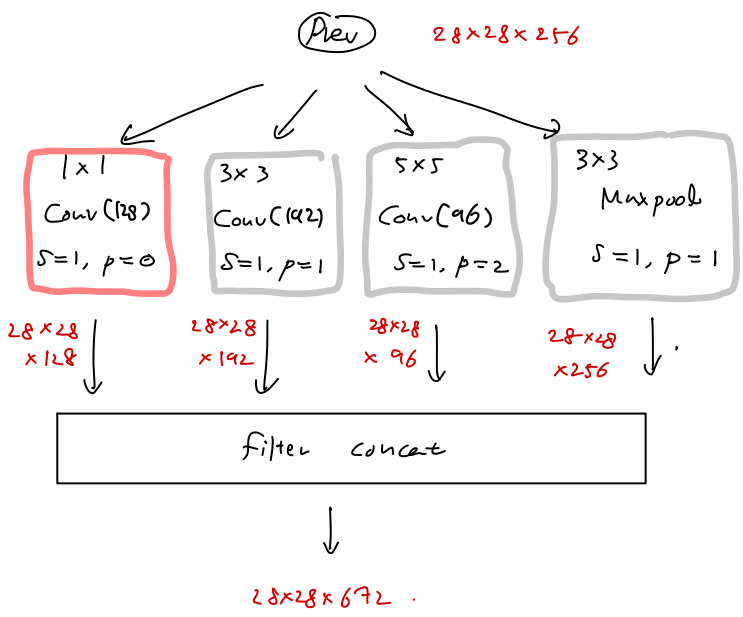
- Conv3, 4, 5 : ~~384, 384, 256~~ 512, 1024, 512 filters

GoogLeNet

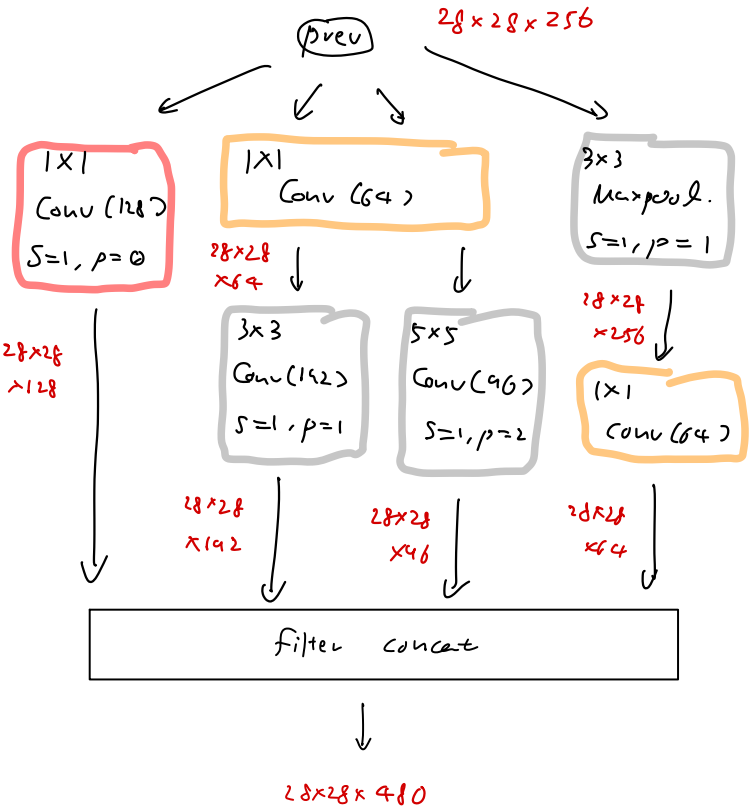
< Inception Modules >

parameter size (filter)

<original>



<improved>



① Stem Network (초기 Conv + Pooling)

② Inception Module Stacle

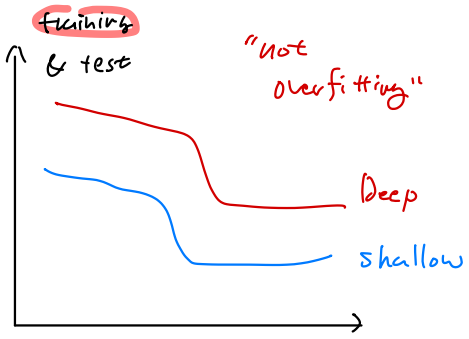
③ Classifier

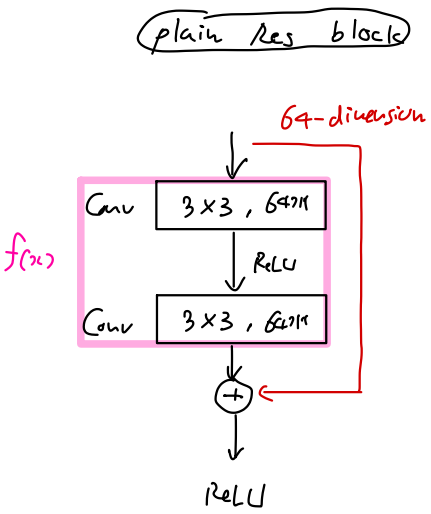
(i) 7×7×1024 —(average pooling)→ 1×1×1024

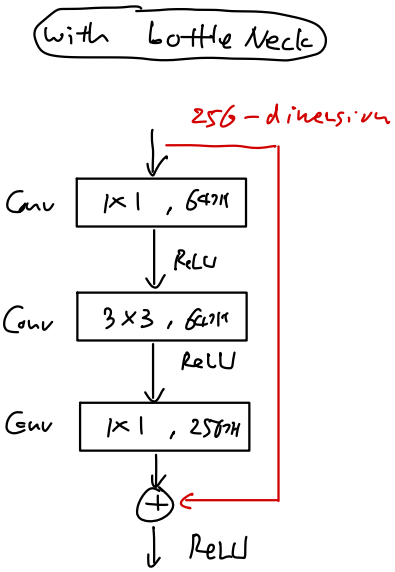
(ii) Auxiliary Classifiers : 중간의 노드 분류기에 그래디언트 주입

Deeper Network?

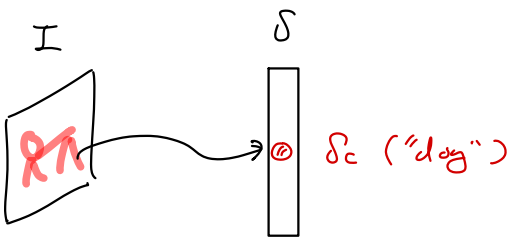
< Residual Connection >







"bypass layers by setting $f(x) = 0$"



Saliency

$$\frac{\partial S_c}{\partial I} = \frac{\partial F^{(L)}}{\partial F^{(L-1)}} \cdot \frac{\partial F^{(L-1)}}{\partial F^{(L-2)}} \cdots \frac{\partial F^{(1)}}{\partial I}$$
