



Convolutional Neural Network

GoogLeNet

2022.02.21

김혜현

GoogLeNet

: Going Deeper with convolutions

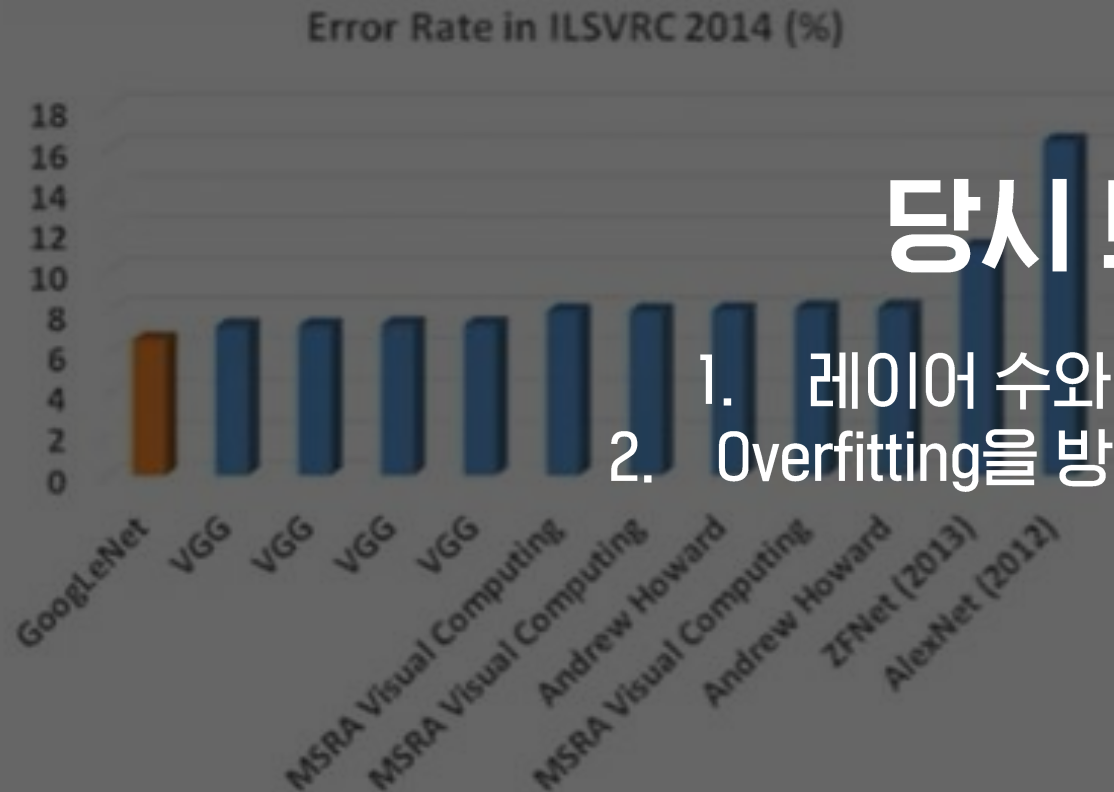
ILSVRC 2014 _ 1등



당시 트렌드?

Team	Year	Place	Error (top-5)	Uses external data
SuperVision	2012	1st	16.4%	no
SuperVision	2012	1st	15.3%	Imagenet 22k
Clarifai	2013	1st	11.7%	no
Clarifai	2013	1st	11.2%	Imagenet 22k
MSRA	2014	3rd	7.35%	no
VGG	2014	2nd	7.32%	no
GoogLeNet	2014	1st	6.67%	no

ILSVRC 2014 _ 1등



당시 트렌드?

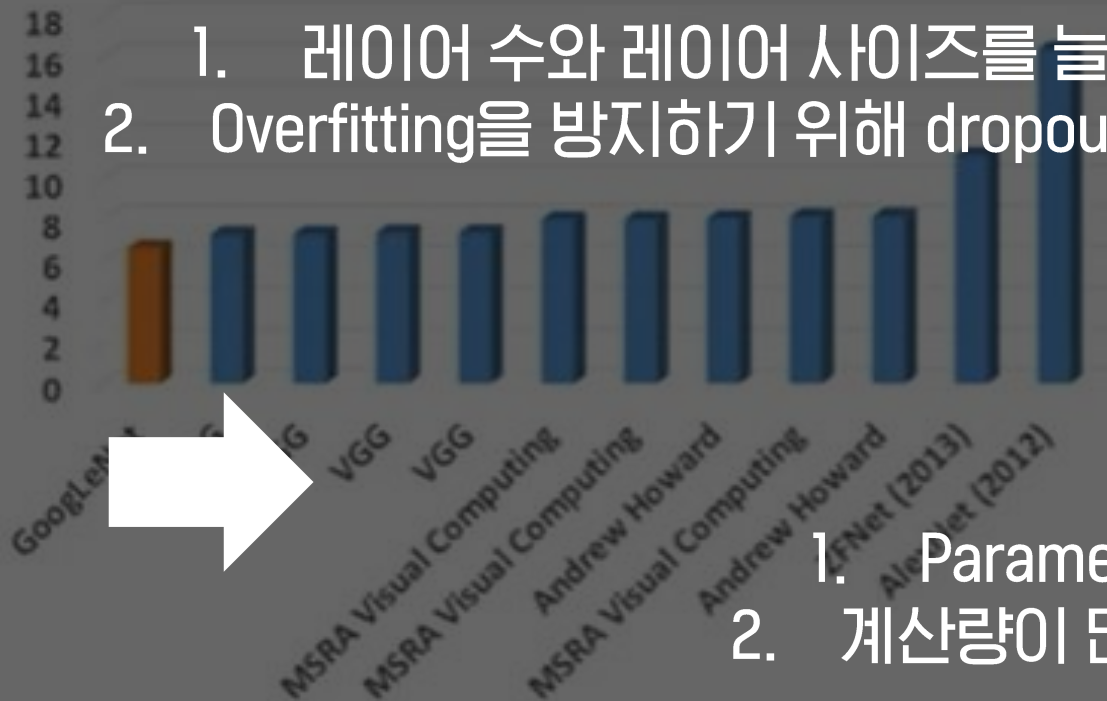
1. 레이어 수와 레이어 사이즈를 늘림
2. Overfitting을 방지하기 위해 dropout 사용

Team	Year	Place	Error (top-5)	Uses external data
SuperVision	2012	1st	16.4%	no
SuperVision	2012	1st	15.3%	Imagenet 22k
Clarifai	2013	1st	11.7%	no
Clarifai	2013	1st	11.2%	Imagenet 22k
MSRA	2014	3rd	7.35%	no
VGG	2014	2nd	7.32%	no
GoogLeNet	2014	1st	6.67%	no

ILSVRC 2014 _ 1등

당시 트렌드?

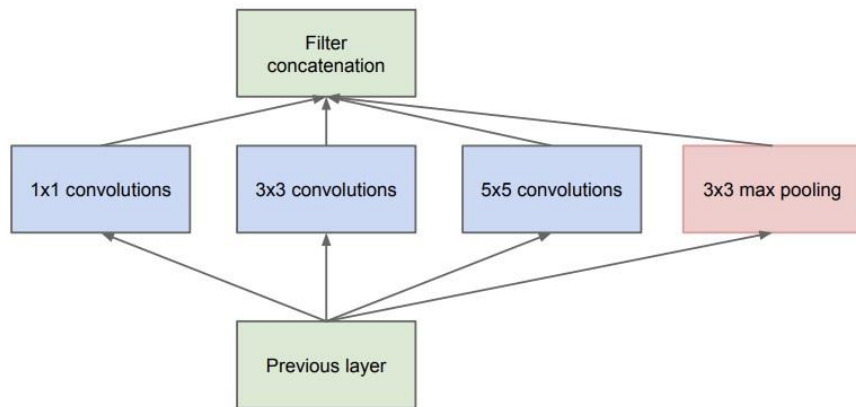
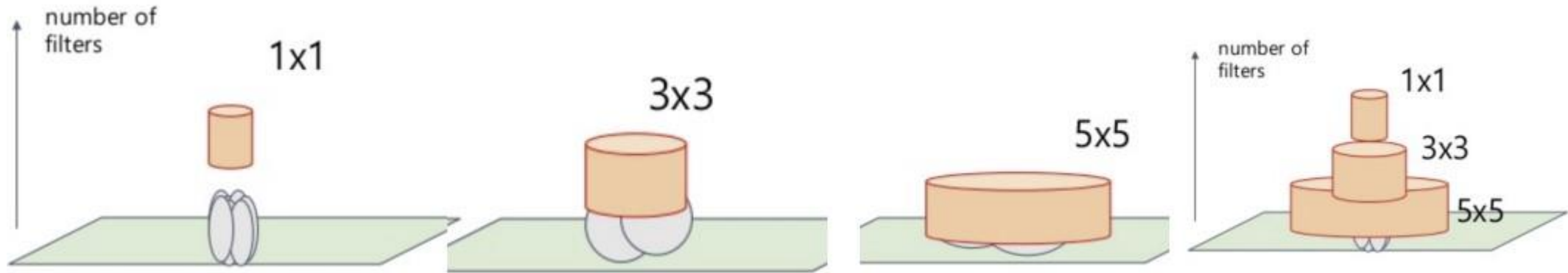
1. 레이어 수와 레이어 사이즈를 늘림
2. Overfitting을 방지하기 위해 dropout 사용



문제점

1. Parameter가 많으면 오버피팅이 되기 쉬움
2. 계산량이 많아지지만, 계산 자원은 제한되어 있음

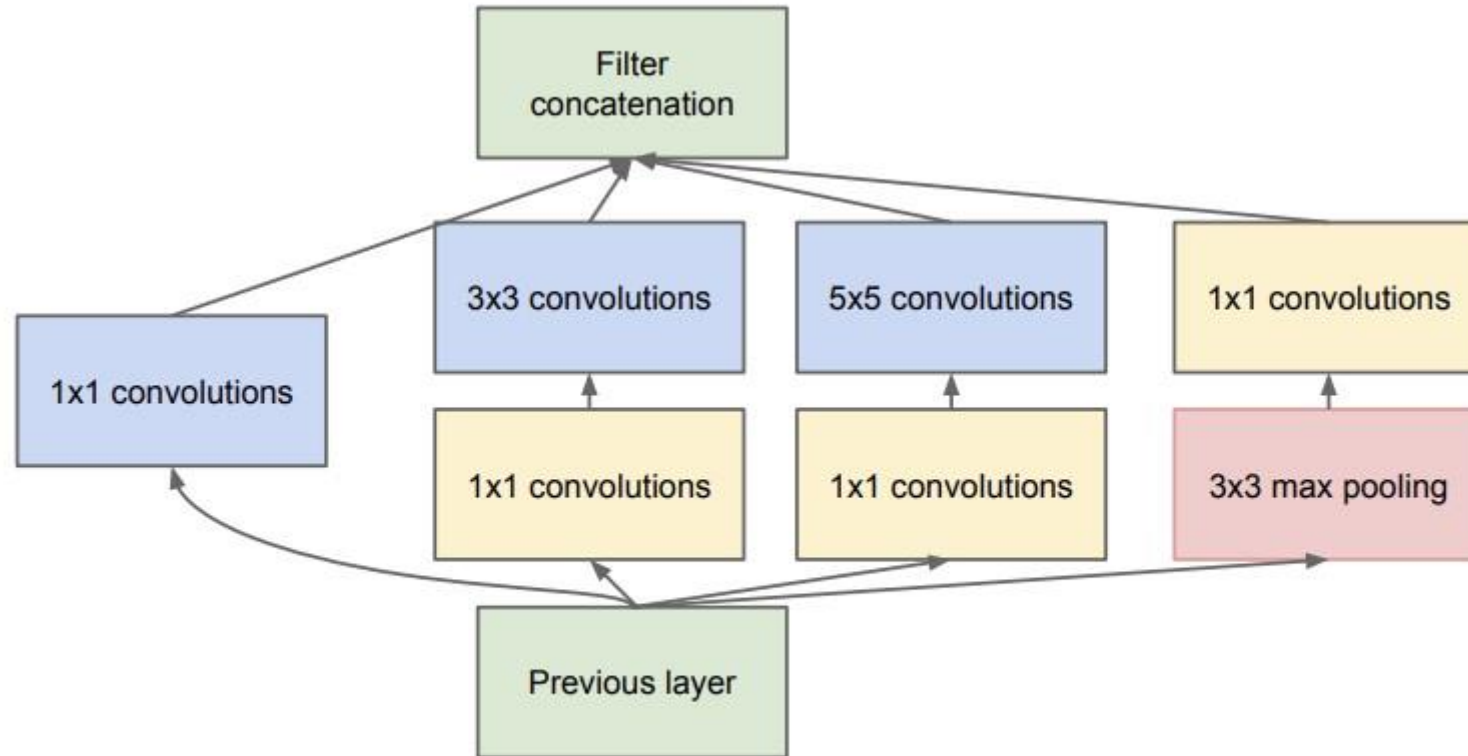
Inception module



(a) Inception module, naïve version

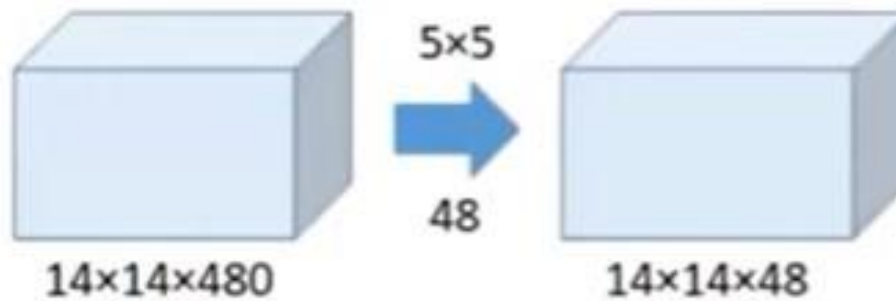
- ✓ Feature Map을 효과적으로 추출하기 위해 1 x 1, 3 x 3, 5 x 5의 convolution 연산을 각각 수행

Inception module



(b) Inception module with dimension reductions

1 x 1 convolution



of operations :
 $(14 * 14 * 48) * (5 * 5 * 480) = 112.9\text{M}$

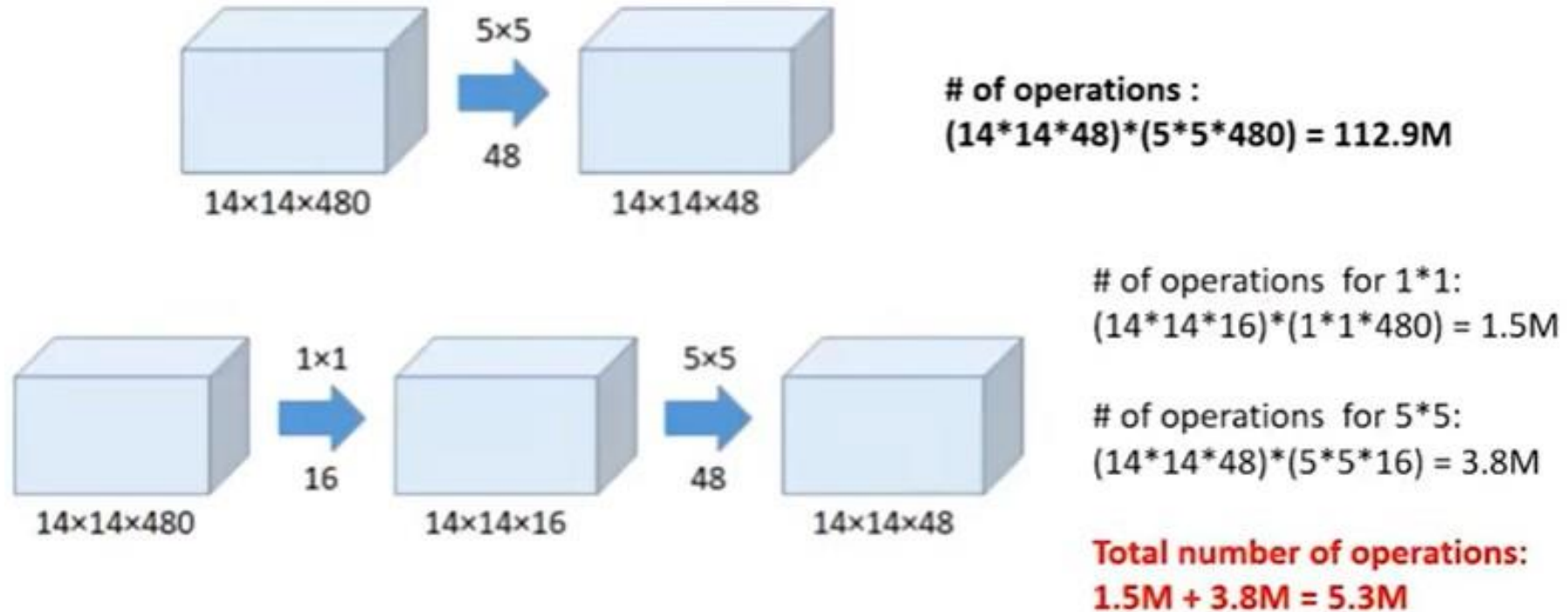


of operations for 1×1 :
 $(14 * 14 * 16) * (1 * 1 * 480) = 1.5\text{M}$

of operations for 5×5 :
 $(14 * 14 * 48) * (5 * 5 * 16) = 3.8\text{M}$

Total number of operations:
 $1.5\text{M} + 3.8\text{M} = 5.3\text{M}$

1 x 1 convolution

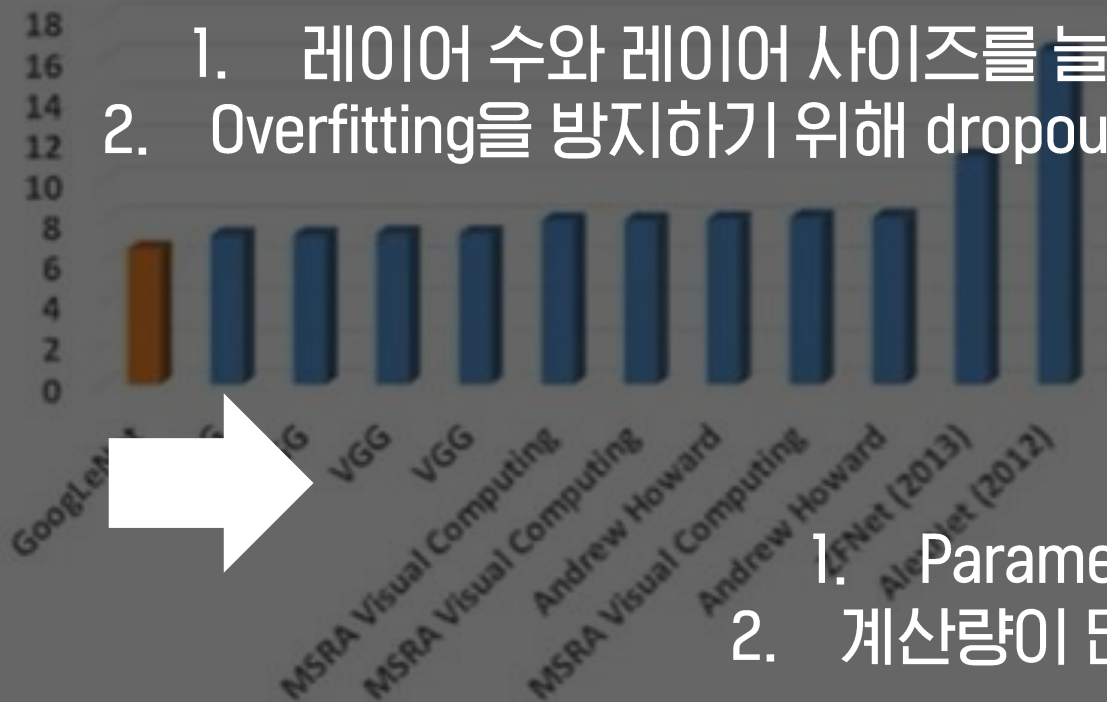


- ✓ 계산을 줄이기 위해 차원 축소 모듈로 사용됨
- ✓ 계산량을 줄임으로써 깊이와 너비를 늘릴 수 있음

ILSVRC 2014 _ 1등

당시 트렌드?

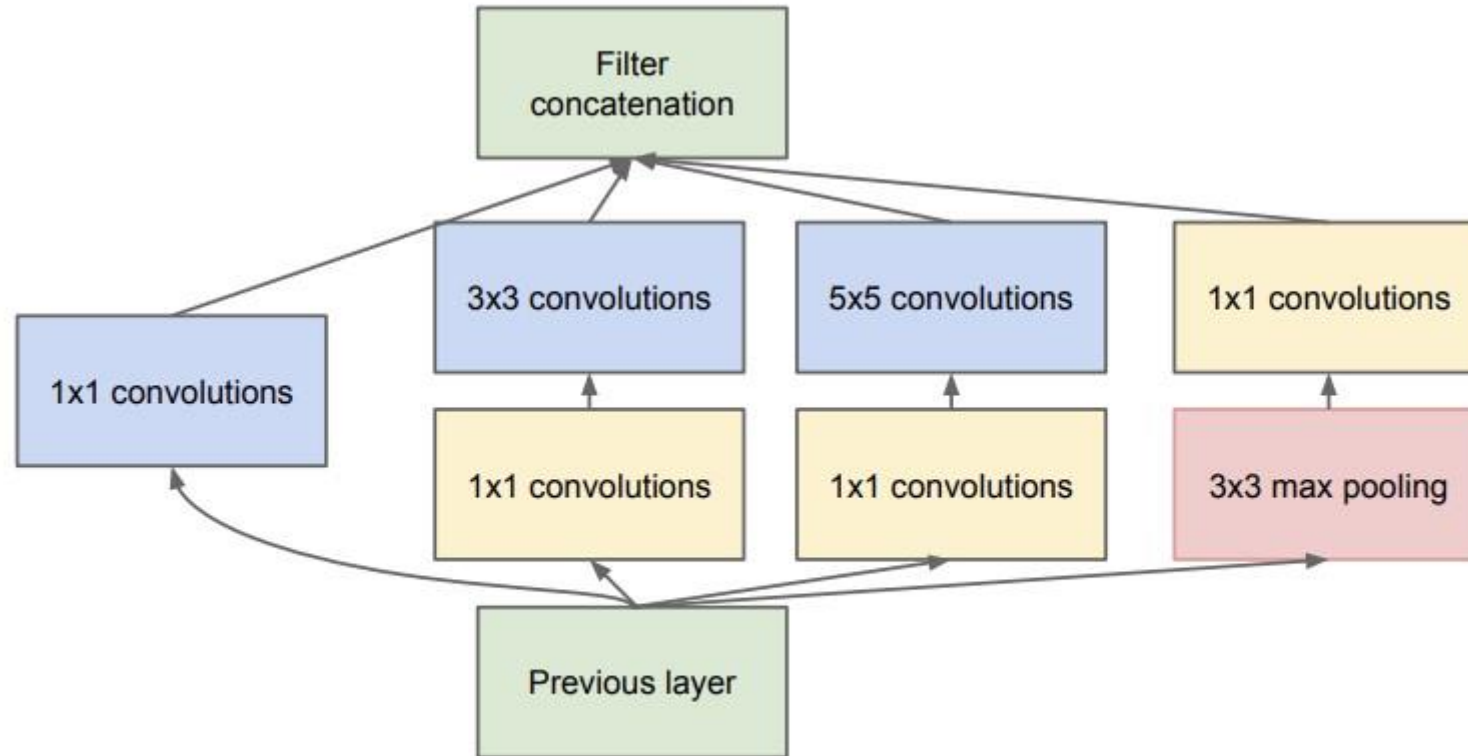
1. 레이어 수와 레이어 사이즈를 늘림
2. Overfitting을 방지하기 위해 dropout 사용



문제점

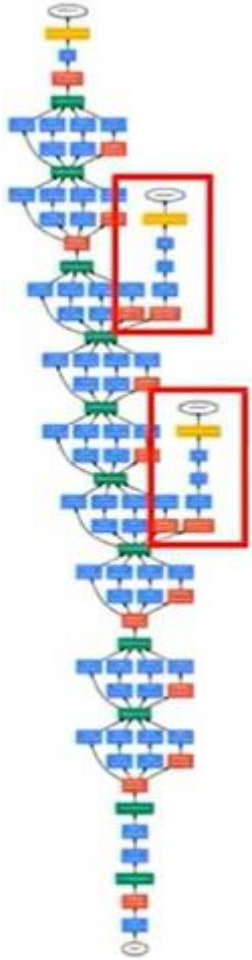
1. Parameter가 많으면 오버피팅이 되기 쉬움
2. 계산량이 많아지지만, 계산 자원은 제한되어 있음

Inception module



(b) Inception module with dimension reductions

Auxiliary Classifier

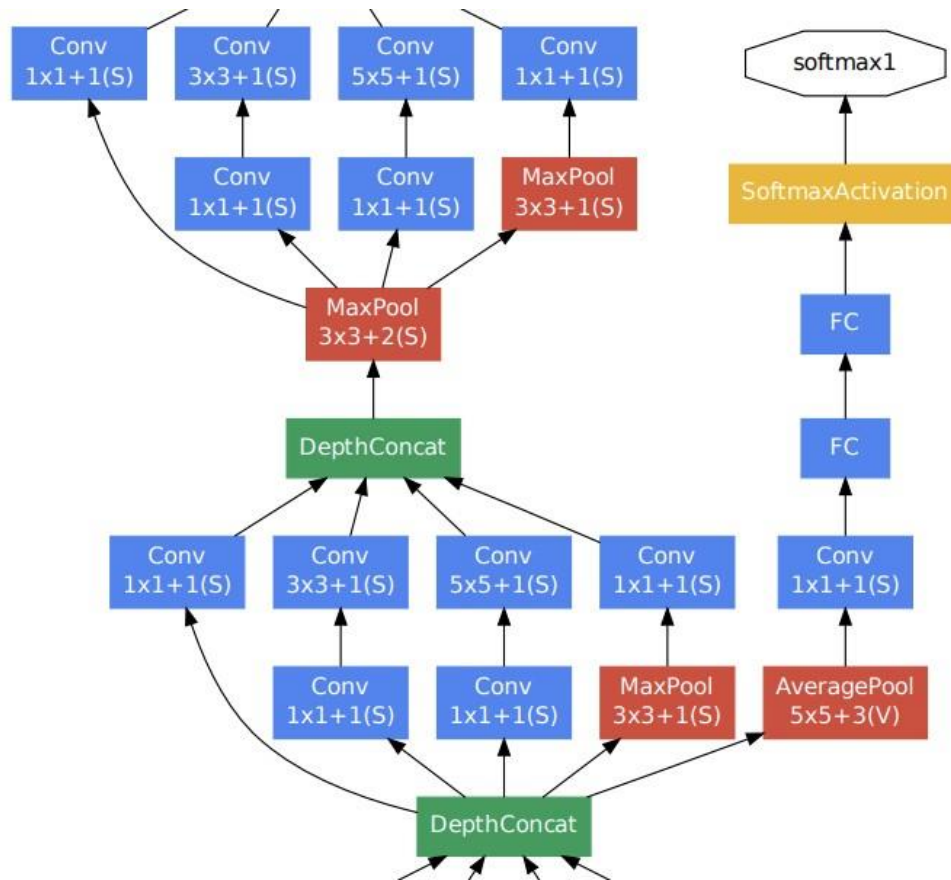


Auxiliary Classifier(보조 분류기)란?

- 중간 단계에서 예측하는 부분
- 깊은 네트워크에서 학습이 잘 안될 수 있는 점을 이를 통해 추가적인 Gradient를 줌으로써 보완

$$\text{total_loss} = \text{real_loss} + 0.3 * \text{aux_loss_1} + 0.3 * \text{aux_loss_2}$$

Auxiliary Classifier



-> Linear layer에 softmax를 사용한 1000-class classifier

-> Dropout layer (0.7)

-> FC layer(1024 nodes) 및 ReLU

-> Dimension reduction을 위한 1 x 1 conv layer(128 filters) 및 ReLU

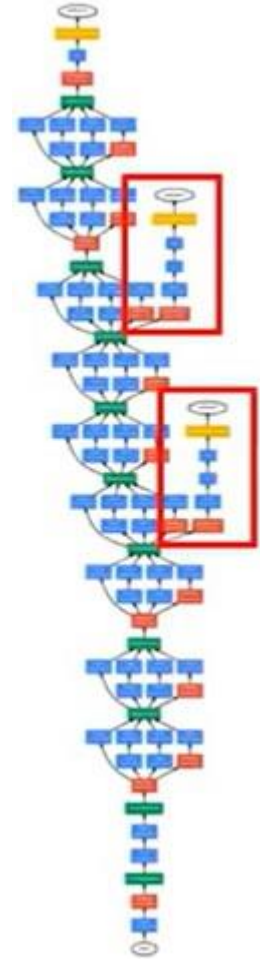
-> Filter size가 5 x 5이고, strides가 3인 average pooling layer.
출력의 shape은 [4a)와 [4d)에서 각각 4 x 4 x 512와 4 x 4 x 5280이다.

GoogLeNet

✓ 성능을 높이기 위해 여러 기법 사용

1. 동일한 GoogLeNet 모델의 7가지 버전(wider version도 하나 포함)을 독립적으로 학습했으며, 이들을 이용한 ensemble prediction을 수행했다.

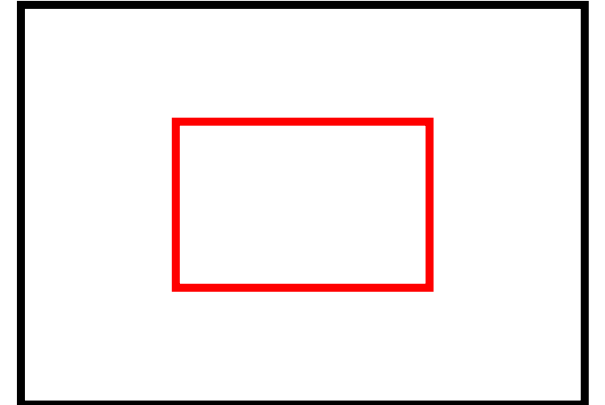
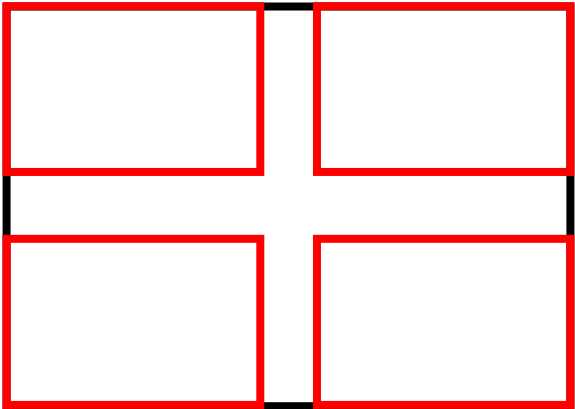
각 모델들은 동일한 weight initialization과 learning rate policy로 학습했으며, sampling 방법과 shuffle로 인한 학습 데이터의 순서에서만 차이가 있다.



✓ 성능을 높이기 위해 여러 기법 사용

2. Cropping 방식 적용

Shorter size가 각각 [256 / 288 / 320 / 352]인
4가지 scale로 이미지의 크기를 조정
[left / center / right] 의 Square를 취한다.
각 square에 대해 [모서리 4개 / 중앙] 에서
224 x 224 크기로 resize한 것과,
이들의 미러링 된 버전을 취한다.



결론



이미지 당 총 $4 \times 3 \times 6 \times 2 = 144$ 개의 crop이 생성

GoogLeNet

✓ 성능을 높이기 위해 여러 기법 사용

Number of models	Number of Crops	Cost	Top-5 error	compared to base
1	1	1	10.07%	base
1	10	10	9.15%	-0.92%
1	144	144	7.89%	-2.18%
7	1	7	8.09%	-1.98%
7	10	70	7.62%	-2.45%
7	144	1008	6.67%	-3.45%

결론

- ✓ 계산량이 조금 증가하는 것에 비해, 성능이 아주 좋아짐
- ✓ AlexNet보다 12배 적은 parameter를 사용하면서, 더 좋은 성능을

BN-Inception

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_1 \dots x_m\}$;

Parameters to be learned: γ, β

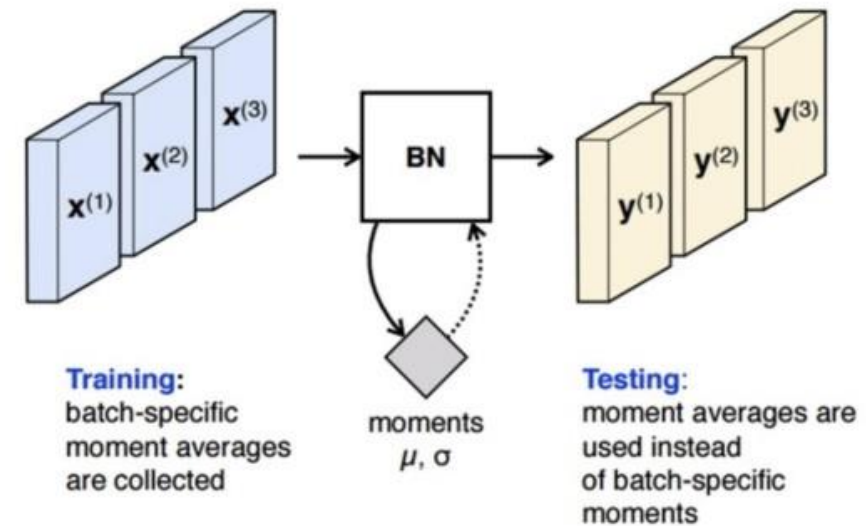
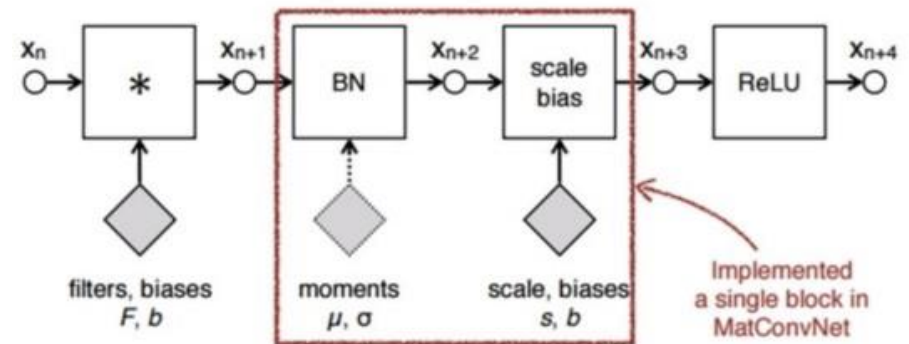
Output: $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad // \text{ mini-batch mean}$$

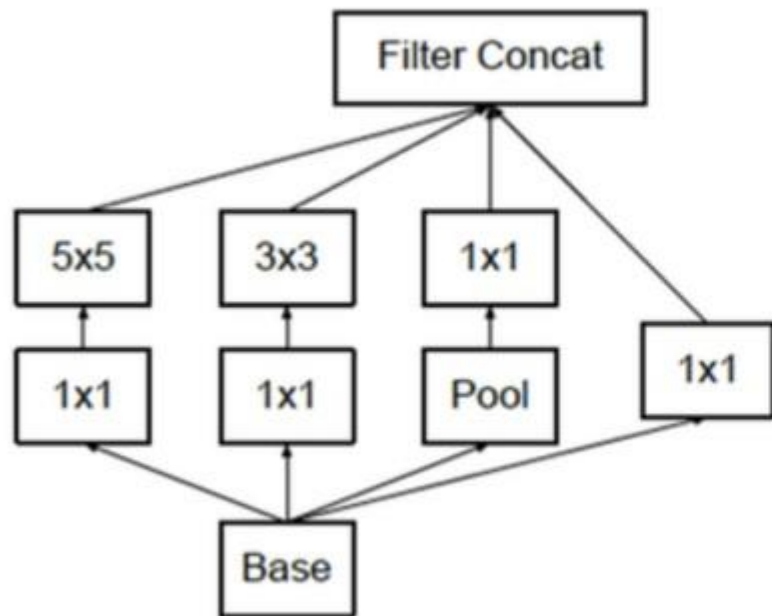
$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\mathcal{B}})^2 \quad // \text{ mini-batch variance}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \quad // \text{ normalize}$$

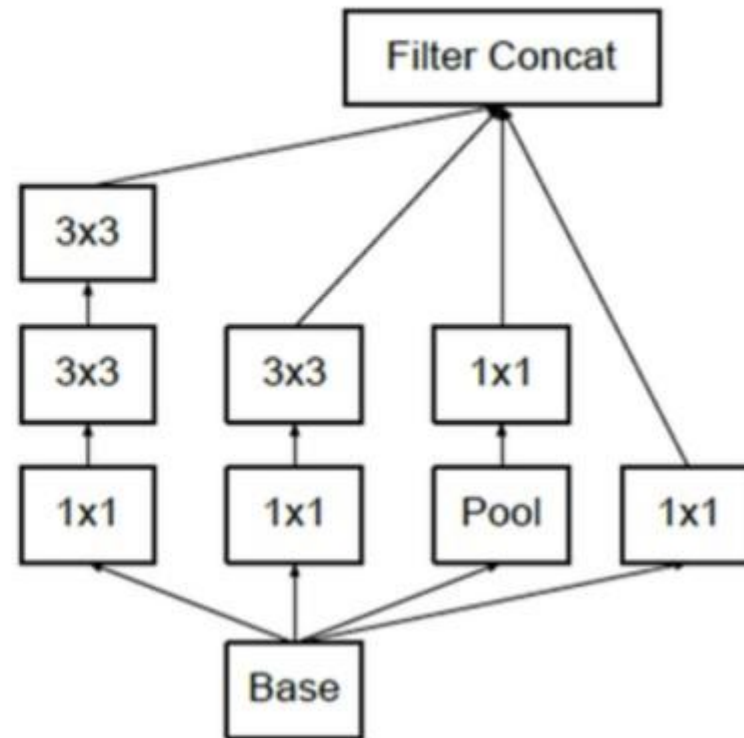
$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i) \quad // \text{ scale and shift}$$



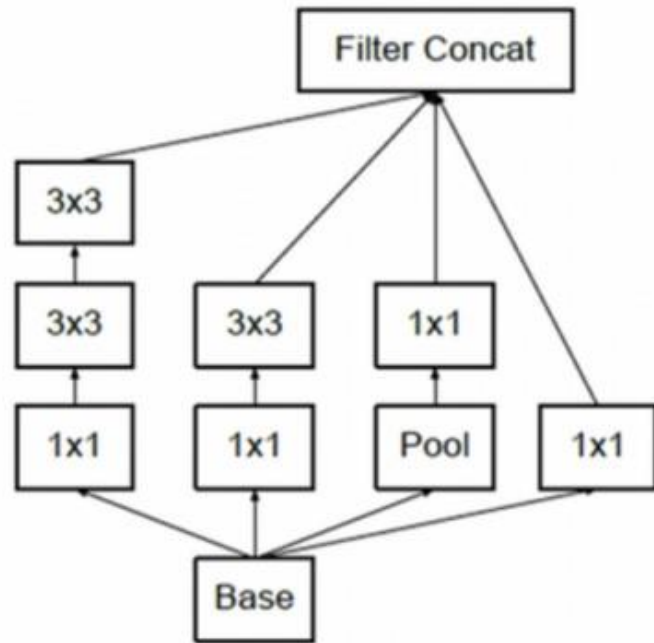
Inception v2



Inception Module A

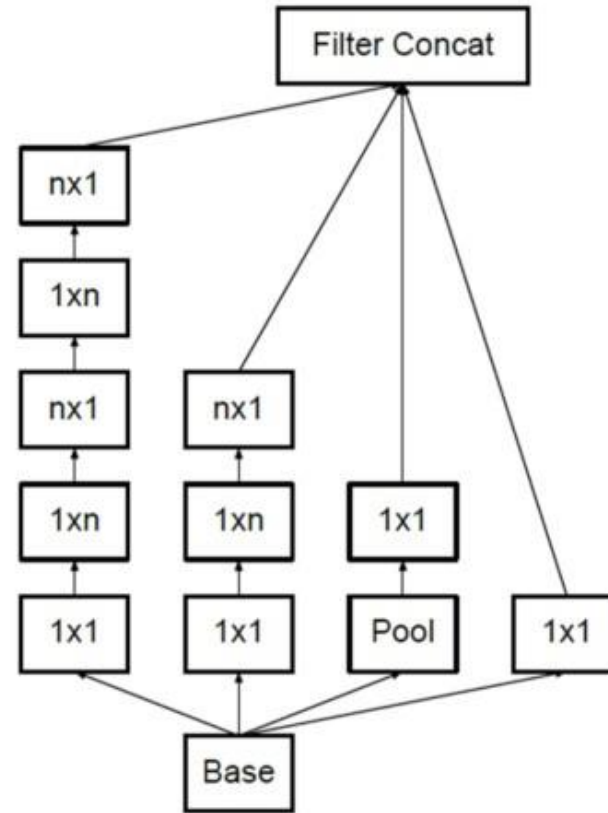


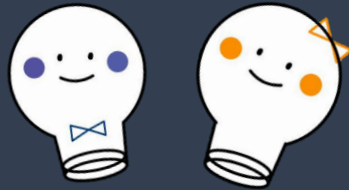
Inception v3



연산량 33% 절감

Inception Module B





우리만 따라와 *Follow ADS*

이상 ADS 김혜현이었습니다. 감사합니다.