D&A

Deep Session 5차시

CNN 심화.

2022 / 04 / 07

D&A 운영진 윤경서

CONTENTS.

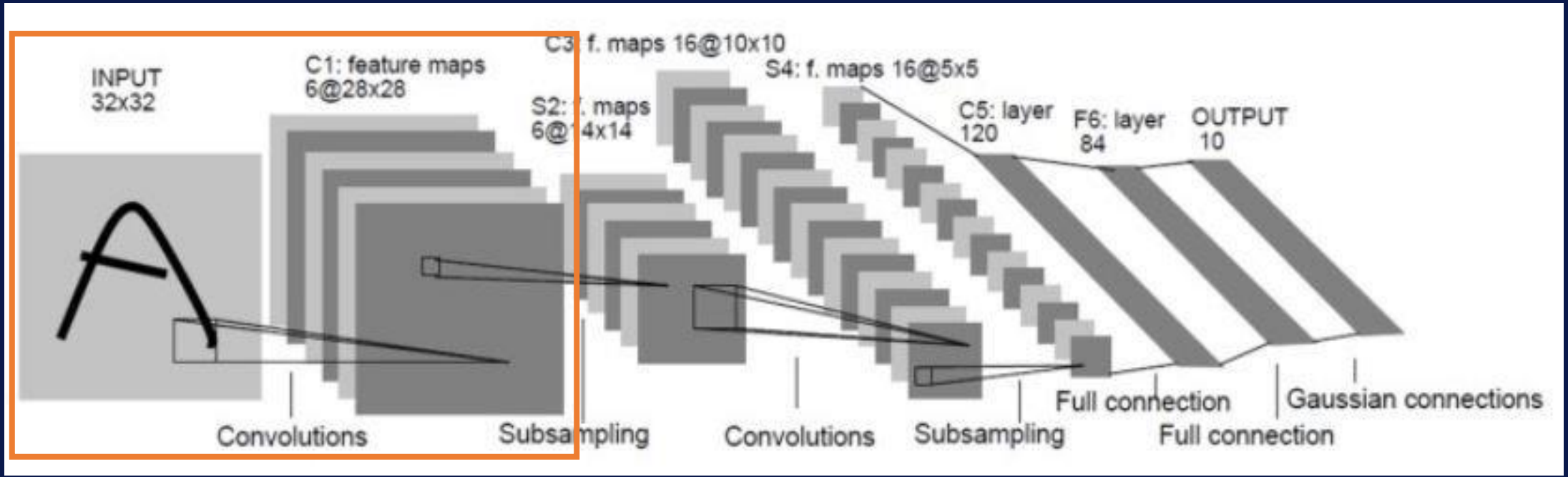
01. LeNet

- 1998년 Yann Lecun 연구팀이 개발한 CNN 알고리즘이다.
- Yann Lecun 팀의 논문 ‘ Gradient-Based Learning Applied to Document Recognition’에 수록되어 있는 LeNet-5가 대표적인 모델이다.

- LeNet-5는 32x32 크기의 흑백 이미지에서 학습된 7 layer CNN이다.
- [Input – Conv(C1) – Subsampling(S2) – Conv(C3) – Subsampling(S4) – Conv(C5) – FC6 – FC7(output)]

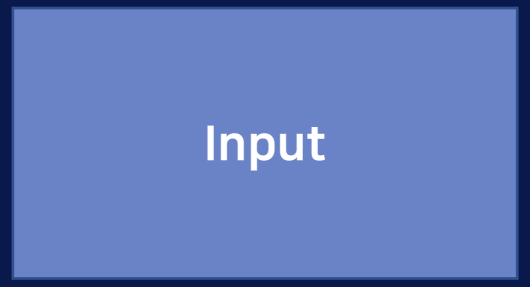
01. Lenet

LeNet Architecture



01. Lenet

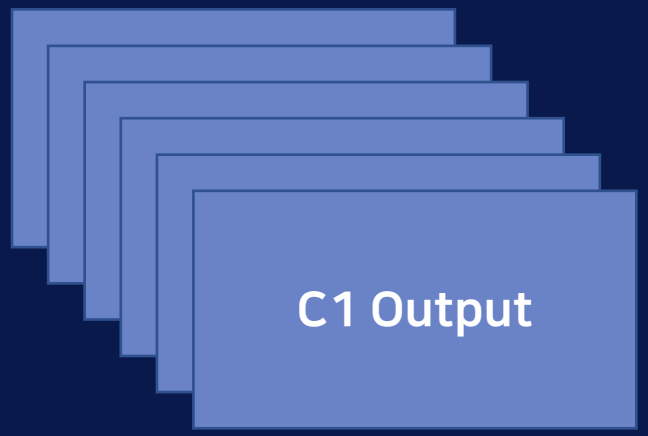
Input



Input size : (1, 32, 32)

Filter size : 5x5
Filter 개수 : 6
Stride : 1

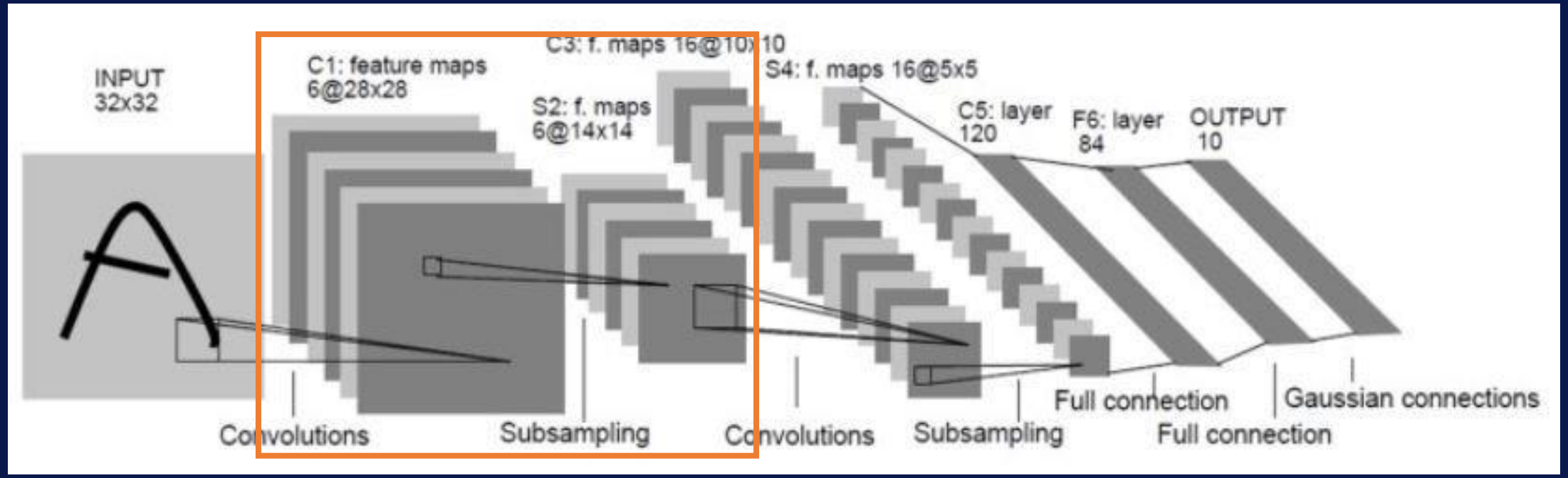
C1 layer



C1 Output size : (6, 28, 28)

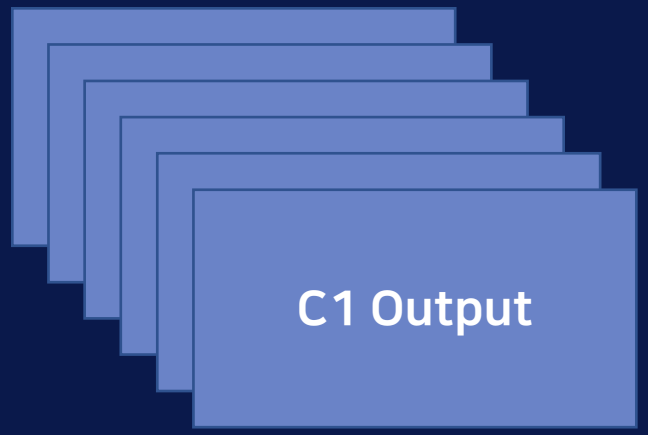
01. Lenet

LeNet Architecture



01. Lenet

C1 layer



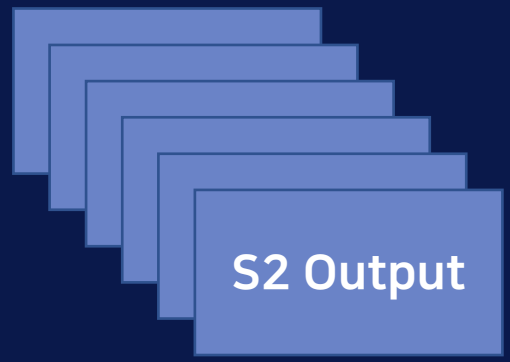
C1 Output size : (6, 28, 28)

Subsampling

Average Pooling

Pooling size : 2x2

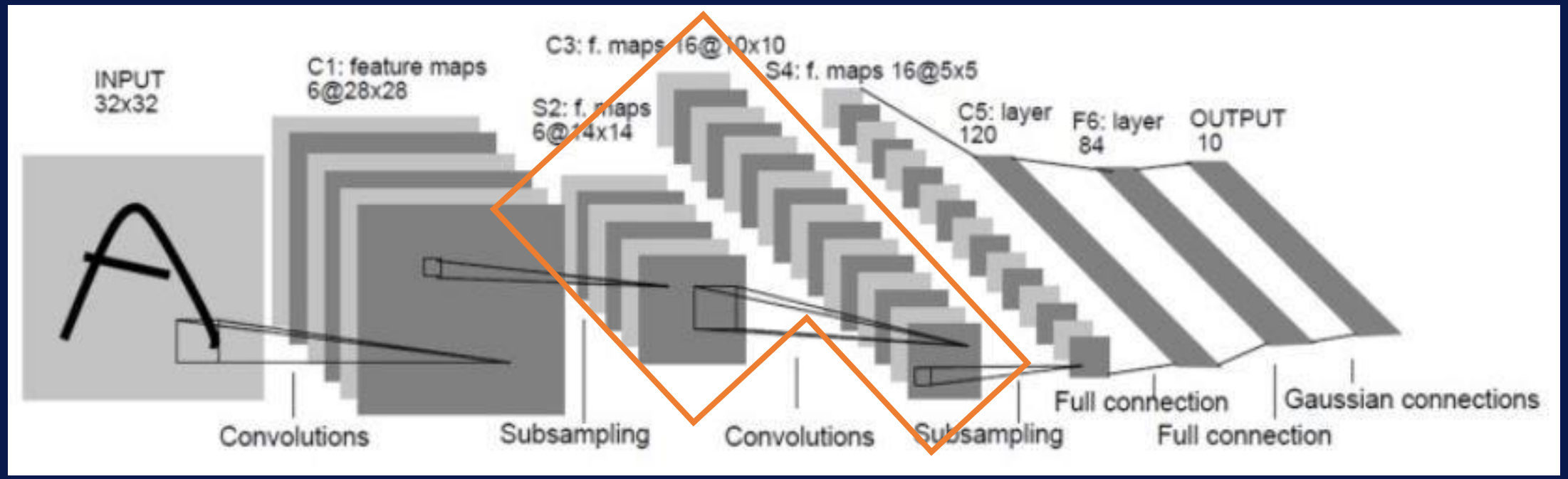
S2 layer



S2 Output size : (6, 14, 14)

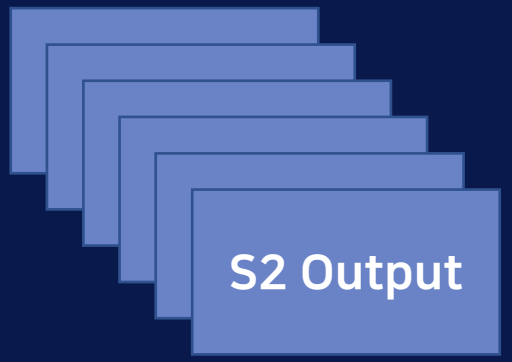
01. Lenet

LeNet Architecture



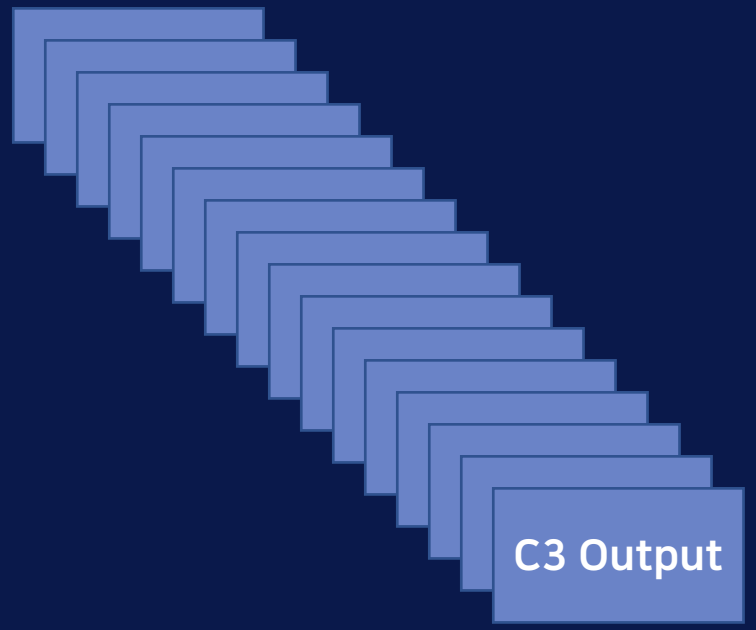
01. Lenet

S2 layer



Convolution

C3 layer



S2 Output size : (6, 14, 14)

Filter size : 5x5
Filter 개수 : 16
Stride : 1

C3 Output size : (16, 10, 10)

01. Lenet

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

TABLE I

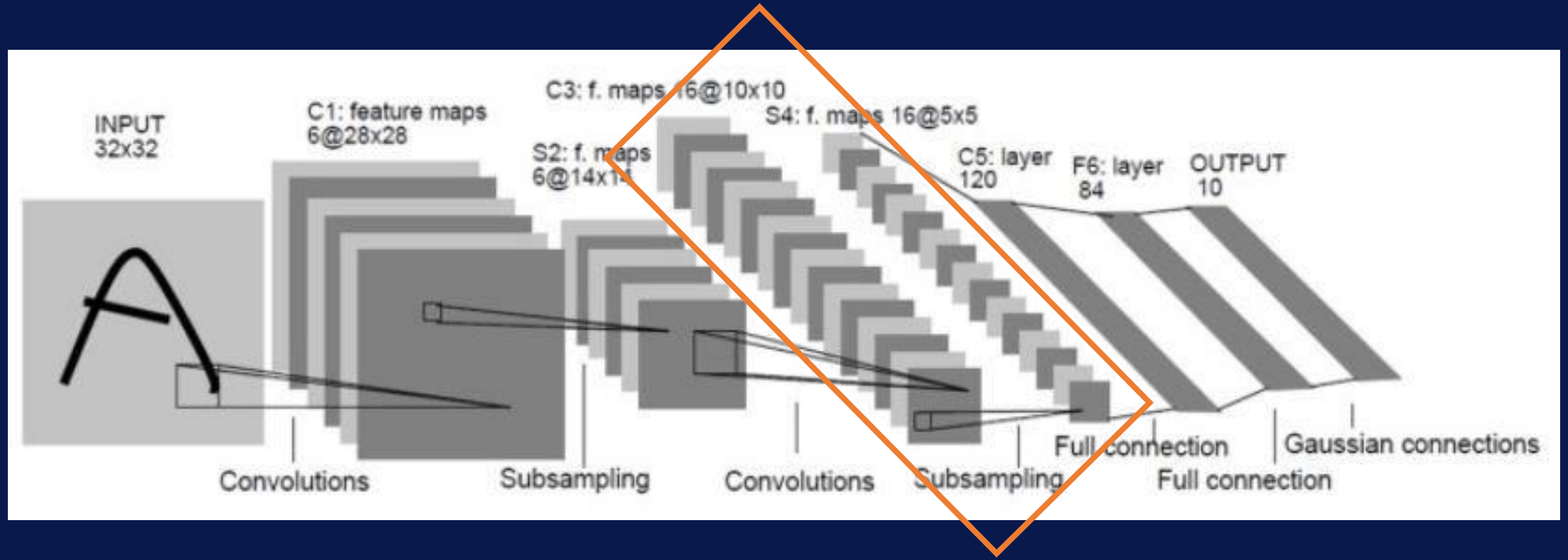
EACH COLUMN INDICATES WHICH FEATURE MAP IN S2 ARE COMBINED BY THE UNITS IN A PARTICULAR FEATURE MAP OF C3.

① 연속된 3장을 모아 Convolution
→ 6장의 10x10 feature map 생성
② 연속된 4장을 모아 Convolution
→ 6장의 10x10 feature map 생성
③ 불연속한 4장을 모아 Convolution
→ 3장의 10x10 feature map 생성
④ 6장 모두 Convolution
→ 1장의 10x10 feature map 생성

⇒ 16장(6장+6장+3장+1장) 의 10x10 feature map 생성

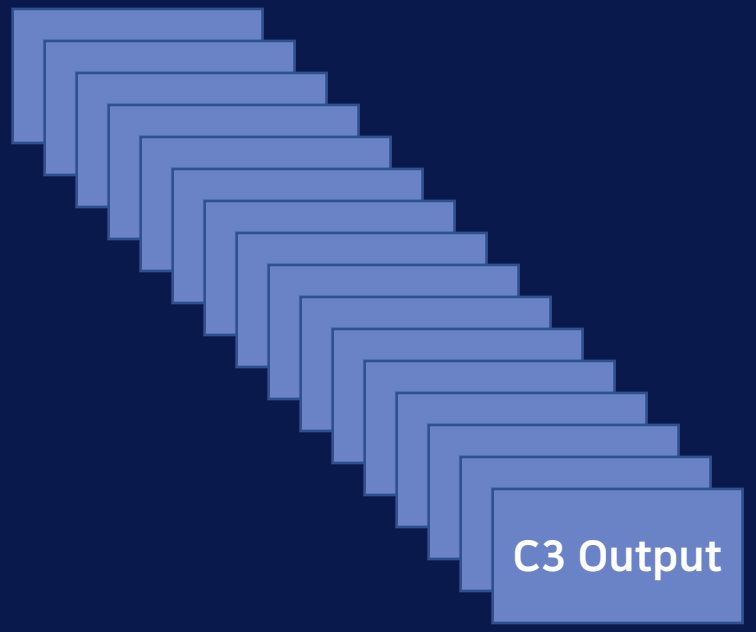
01. Lenet

LeNet Architecture



01. Lenet

C3 layer

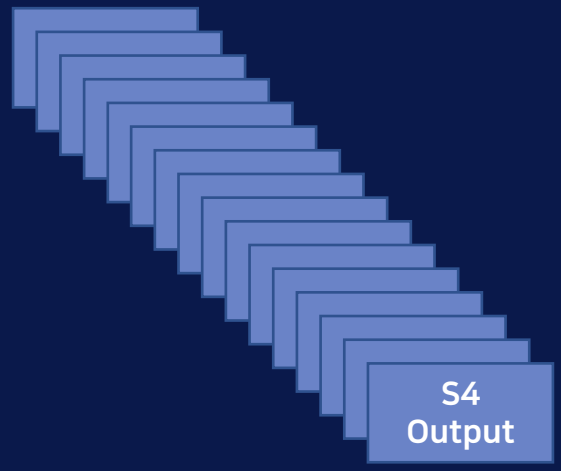


C3 Output size : (16, 10, 10)

Subsampling

Pooling size : 2x2

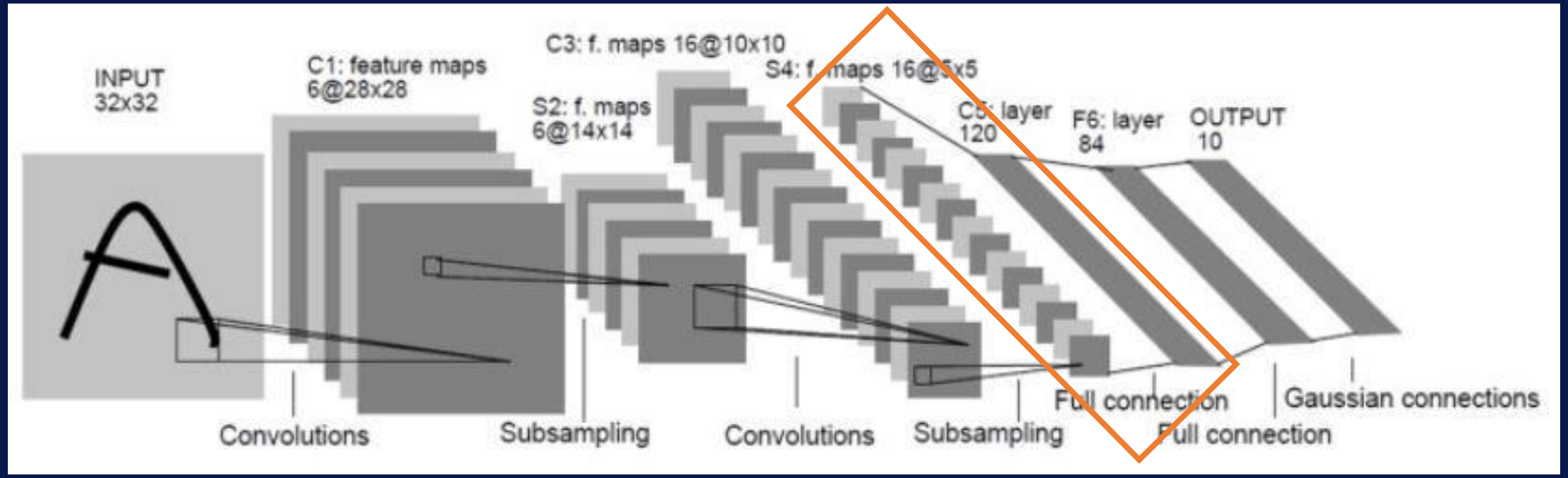
S4 layer



S4 Output size : (16, 5, 5)

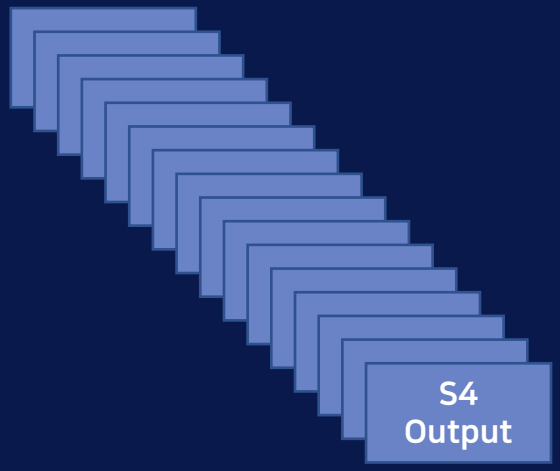
01. Lenet

LeNet Architecture



01. Lenet

S4 layer



S4 Output size : (16, 5, 5)

Convolution

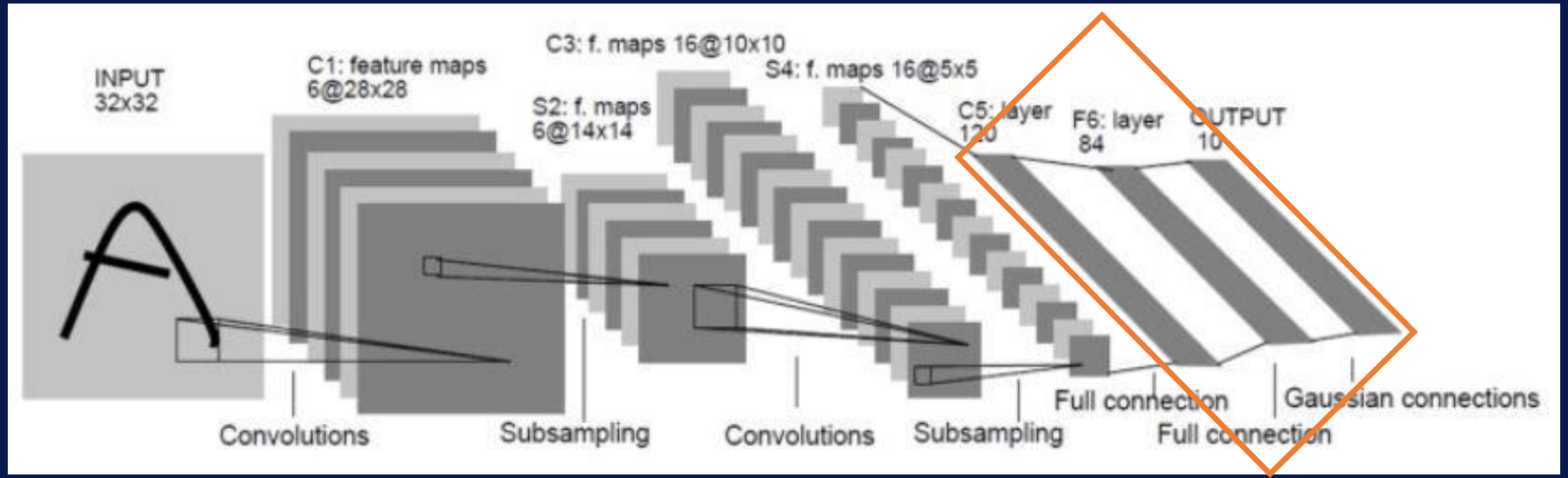
Filter size : 5x5
Filter 개수 : 120
Stride : 1

C5 layer

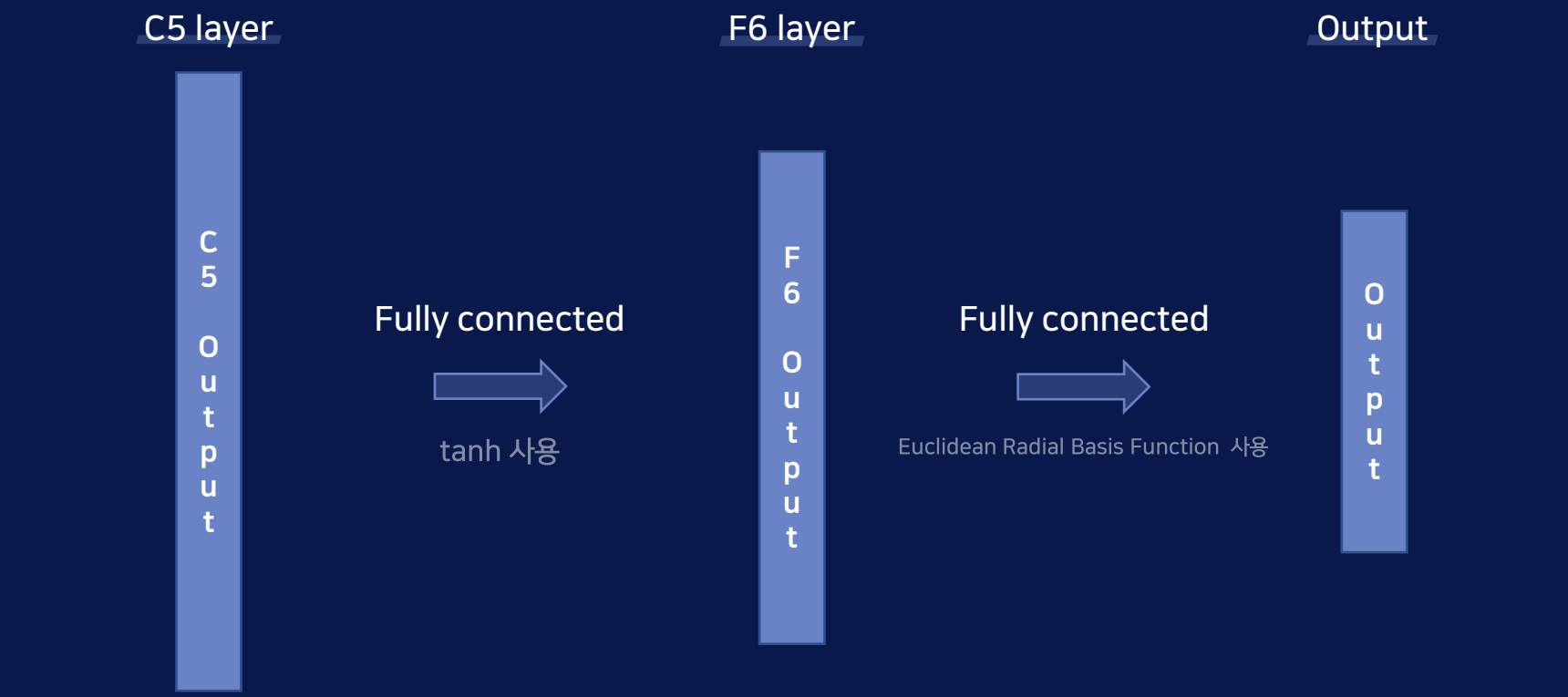
C5 Output size : (120, 1, 1)

01. Lenet

LeNet Architecture



01. Lenet



C5 Output size : (120)

F6 Output size : (84)

F7 Output size : (10)

02. Alexnet

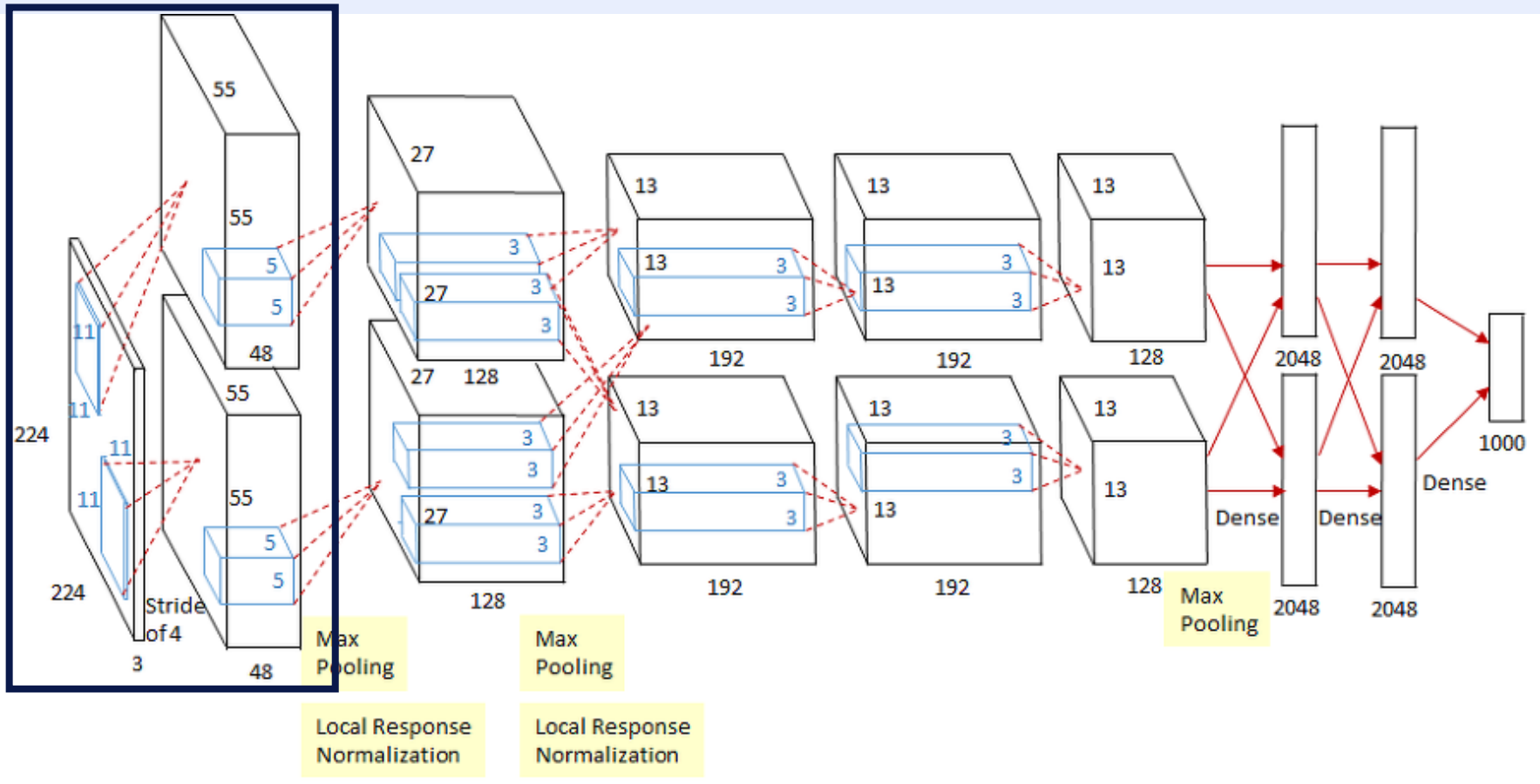
- ILSVRC(ImageNet Large-Scale Visual Recognition Challenge)의 2012년 대회에서 1위를 한 CNN 모델이다.
  - ILSVRC란?
    이미지 인식(image recognition) 경진대회로 대용량의 이미지 데이터셋을 주고 이미지 분류 알고리즘의 성능을 평가한다.

- 논문 'ImageNet Classification with Deep Convolutional Neural Networks'에 수록되어 있다.
- 2개의 GPU로 병렬연산을 수행하기 위해 병렬적인 구조로 설계되었다.

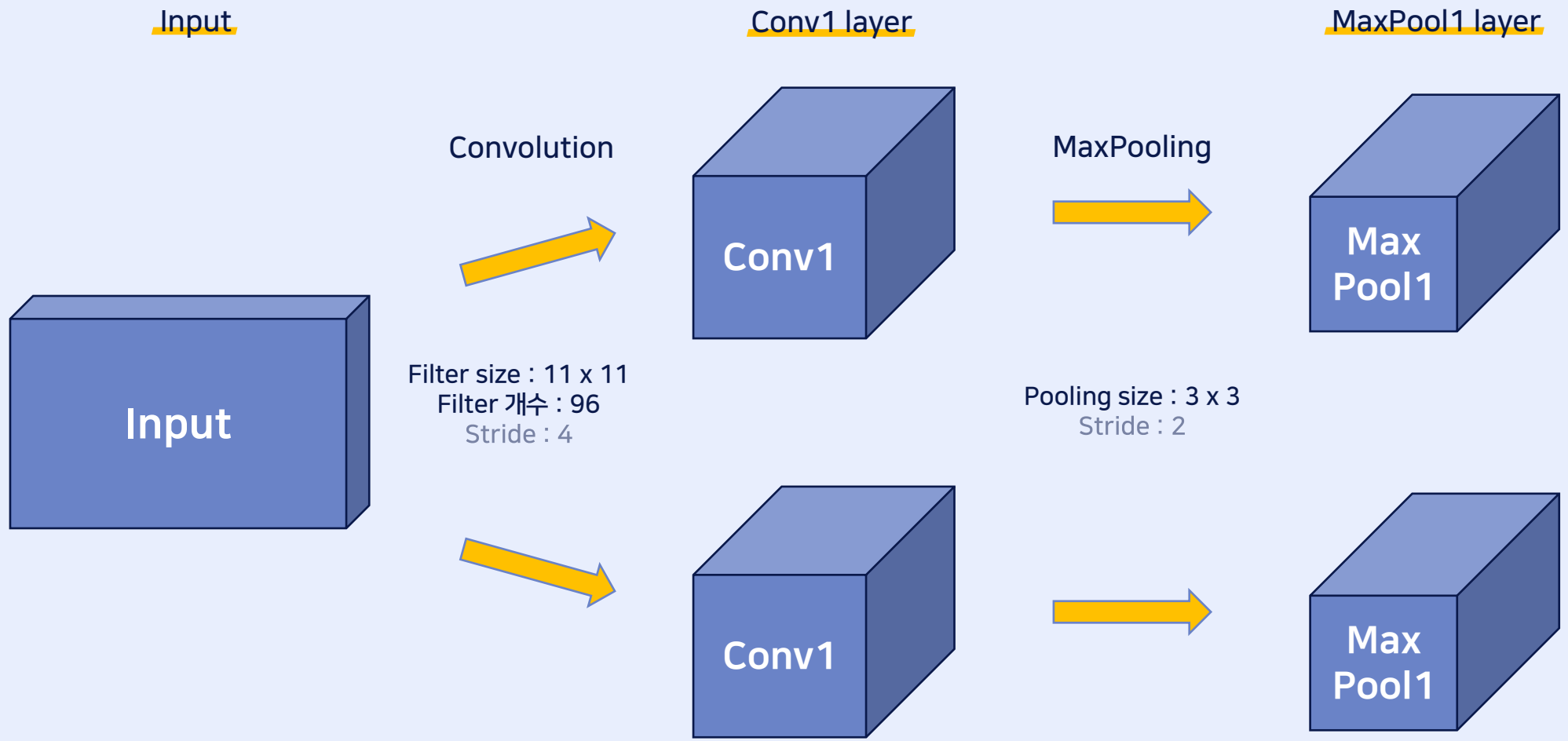
- [Input – Conv1 – MaxPool1 – Norm1 – Conv2 – MaxPool2 – Norm2 – Conv3 – Conv4 – Conv5 – MaxPool5 – FC6- FC7- FC8(output)]

02. Alexnet

AlexNet Architecture



02. Alexnet



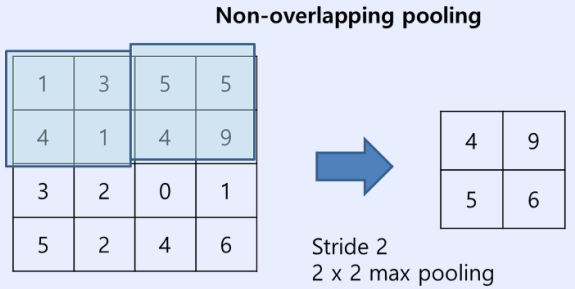
Input size : (3, 227, 227)

Conv1 Output size : (48, 55, 55)

MaxPool1 Output size : (48, 27, 27)

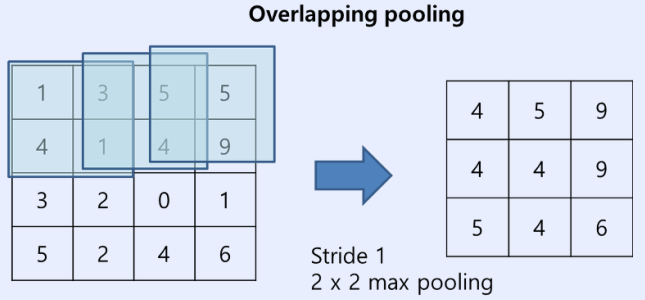
02. Alexnet

[Input. . .– **Conv1 – MaxPool1 – Norm1 -** . . .Output]



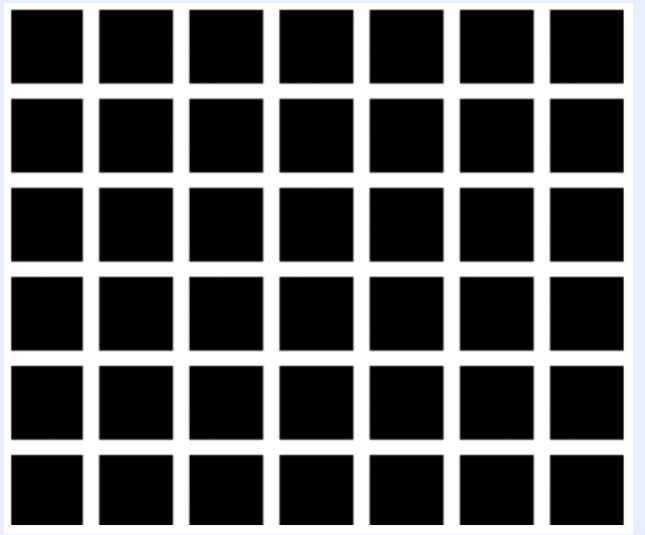
MaxPooling layer : overlapping maxpooling 사용

- LeNet-5의 경우, non-overlapping average pooling을 사용하며, 일반적으로 다른 CNN모델도 non-overlapping pooling을 사용한다.



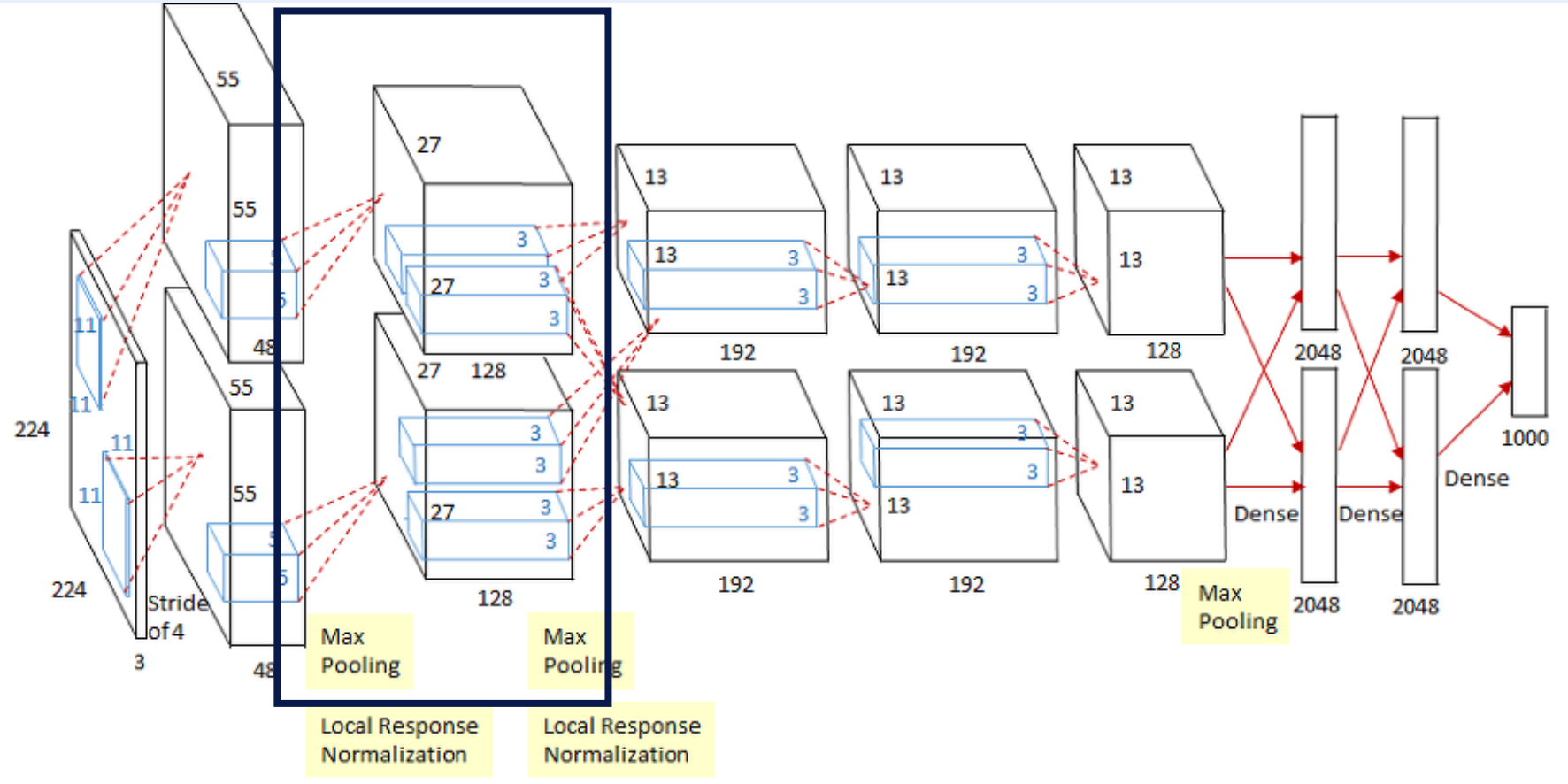
Normalization layer : LRN (local response normalization) 시행

- AlexNet에서 처음 도입했다.
- 인접 화소들을 억제시키고, 특징을 부각시키기 위한 방법이다.
- 신경생물학에서 원리를 가져온 것으로, 예를 들어 우측 그림의 검은 부분을 집중하여 보면 회색의 점이 보인다. 이는 강한 자극인 검정색이 약한 자극인 흰색의 인식을 막아 발생하는 '측면억제' 현상이다.
- 위 예시처럼 지역적인 현상을 약화하고 학습을 일반화 하기 위해 고안된 방법이다.

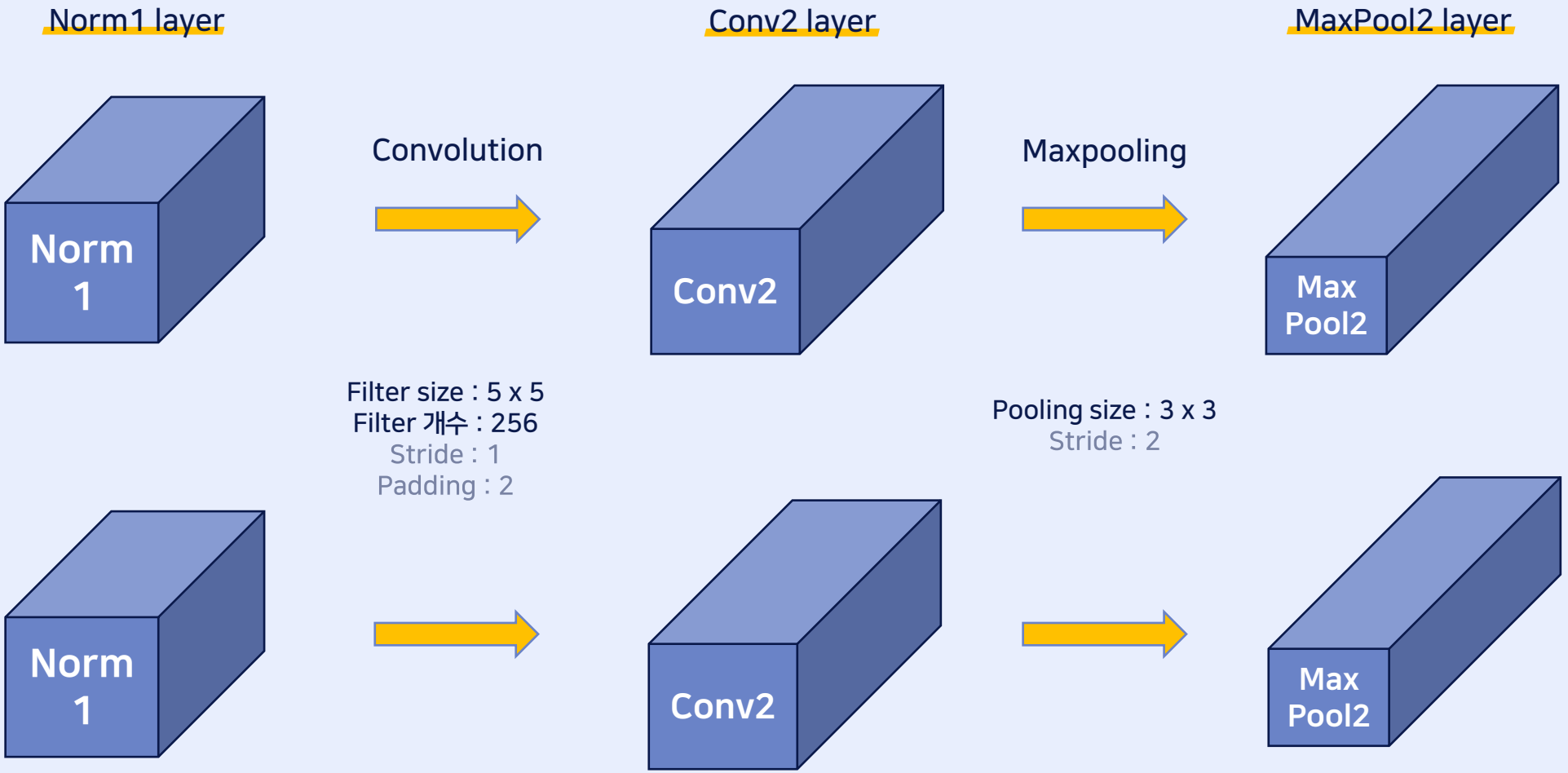


02. Alexnet

AlexNet Architecture



02. Alexnet



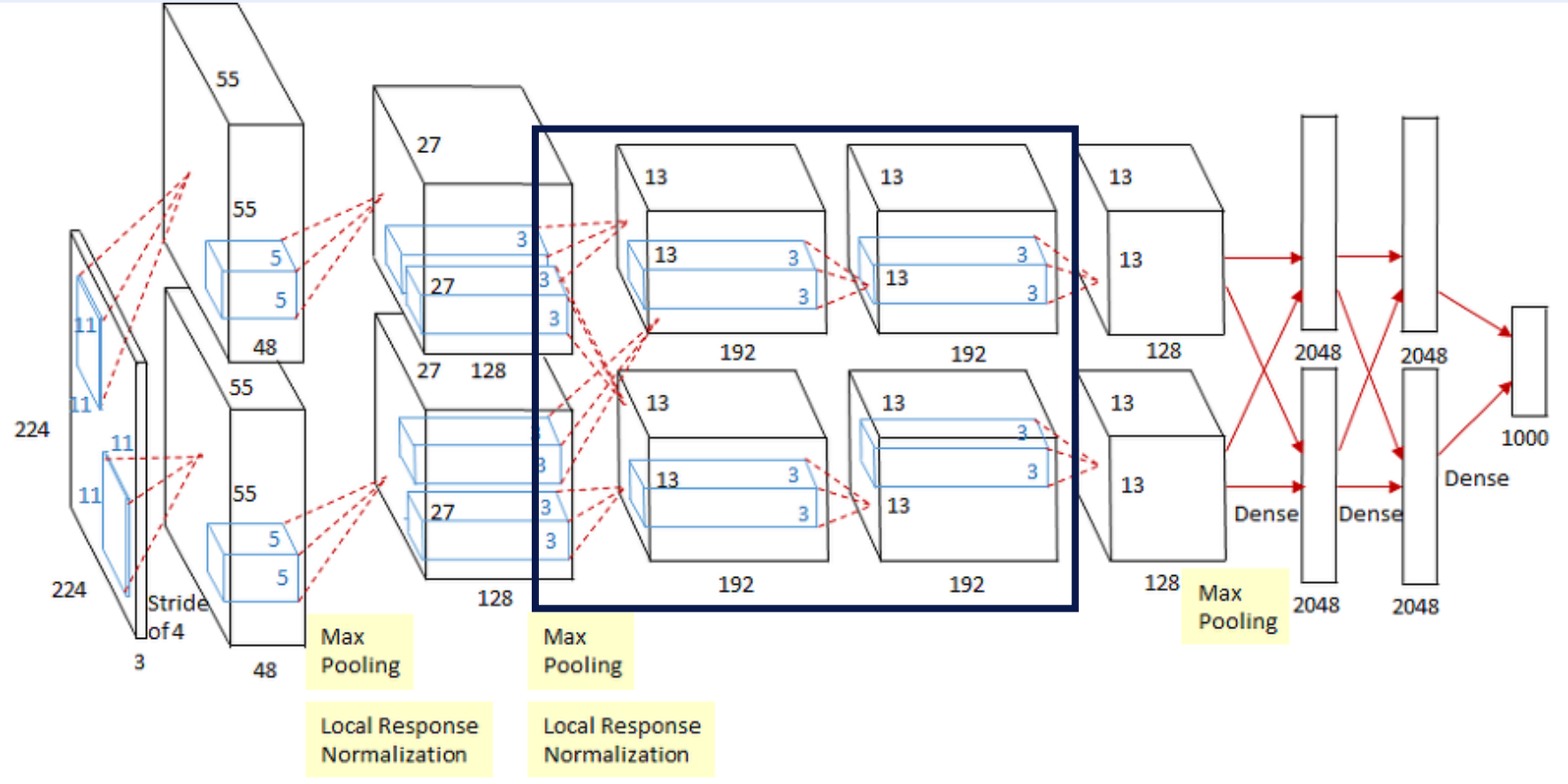
Norm1 Output size : (48, 27, 27)

Conv2 Output size : (128, 27, 27)

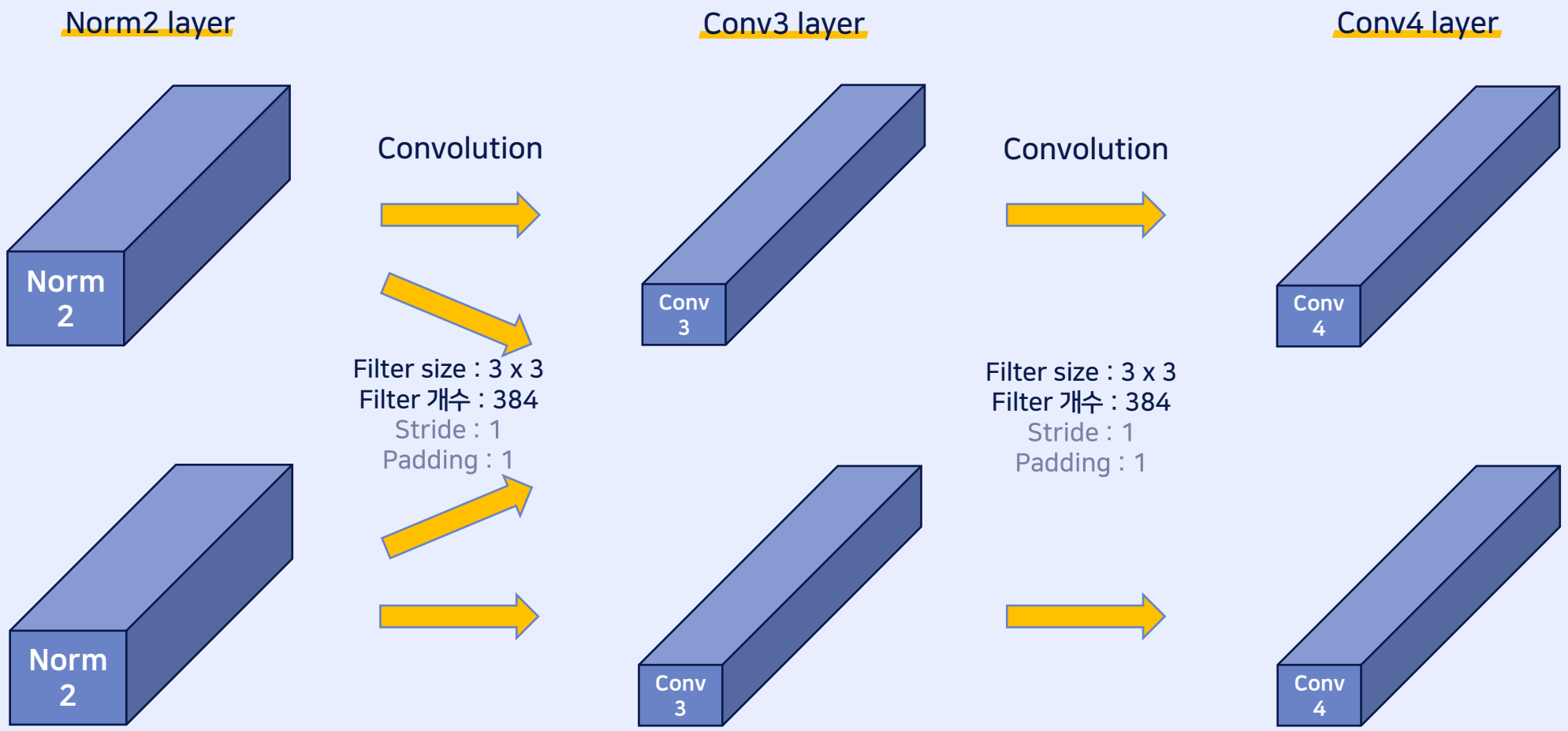
MaxPool2 Output size : (128, 13, 13)

02. Alexnet

AlexNet Architecture



02. Alexnet



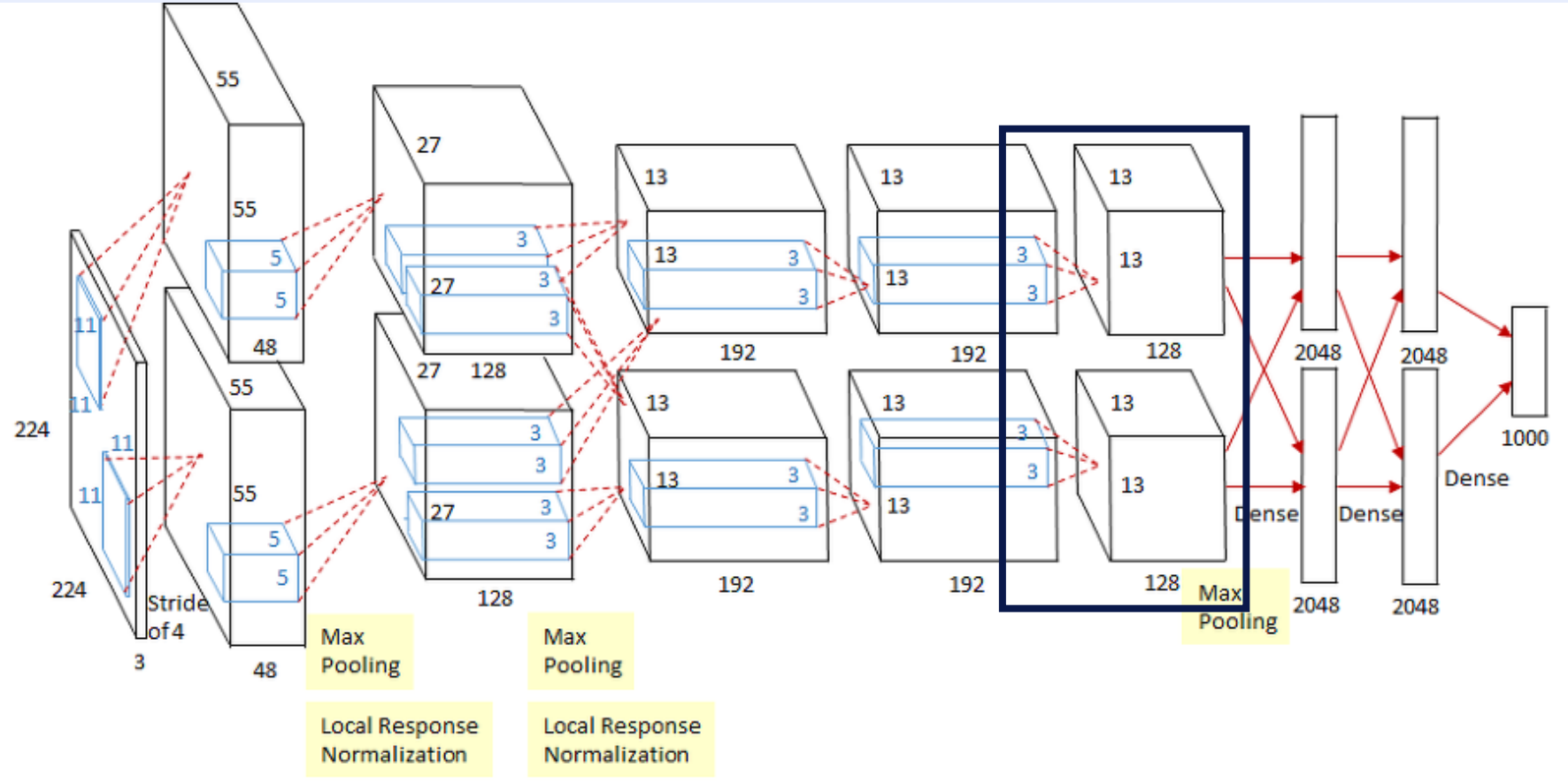
Norm2 Output size : (128, 13, 13)

Conv3 Output size : (192, 13, 13)

Conv4 Output size : (192, 13, 13)

02. Alexnet

AlexNet Architecture

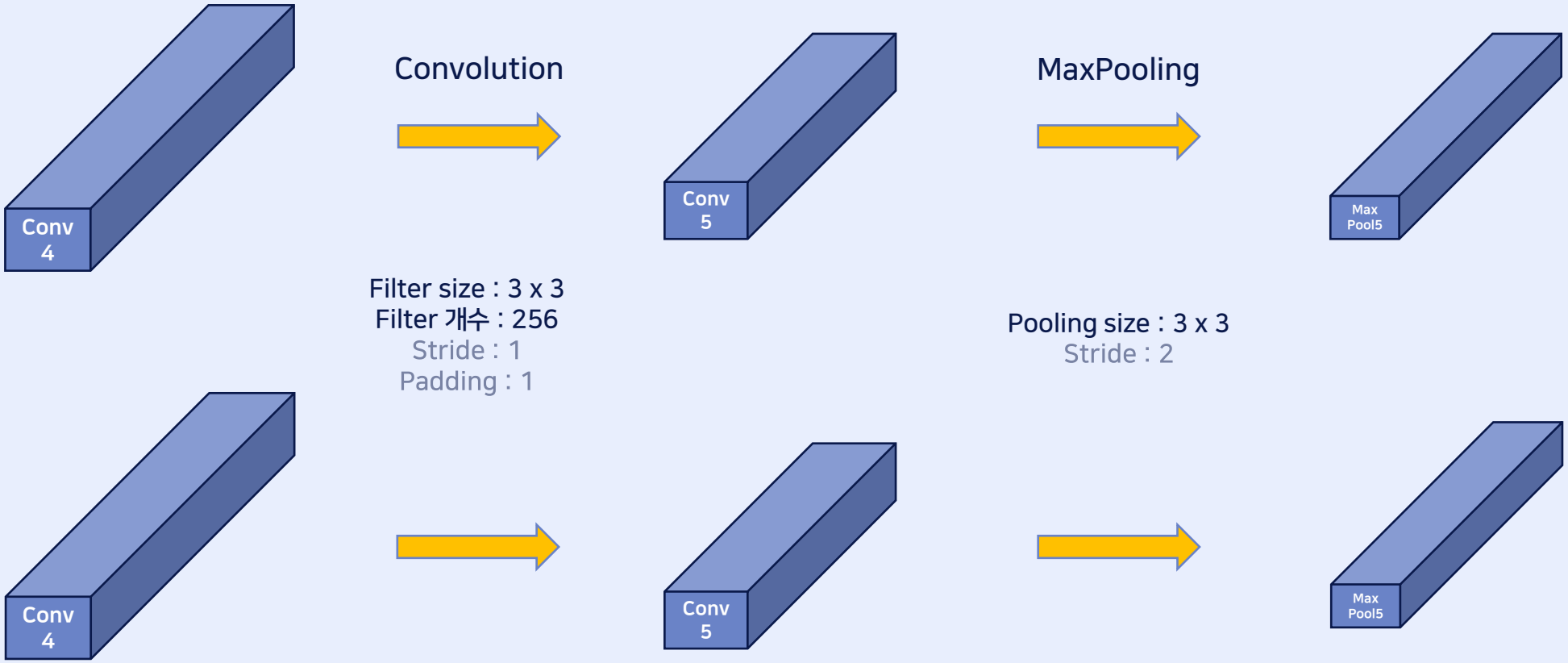


02. Alexnet

Conv4 layer

Conv5 layer

MaxPool5 layer



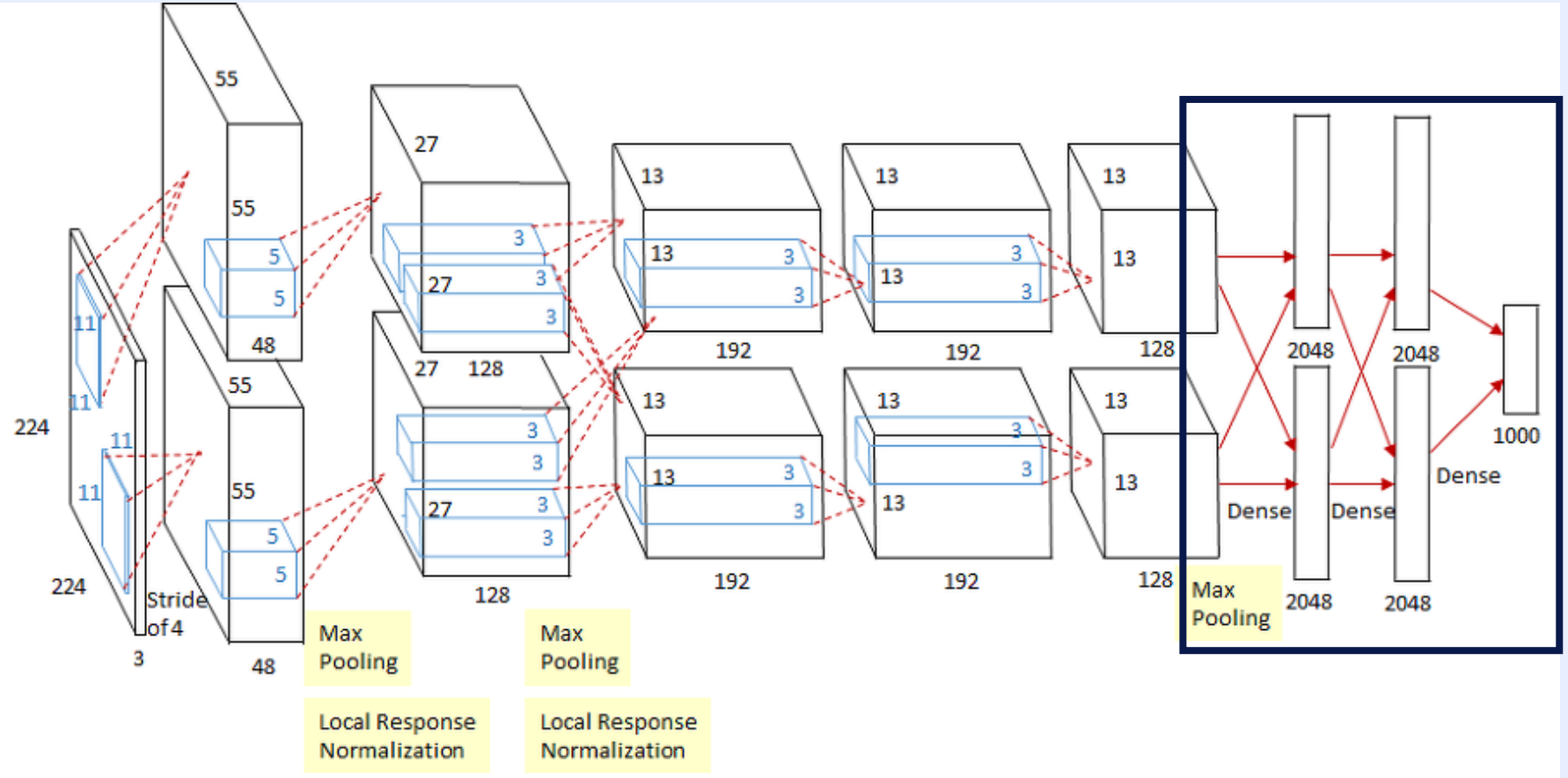
Conv4 Output size : (192, 13, 13)

Conv5 Output size : (128, 13, 13)

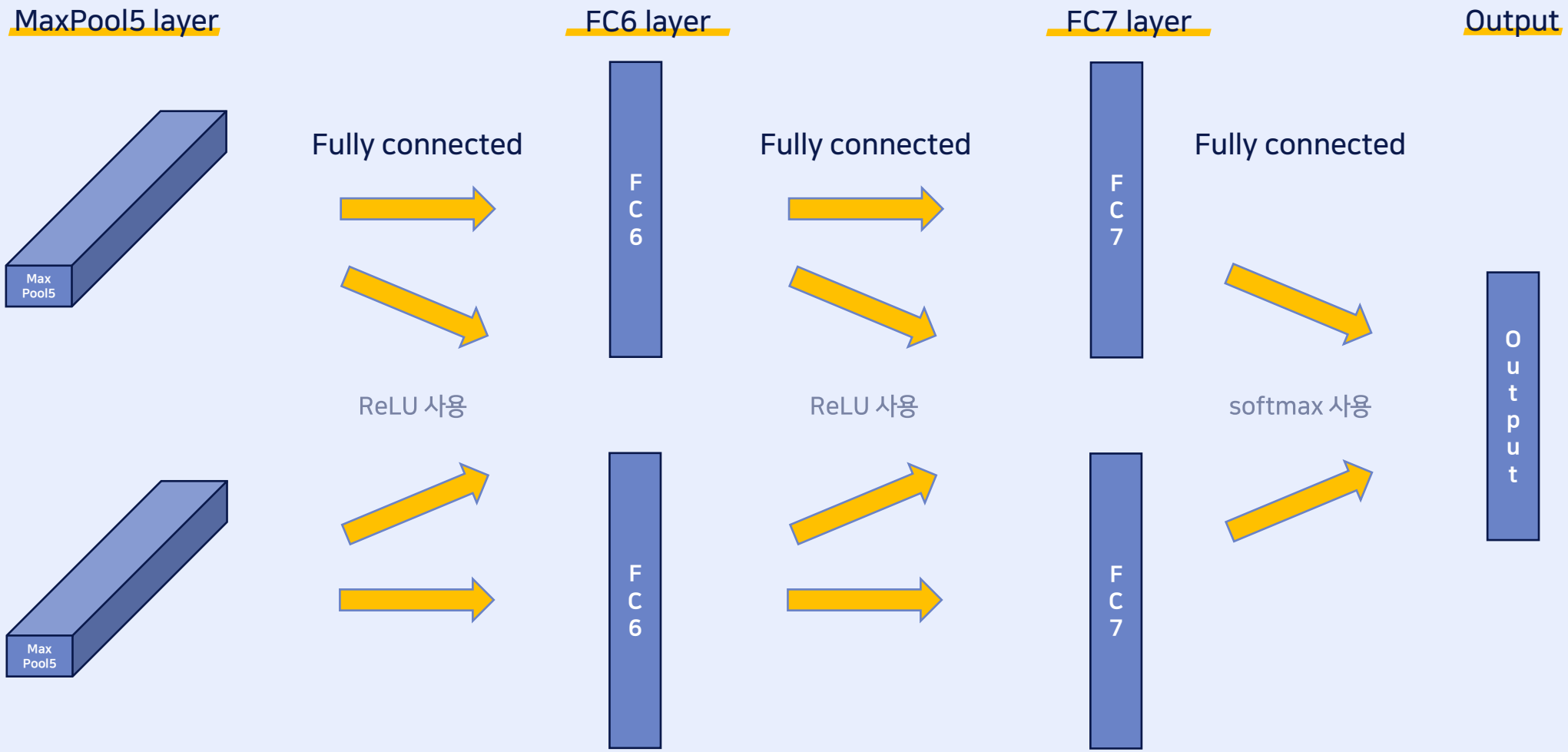
MaxPool5 Output size : (128, 6, 6)

02. Alexnet

AlexNet Architecture



02. Alexnet



MaxPool5 Output size : (128, 6, 6)

FC6 Output size : (2048)

FC7 Output size : (2048)

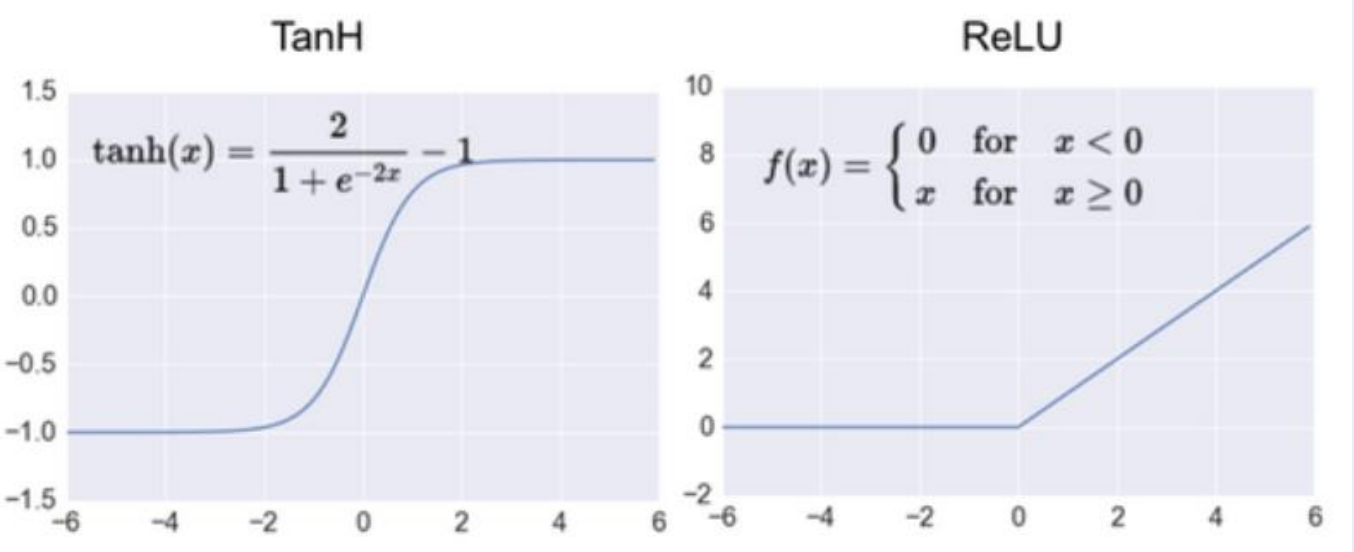
FC8 Output size : (1000)

02. Alexnet

[Input. . . – **FC6 – FC7** - Output]

Fully connected layer : ReLU function 사용

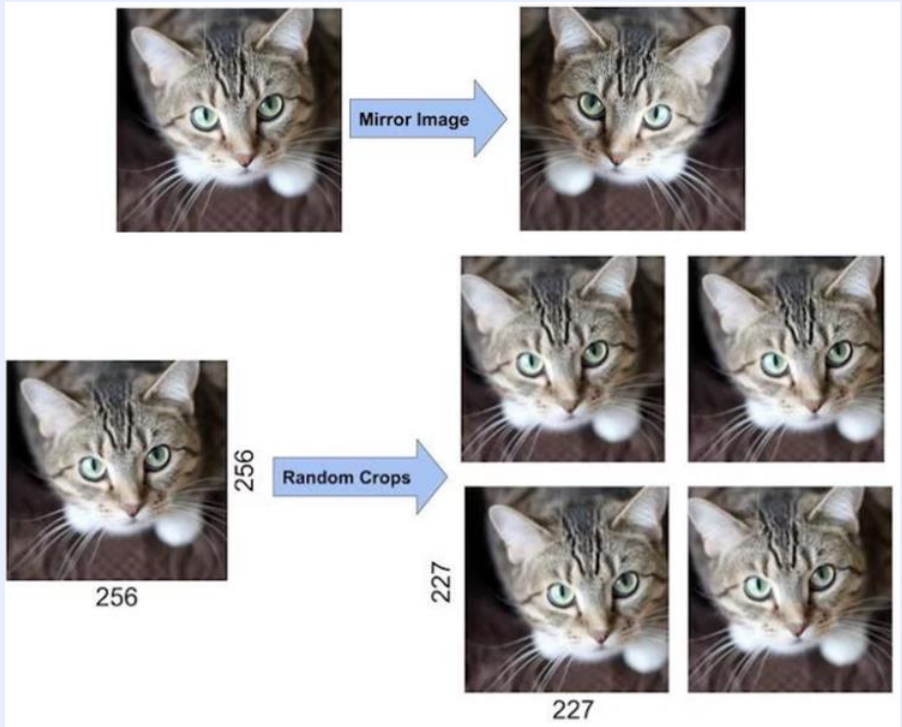
- LeNet-5에서는 tanh 사용
- ReLU를 사용하는 것이 같은 정확도를 유지하면서 tanh을 사용하는 것보다 6배나 빨라 AlexNet 이후로는 ReLU 함수 주로 사용했다.



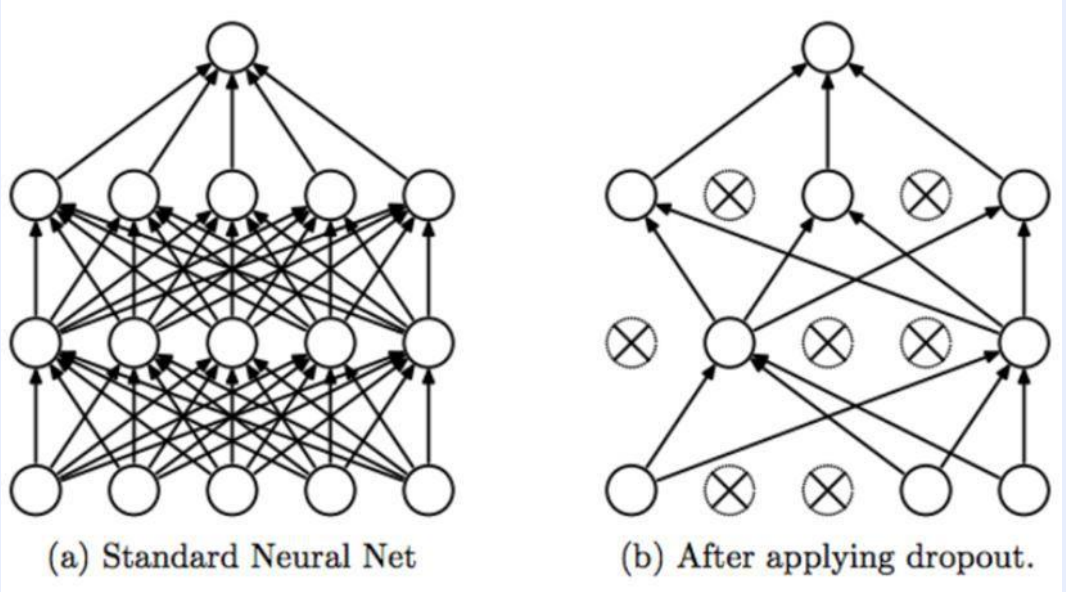
02. Alexnet

Overfitting을 막기위한 방법 2가지

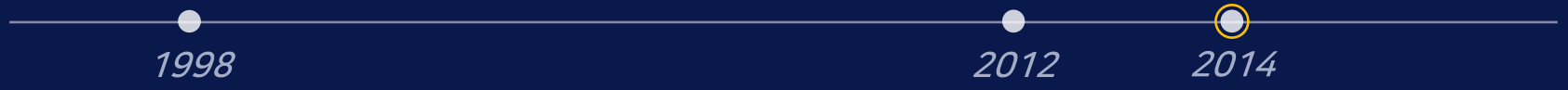
1. Data Augmentation



2. Dropout



03. VGG

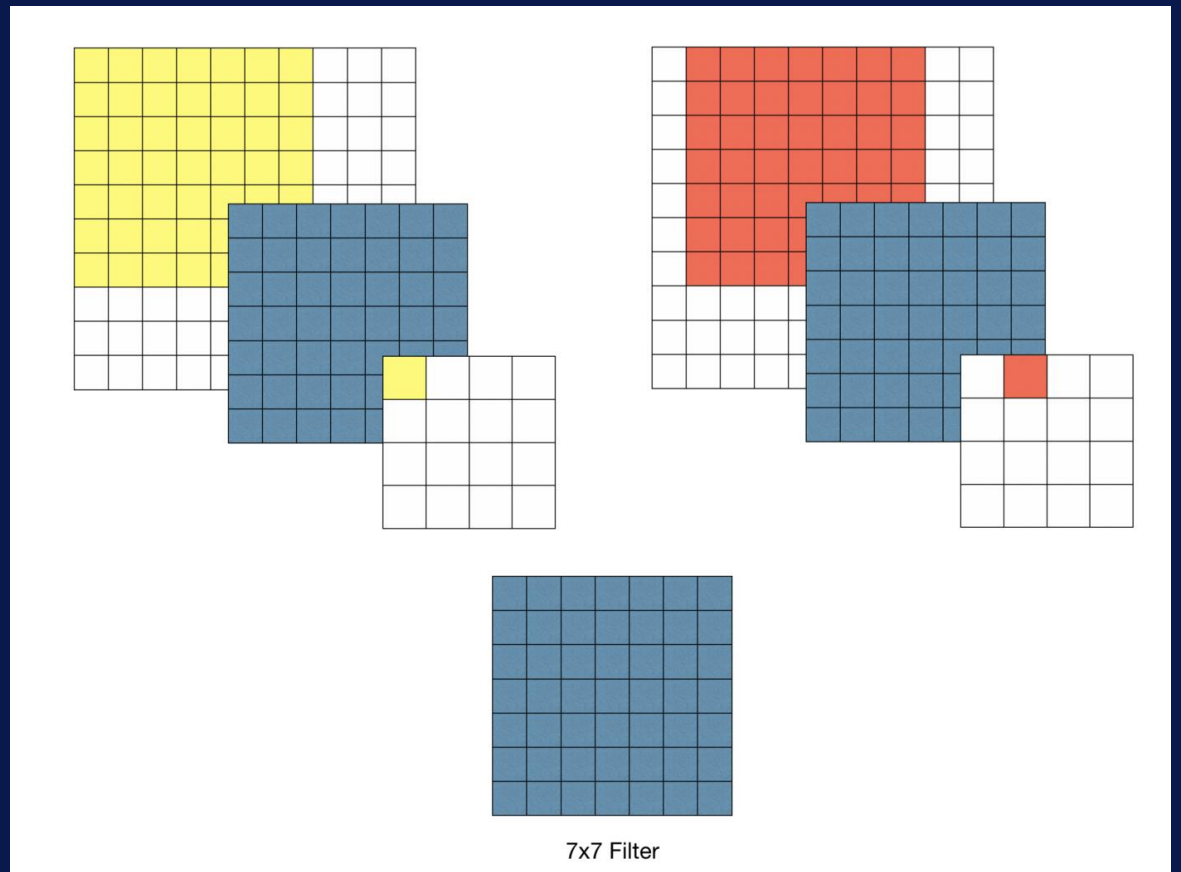


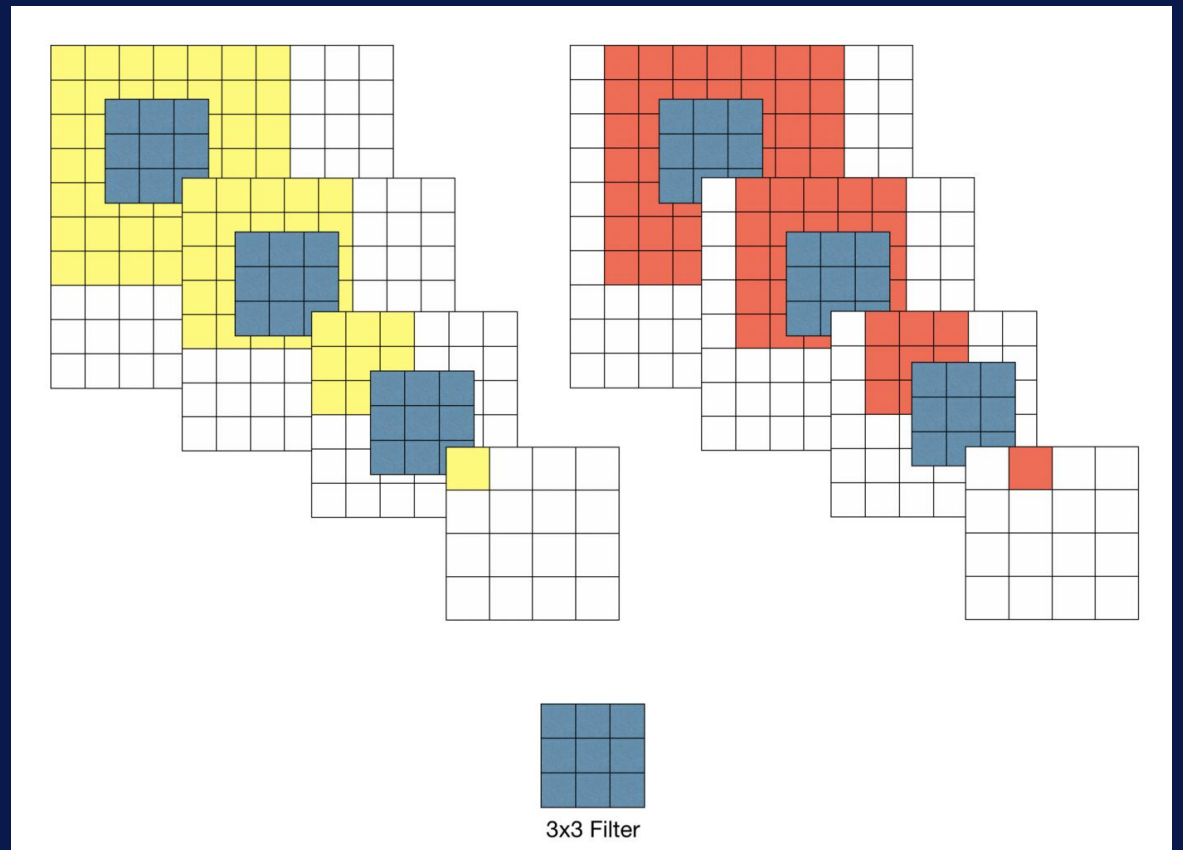
- ILSVRC(ImageNet Large-Scale Visual Recognition Challenge)의 2014년 대회에서 2위를 한 CNN 모델이다.
- 논문 'Very deep convolutional networks for large-scale image recognition'에 수록되어 있다.

- VGG부터 네트워크의 깊이가 확 깊어졌다. (VGG16 : 16개층, VGG19 : 19개층)
- [Input – C1 – C2 – MaxPool2 – C3 – C4 – MaxPool4 – C5 – C6 – C7 – MaxPool7 – C8 – C9 – C10 – MaxPool10 – C11 – C12 – C13 – MaxPool13 – FC14 – FC15 – FC16(Output)]

03. VGG

3*3 filter 사용





03. VGG

학습 이미지 data augmentation

256x256

224x224

224x224

224x224

224x224

512x512

224x224

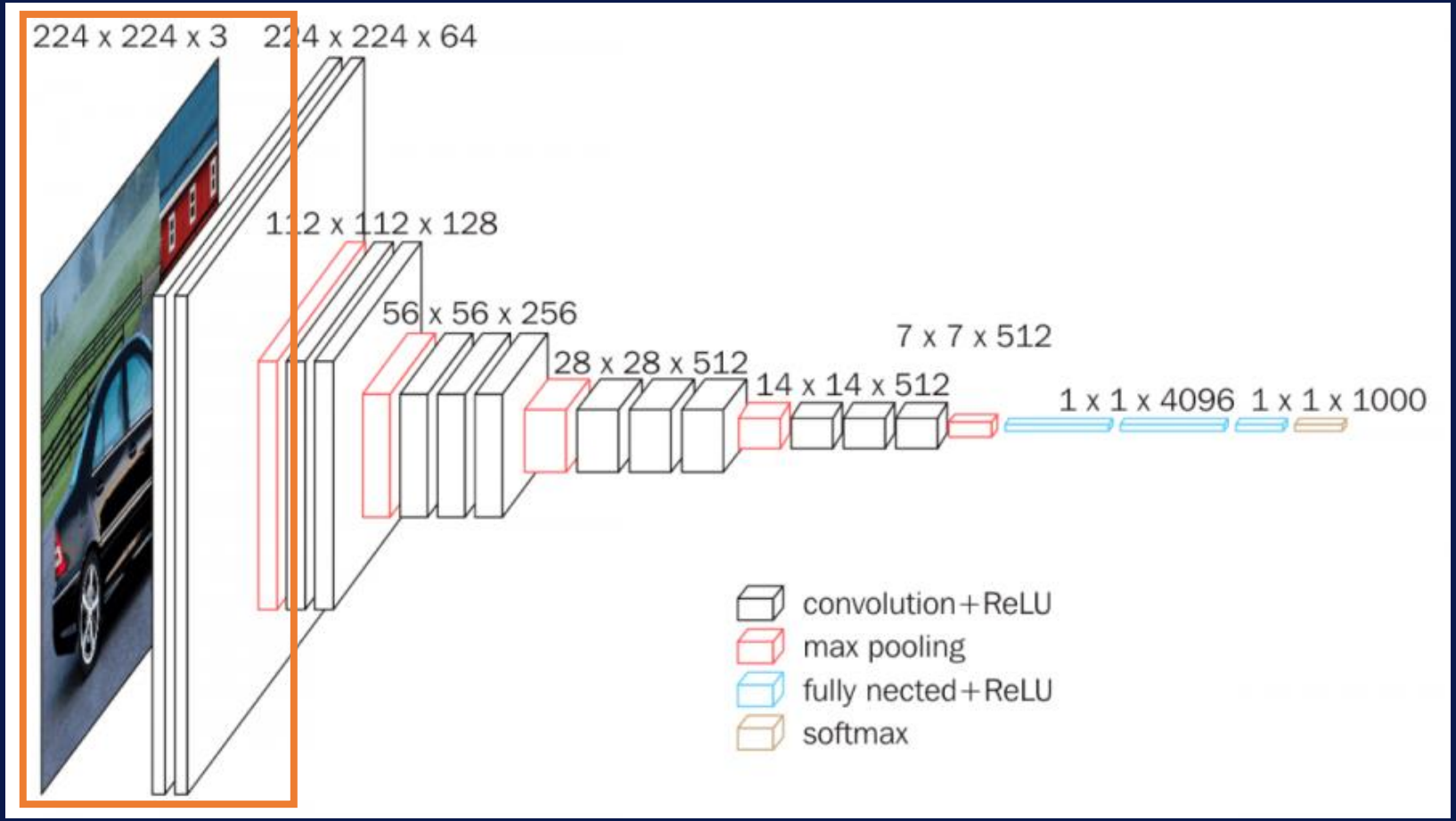
224x224

224x224

224x224

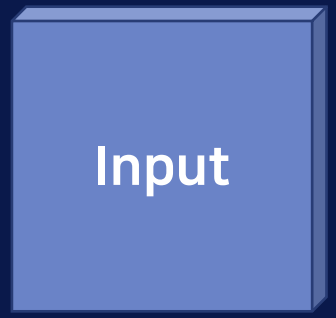
03. VGG

VGG Architecture



03. VGG

Input



Input size : (3, 224, 224)

convolution

Filter size : 3x3
Filter 개수 : 64
Stride : 1
Padding : 1

C1 layer

C1 Output size : (64, 224, 224)

convolution

Filter size : 3x3
Filter 개수 : 64
Stride : 1
Padding : 1

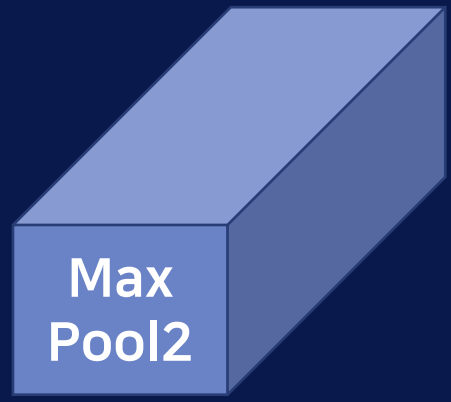
C2 layer

C2 Output size : (64, 224, 224)

Maxpooling

Pooling size : 2x2
Stride : 2

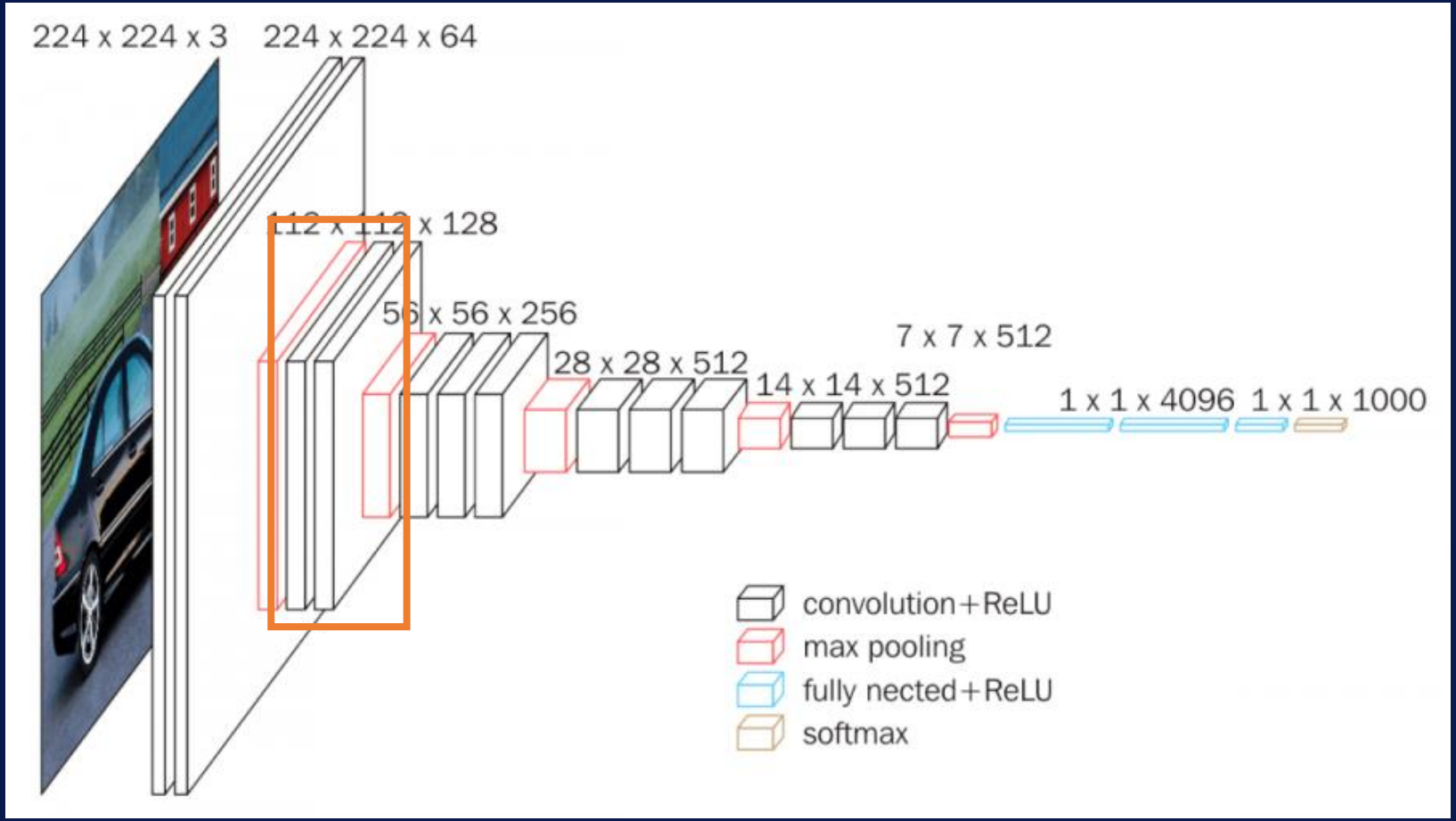
MaxPool2 layer



MaxPool2 Output size : (64, 112, 112)

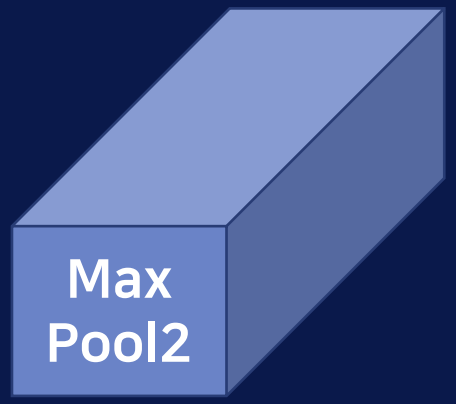
03. VGG

VGG Architecture



03. VGG

MaxPool2 layer

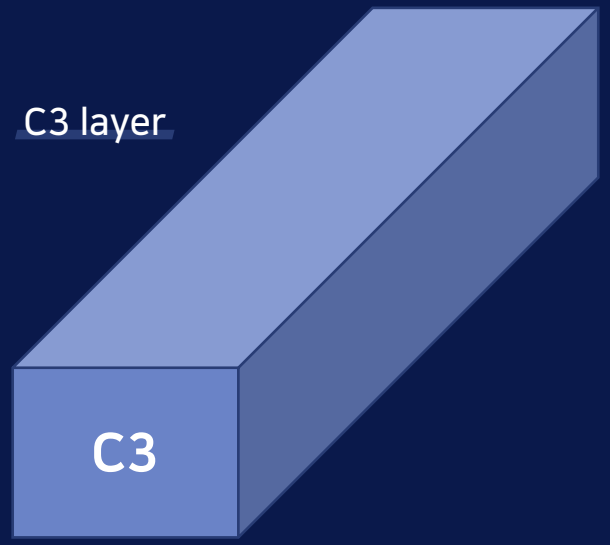


MaxPool2 Output size : (64, 112, 112)

convolution

Filter size : 3x3
Filter 개수 : 128
Stride : 1
Padding : 1

C3 layer

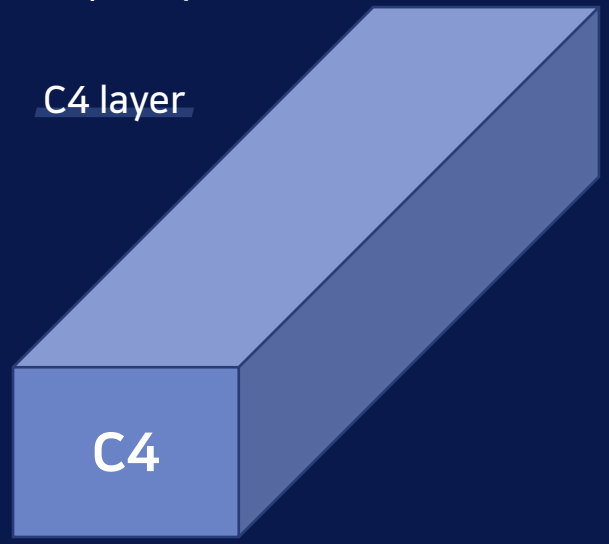


C3 Output size : (128, 112, 112)

convolution

Filter size : 3x3
Filter 개수 : 128
Stride : 1
Padding : 1

C4 layer

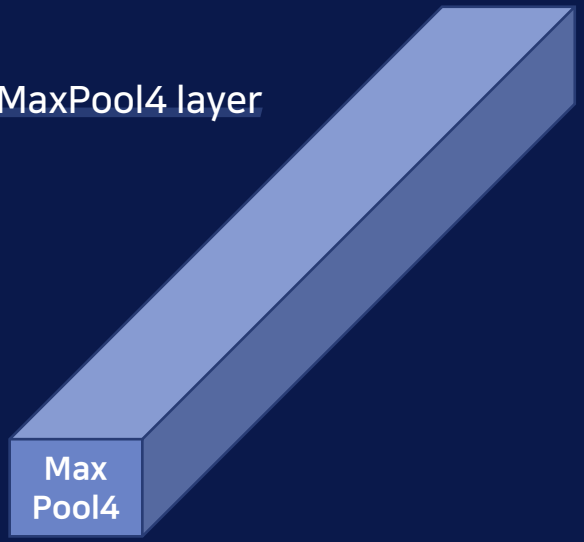


C4 Output size : (128, 112, 112)

Maxpooling

Pooling size : 2x2
Stride : 2

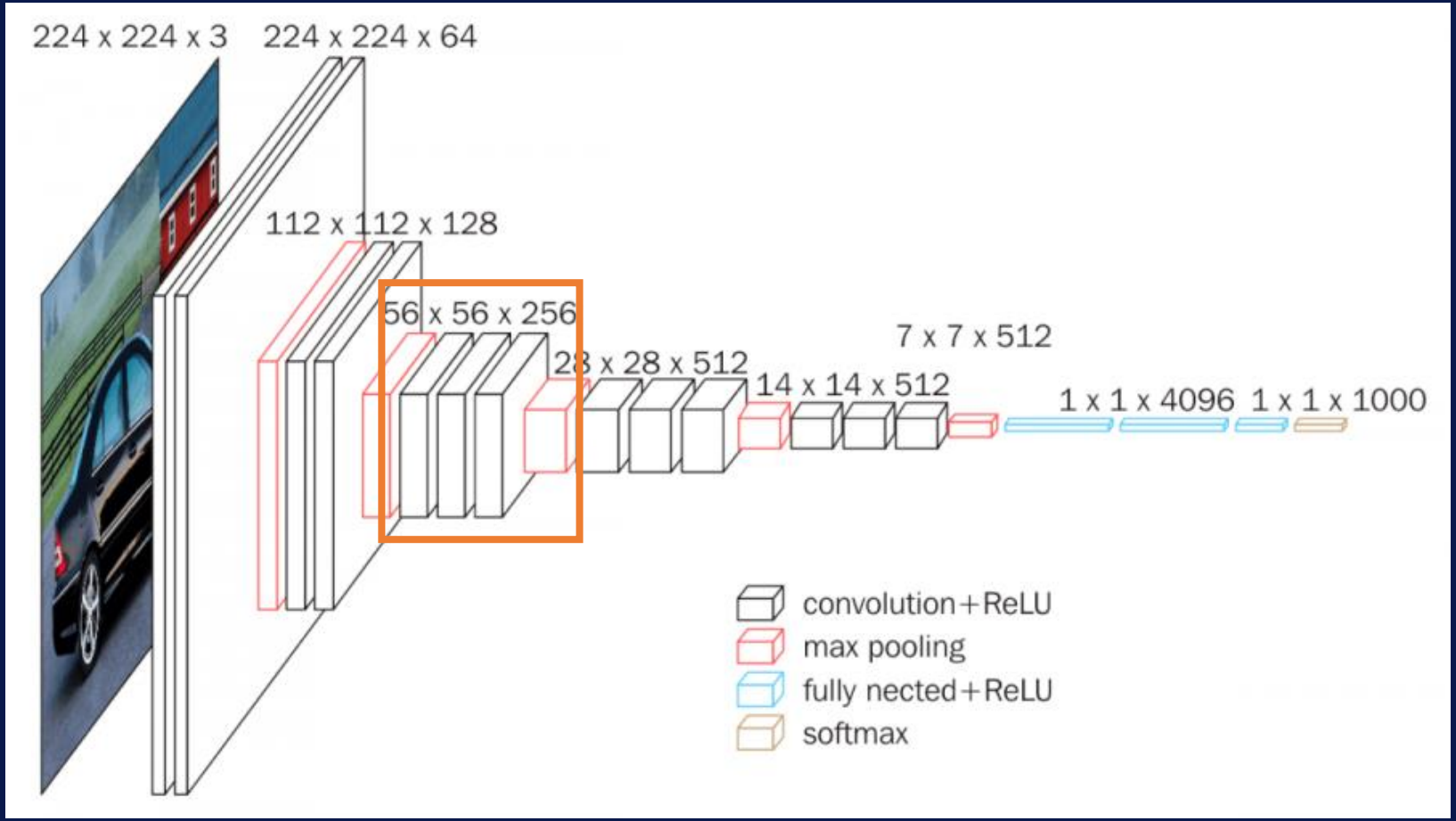
MaxPool4 layer



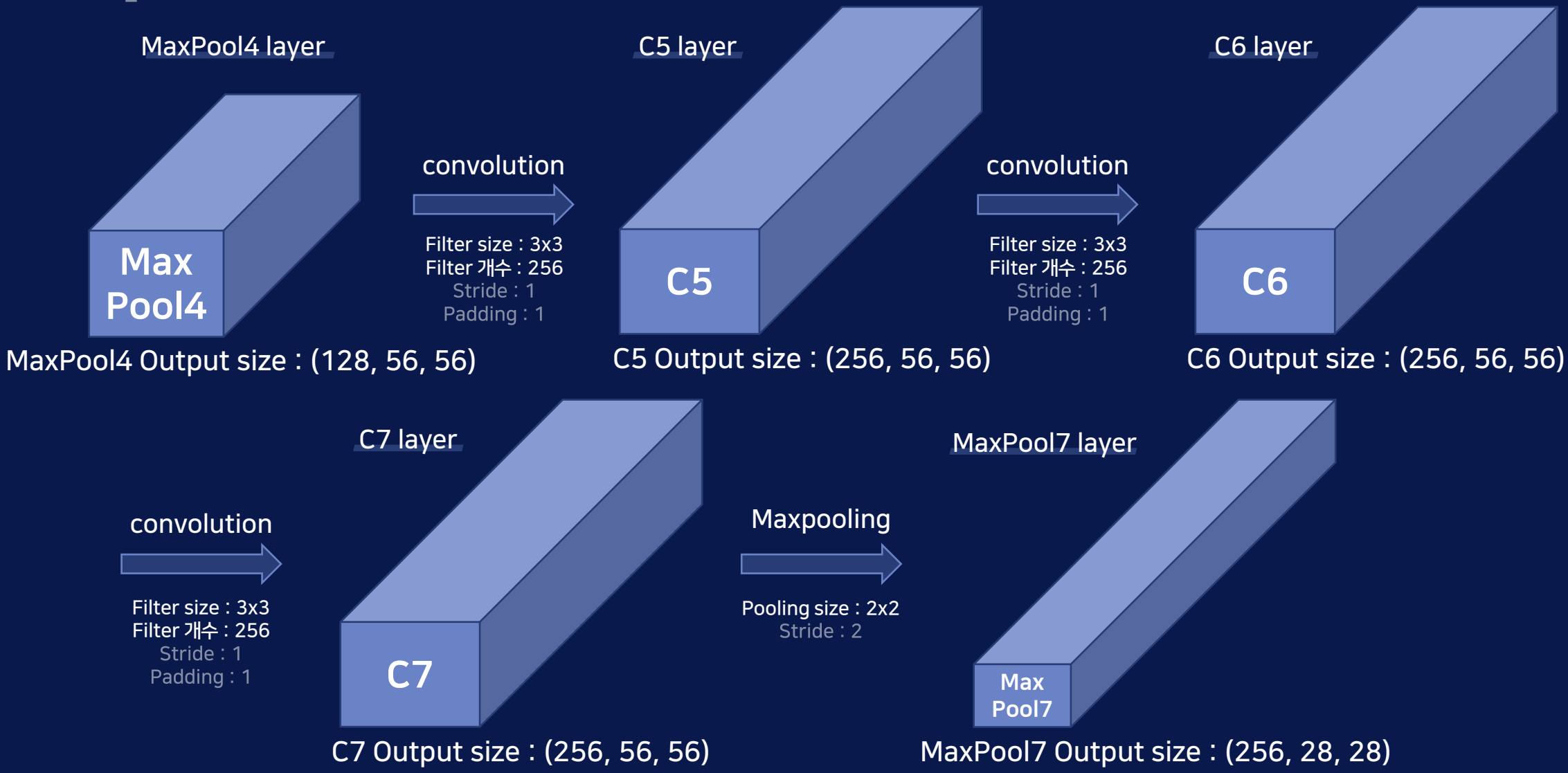
MaxPool4 Output size : (128, 56, 56)

03. VGG

VGG Architecture

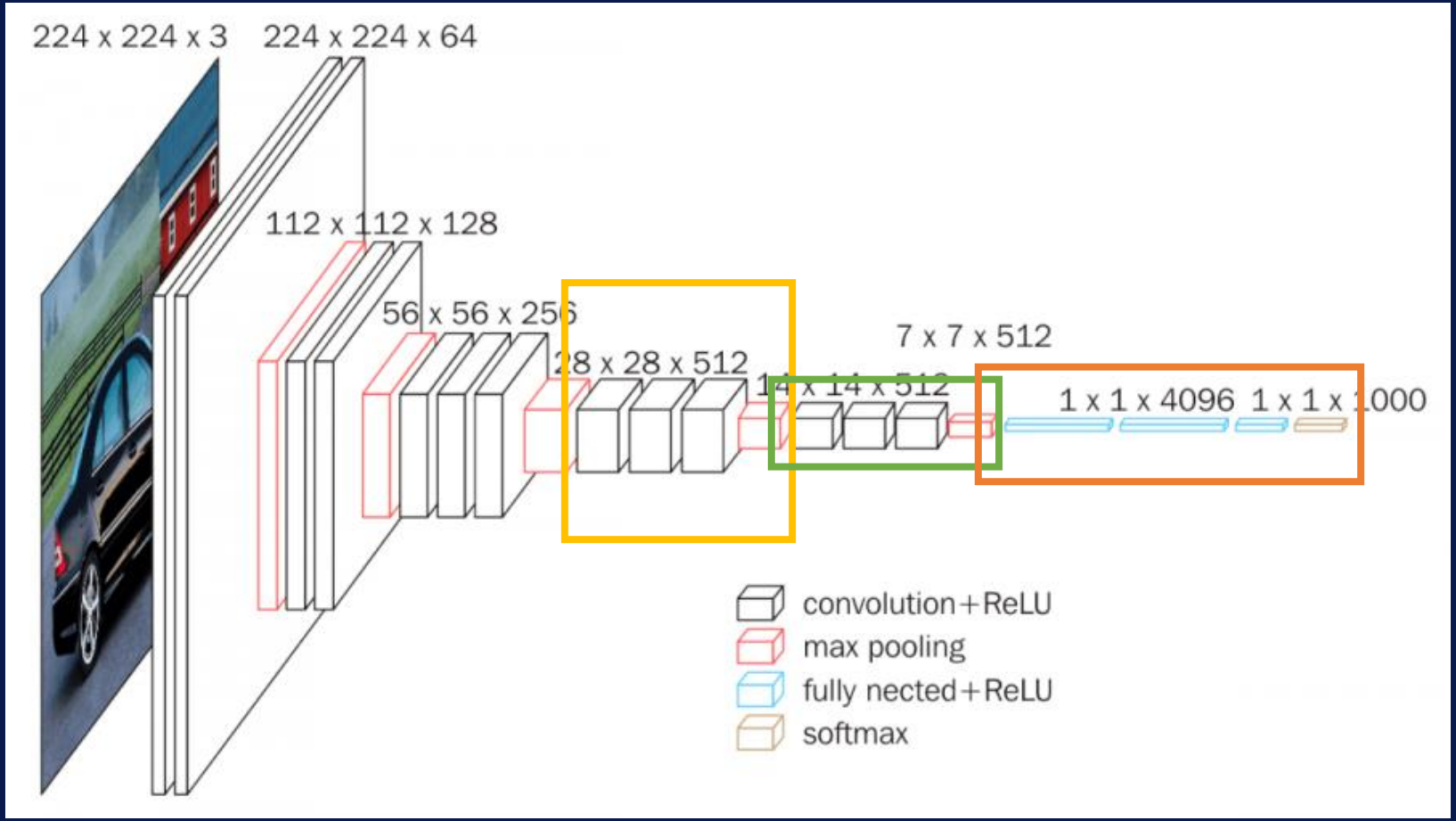


03. VGG


MaxPool4 layer
Max Pool4
convolution
Filter size : 3x3
Filter 개수 : 256
Stride : 1
Padding : 1
C5 layer
C5
convolution
Filter size : 3x3
Filter 개수 : 256
Stride : 1
Padding : 1
C6 layer
C6
MaxPool4 Output size : (128, 56, 56)
C5 Output size : (256, 56, 56)
C6 Output size : (256, 56, 56)
convolution
Filter size : 3x3
Filter 개수 : 256
Stride : 1
Padding : 1
C7 layer
C7
Maxpooling
Pooling size : 2x2
Stride : 2
MaxPool7 layer
Max Pool7
C7 Output size : (256, 56, 56)
MaxPool7 Output size : (256, 28, 28)

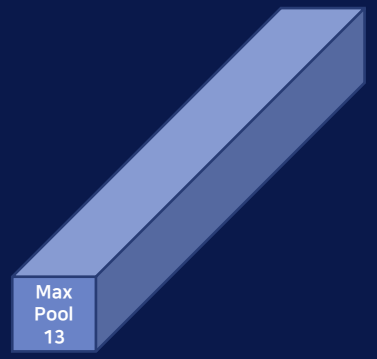
03. VGG

VGG Architecture



03. VGG

MaxPool13 layer



Fully connected

ReLU 사용

FC14 layer

F
C
1
4

MaxPool13 Output size : (512, 7, 7)

FC14 Output size : (4096, 1, 1)

FC15 layer

Fully connected

ReLU 사용

F
C
1
5

Output

Fully connected

Softmax 사용

O
u
t
p
u
t

FC15 Output size : (4096, 1, 1)

FC16 Output size : (1000, 1, 1)

과제

1. Alexnet 주석달기
2. VGG 논문 리뷰하기

Reference

01. LeNet

- LeNet architecture
https://deep-learning-study.tistory.com/368

- Convoution layer Table
https://deep-learning-study.tistory.com/368

- 논문
http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf

02. AlexNet

- Maxpooling layer
https://bskyvision.com/421

- Normalization layer
https://daeun-computer-uneasy.tistory.com/33

- Alexnet architecture
https://wiserloner.tistory.com/1126

- Relu function 사용
https://daeun-computer-uneasy.tistory.com/33

- Data augmentation / Drop out
https://bskyvision.com/421

- 논문
https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

03. VGG

- 3*3 filter 사용
https://medium.com/@msmapark2/vgg16-%EB%85%BC%EB%AC%B8-%EB%A6%AC%EB%B7%B0-very-deep-convolutional-networks-for-large-scale-image-recognition-6f748235242a

- Data augmentation
https://medium.com/@msmapark2/vgg16-%EB%85%BC%EB%AC%B8-%EB%A6%AC%EB%B7%B0-very-deep-convolutional-networks-for-large-scale-image-recognition-6f748235242a

- VGG architecture
https://hnsuk.tistory.com/30

- 논문
https://arxiv.org/pdf/1409.1556.pdf%20http://arxiv.org/abs/1409.1556.pdf

04. 폰트

- 네이버 글꼴
https://hangeul.naver.com/font

D&A

Deep Session 5차시 CNN 심화

Thank You.

2022 / 04 / 07

D&A 운영진 윤경서