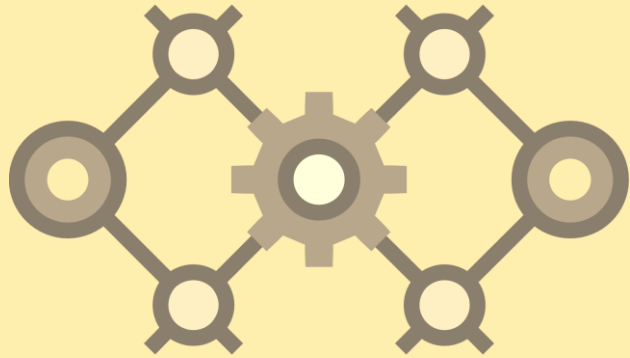


Python을 활용한 이미지 분석

7. 다양한 합성곱 신경망 모델



다양한 합성곱 신경망 모델

01 AlexNet(2012)

02 VGGNet(2014)

03 GoogLeNet(2014)

04 ResNet(2015)

1

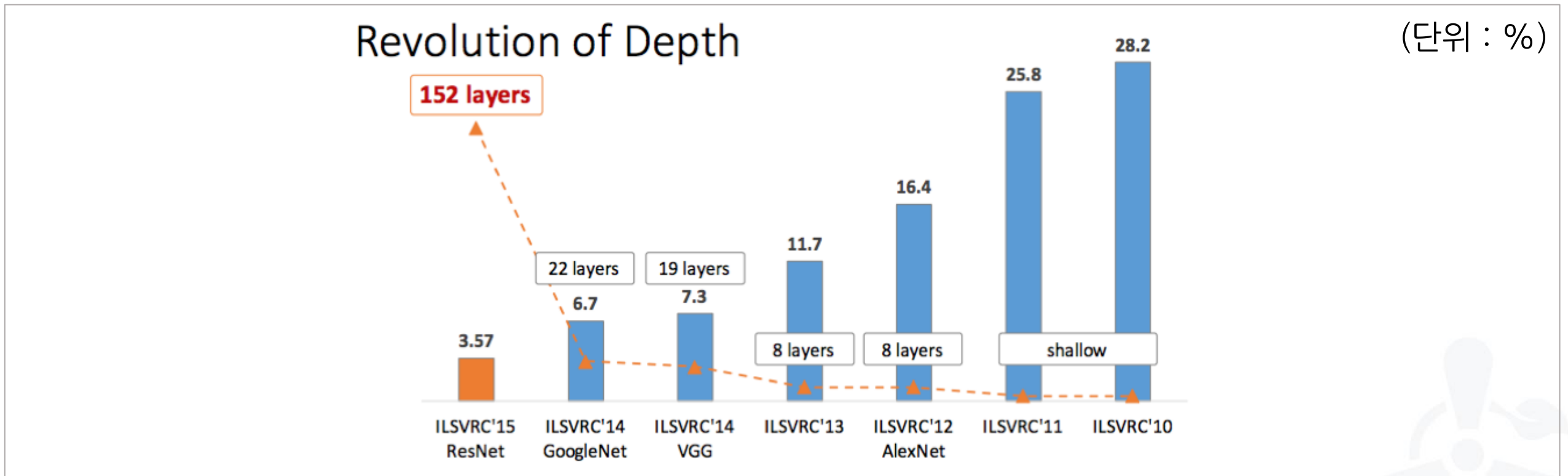
AlexNet(2012)

1. AlexNet(2012)



ILSVRC 이미지넷 역대 우승 알고리즘

- 역대 우승 알고리즘의 이미지 분류 오류율
 - ▶ 층의 깊이가 깊어질수록 오차가 줄어듦
 - ▶ AlexNet 덕분에 합성곱 신경망이 주목을 받게 됨



[출처 : Stanford University CS231n, <http://cs231n.stanford.edu/>]

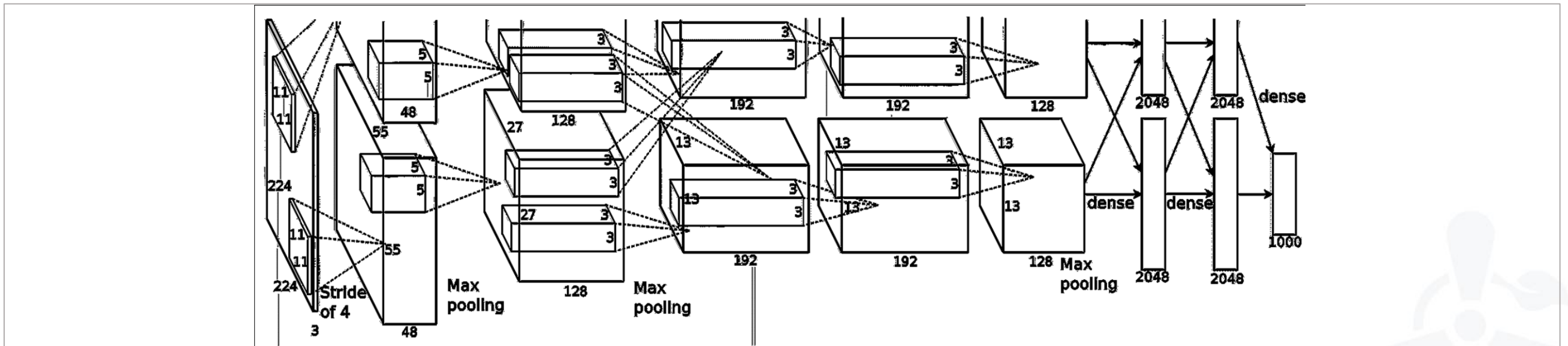


1. AlexNet(2012)



AlexNet 개요

- ILSVRC(ImageNet Large-Scale Visual Recognition Challenge)의 2012년 대회 우승
- 의미있는 성능을 낸 첫 번째 합성곱 신경망 구조
- 당시 GPU가 지금보다 발달하지 않아 2대의 GPU를 사용함
- 활성화 함수로 Relu를 사용함
- 드롭아웃 기법을 사용함



[출처 : <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>]

2

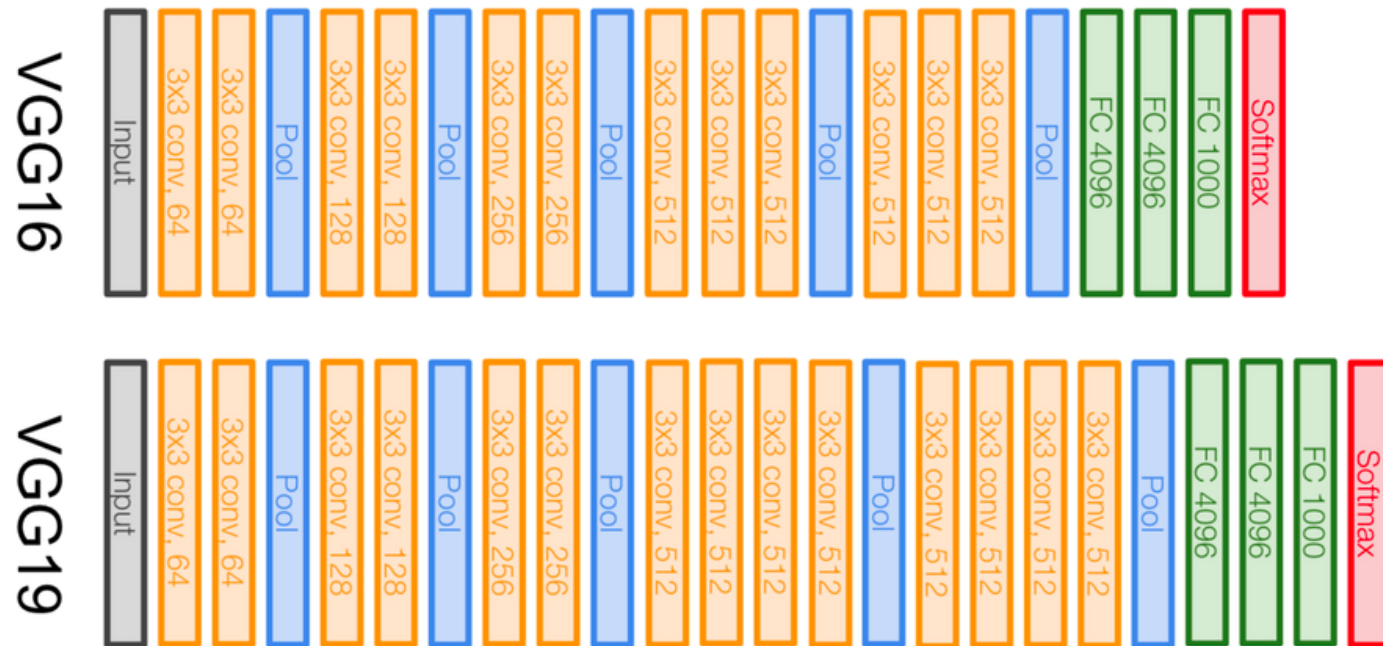
VGGNet(2014)

2. VGGNet(2014)



VGGNet 개요

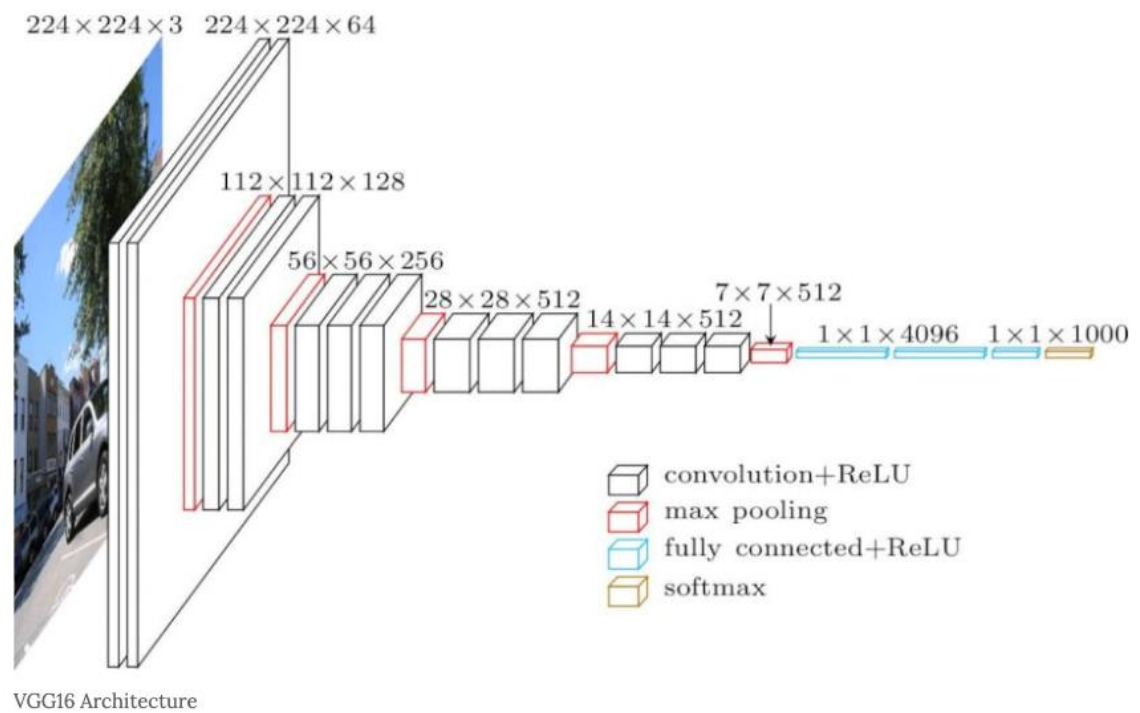
- 옥스포드 대학의 연구팀 VGG에 의해 개발된 모델
- ILSVRC(ImageNet Large-Scale Visual Recognition Challenge)의 2014년 대회 준우승
- VGGNet을 시작으로 네트워크 깊이가 엄청 깊어지게 됨



2. VGGNet(2014)

VGGNet-16 구조

- 13 Convolution Layers + 3 Fully-connected Layers
- 3x3 Convolution Filters
- Stride: 1 & Padding: 1
- 2x2 Max Pooling (Stride : 2)
- ReLU



ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					



실습

예제

문제 상황

VGGNet으로 이미지 분류하기

실습 코드

```
from tensorflow.keras.applications.  
vgg19 import VGG19
```

예
시
화
면

```
from tensorflow.keras.applications.vgg19 import VGG19  
vgg = VGG19(weights = "imagenet", include_top=False, input_shape=(img_size, img_size, 3))
```

```
Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/vgg19/vgg19\_weights\_tf\_dim\_orderin  
g\_tf\_kernels\_notop.h5  
80142336/80134624 [=====] - 18s 0us/step
```

```
for layer in vgg.layers:  
    layer.trainable = False
```

```
from tensorflow.keras import Sequential  
from tensorflow.keras.layers import Flatten, Dense  
model = Sequential()  
model.add(vgg)  
model.add(Flatten())  
model.add(Dense(5, activation="softmax"))
```

3

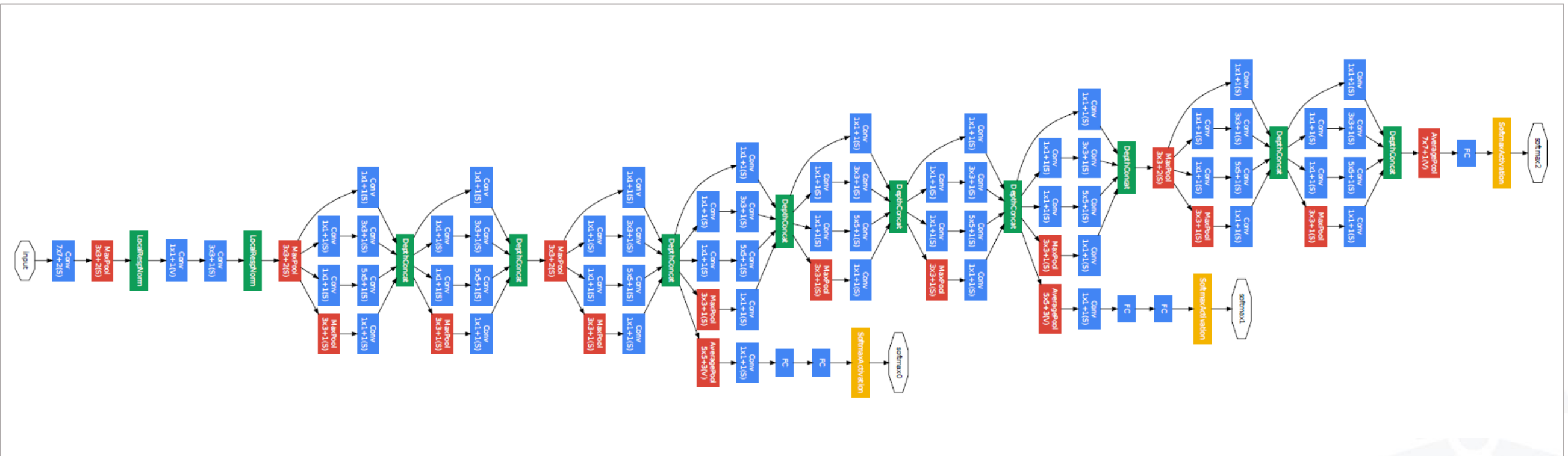
GoogLeNet(2014)

3. GoogLeNet(2014)



GoogLeNet 개요

- 합성곱 신경망의 아버지 ‘안르쿰’이 구글에서 개발함
- VGGNet 보다 구조가 복잡해 널리 쓰이진 않았지만, 구조 면에서 주목받음
- 하나의 계층에서도 다양한 종류의 필터, 풀링을 도입함



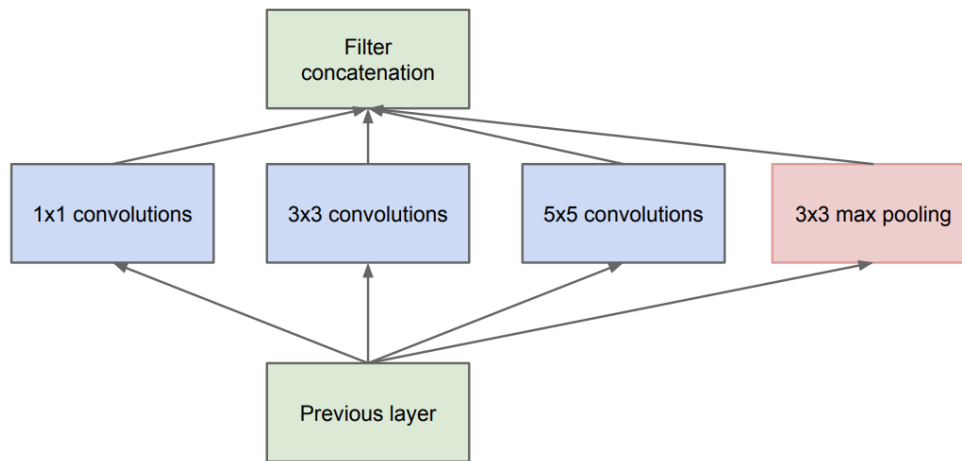
[출처 : <https://arxiv.org/pdf/1409.4842.pdf>]

인셉션 모듈

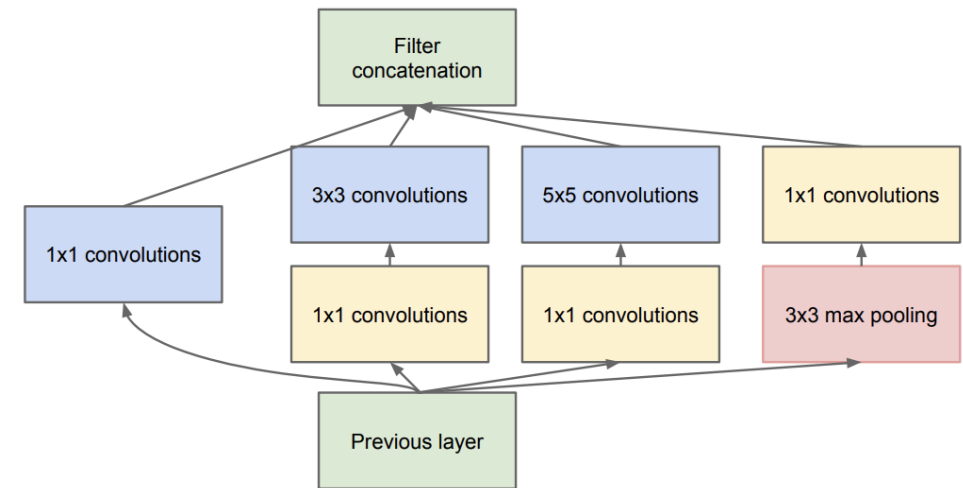
인셉션 모듈

차원(채널) 축소를 위한 1x1 합성곱 계층 아이디어

- ▶ 갈림길이 생김으로써 조금 더 다양한 특성을 모델이 찾을 수 있게 함
- ▶ 인공지능이 사람이 보는 것과 비슷한 구조로 볼 수 있게 함
- ▶ VGGNet 보다 신경망이 깊어졌음에도, 사용된 파라미터는 절반 이하임



(a) Inception module, naïve version



(b) Inception module with dimension reductions

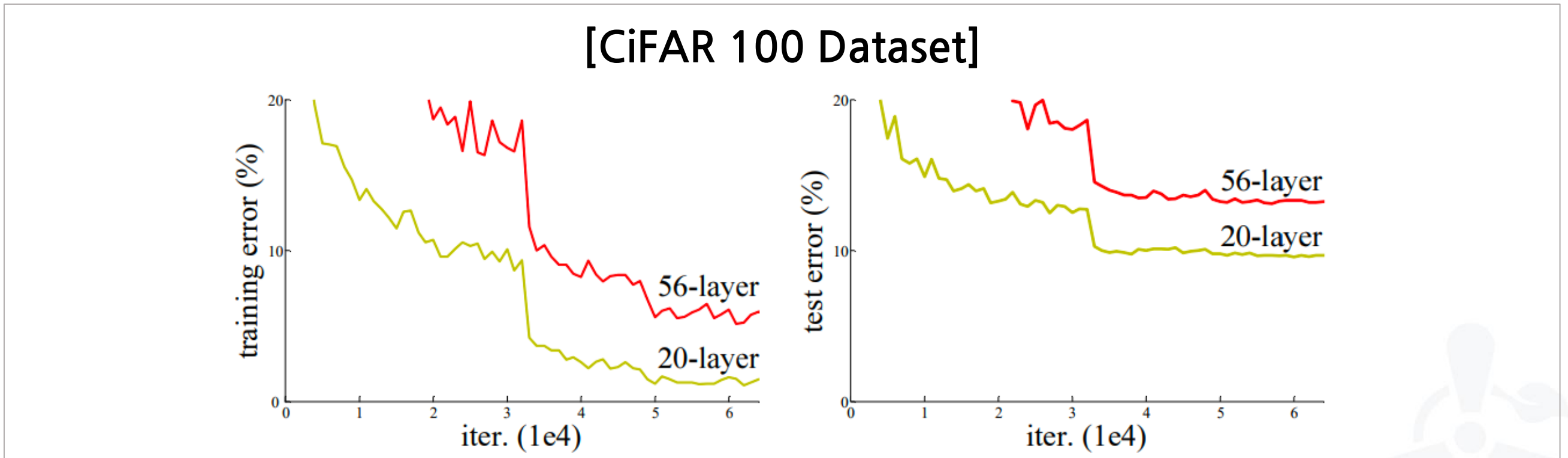
4

ResNet(2015)



Degradation 문제

- 20층의 신경망이 56층의 신경망보다 낮은 오류율
- 층이 깊어질수록 기울기가 점점 사라져 학습이 잘 되지 않는 문제
- 기울기 소실에 의한 Degradation 문제를 해결하기 위해 제안



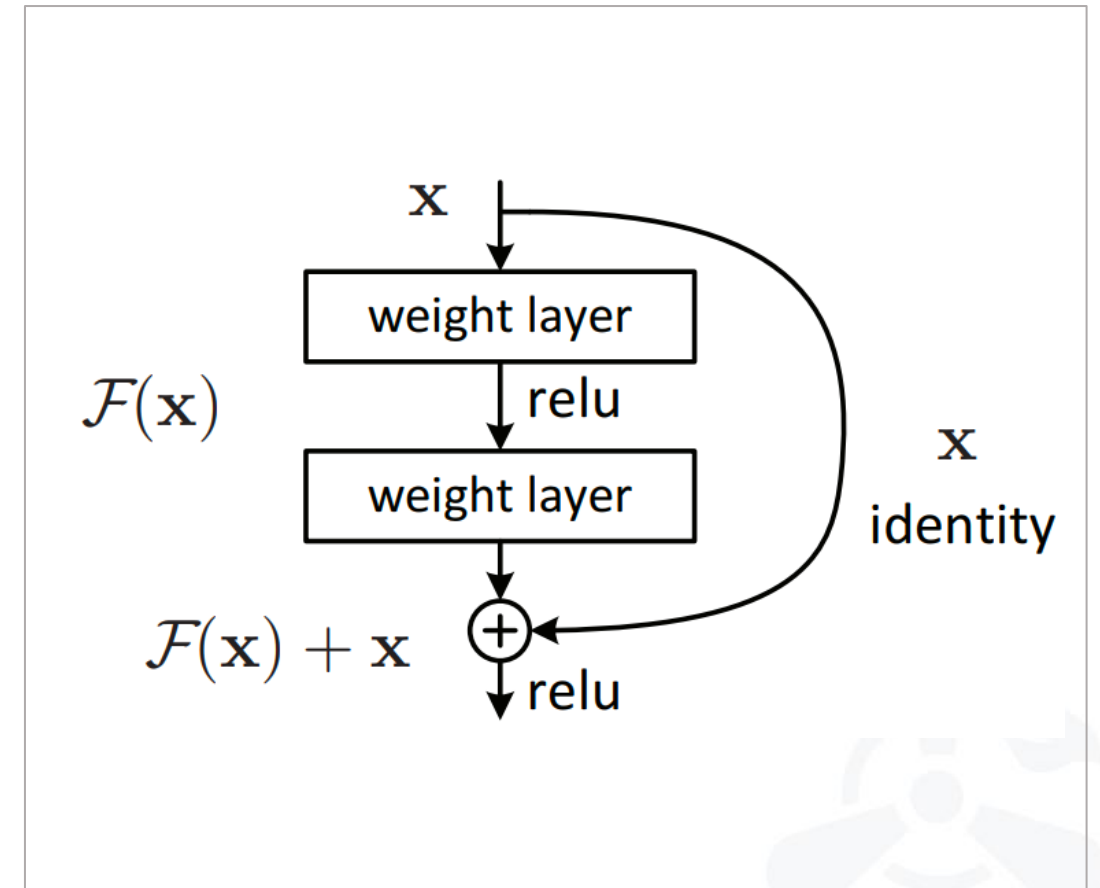


Residual Block

Residual Block

기울기가 잘 학습될 수 있도록 일종의 지름길을 만들어주는 방법

- ▶ $y=F(x)+x$ 를 다시 쓰면 $F(x)=y-x$ 로 표현할 수 있음
- ▶ 입력과 출력 간의 차이를 학습하도록 설계됨
- ▶ Identity Short Connection은 하나 이상의 층을 건너 뛸 수 있게 함
- ▶ 추가적인 파라미터도 필요하지 않고 복잡한 곱셈 연산도 필요하지 않는 것이 장점임



Degradation Problem

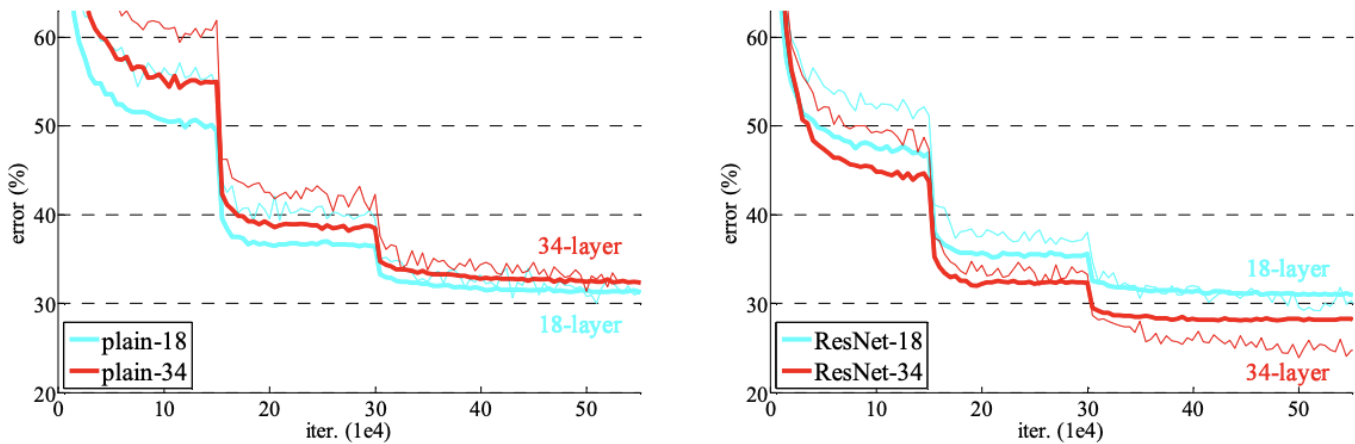
PlainNet

기울기 소실 문제 때문에 18-layer의 성능이 34-layer 보다 뛰어남

ResNet

기울기 소실 문제가 Skip Connection에 의해 해결되어 34-layer의 성능이 더 뛰어남

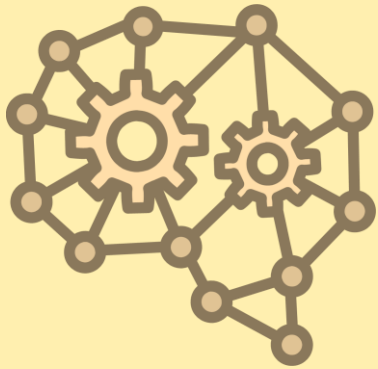
[PlainNet vs ImageNet]



	plain	ResNet
18 layers	27.94	27.88
34 layers	28.54	25.03

Table 2. Top-1 error (% , 10-crop testing) on ImageNet validation. Here the ResNets have no extra parameter compared to their plain counterparts. Fig. 4 shows the training procedures.

[출처 : <https://arxiv.org/pdf/1512.03385.pdf>]



KEY POINT

- ▶ AlexNet(2012)
 - 의미있는 성능을 낸 첫 번째 합성곱 신경망 구조
 - 활성화 함수로 Relu를 사용, 드롭아웃 기법 사용
- ▶ VGGNet(2014)
 - 네트워크 깊이가 엄청 깊어지게 됨
 - 컨볼루션 필터커널의 사이즈는 가장 작은 3x3으로 고정
- ▶ GoogLeNet(2014)
 - VGGNet 보다 구조가 복잡해 널리 쓰이진 않았지만, 구조 면에서 주목받음
 - 하나의 계층에서도 다양한 종류의 필터, 풀링을 도입
- ▶ ResNet(2015)
 - 기울기 소실에 의한 Degradation 문제를 해결하기 위해 제안
 - Residual Block으로 기울기가 잘 학습될 수 있도록 일종의 지름길을 만들어줌



Python을 활용한 이미지 분석

7. 다양한 합성곱 신경망 모델

“이번 시간을 모두 마치셨습니다.
수고하셨습니다.”