

AI Expert 프로그램 실습

7/10 Basic of CNN, Image Classification, Style Transfer

Overview

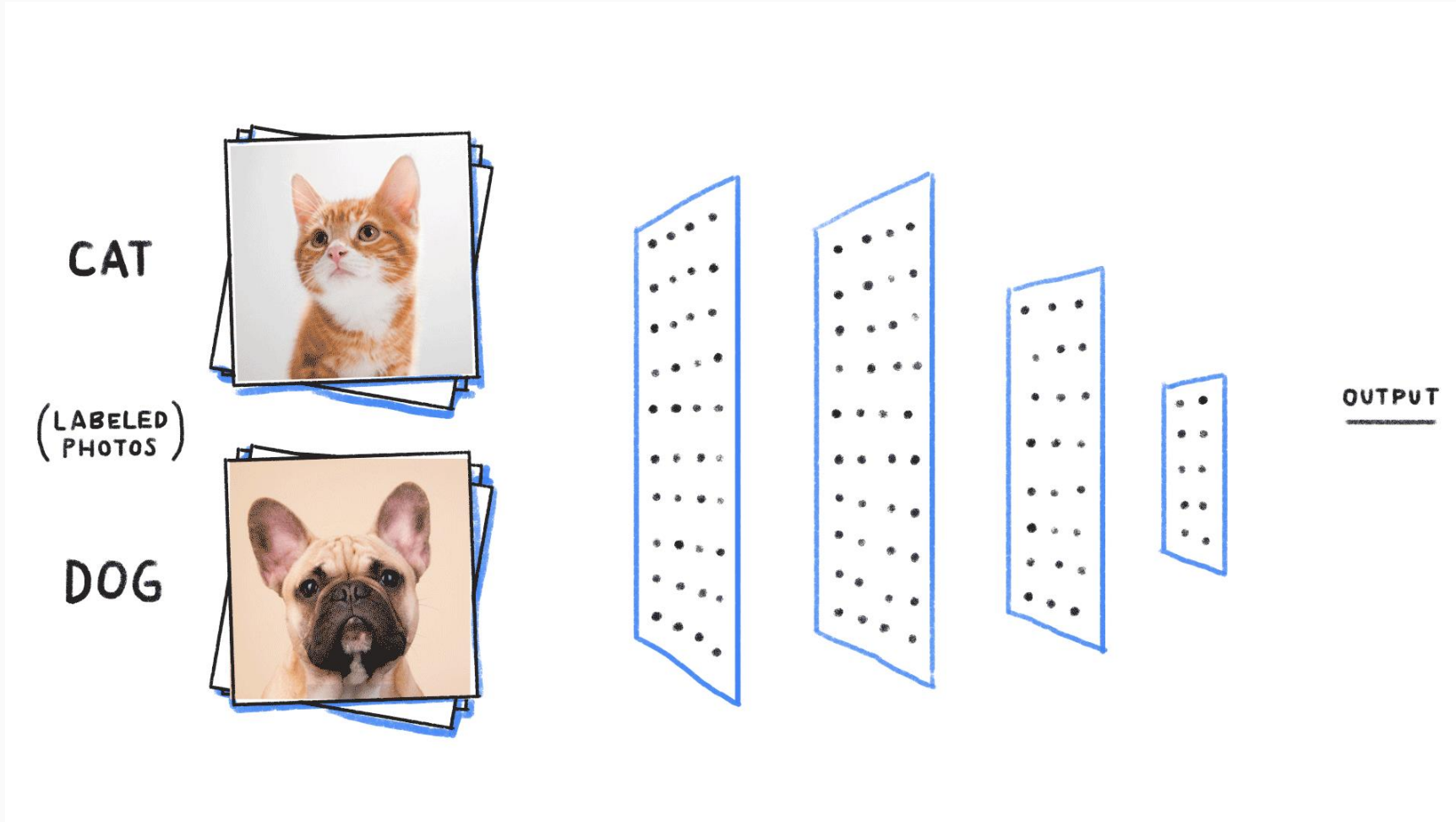
This tutorial is three-fold as follows:

1. Basics of CNN
2. Image Classification
3. Style Transfer

* 10 minutes break between each part.

Part 2. Image Classification

What is Image Classification ?

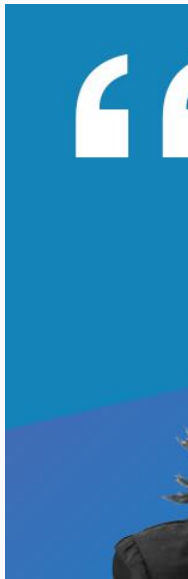


History of Neural Networks

Launch of ImageNet (2009)

ILSVRC: ImageNet Large Scale Visual Recognition Competition

ImageNet Challenge

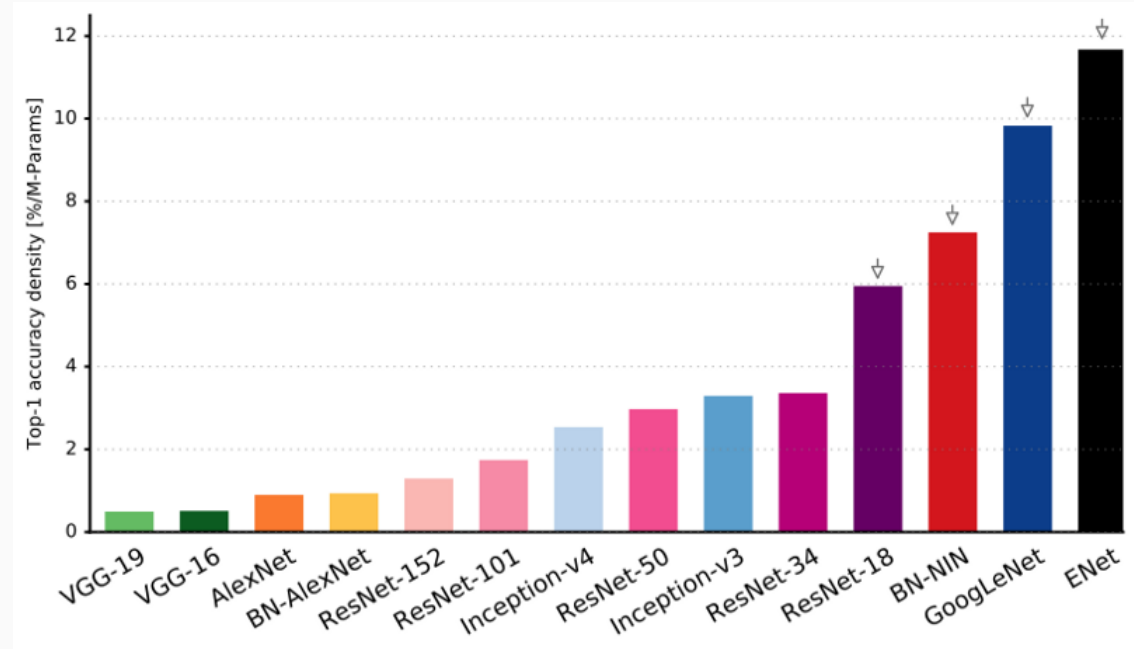
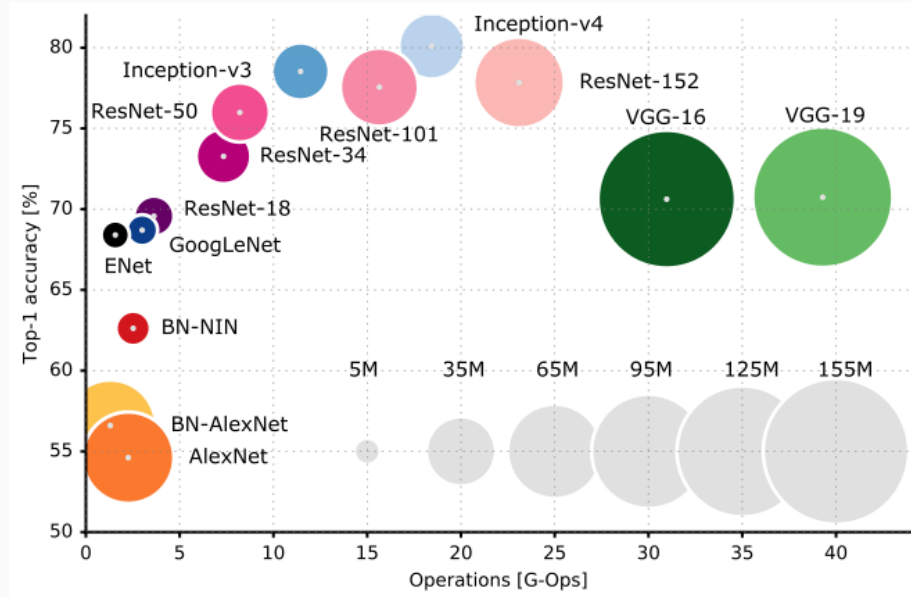


- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



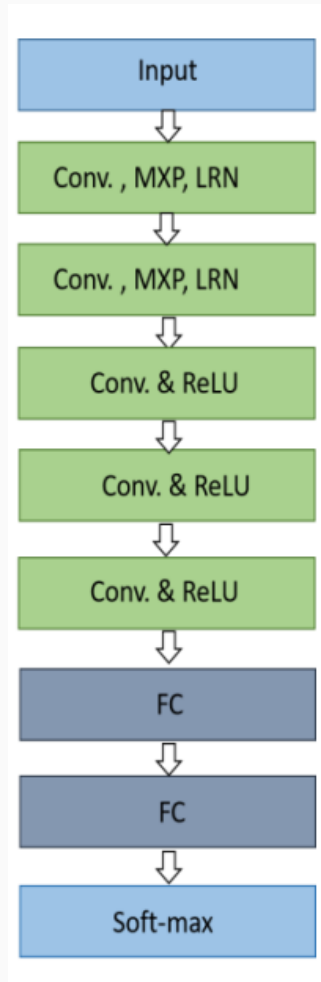
History of Neural Networks

Top-1 one-crop accuracy versus amount of operations required for a single forward pass



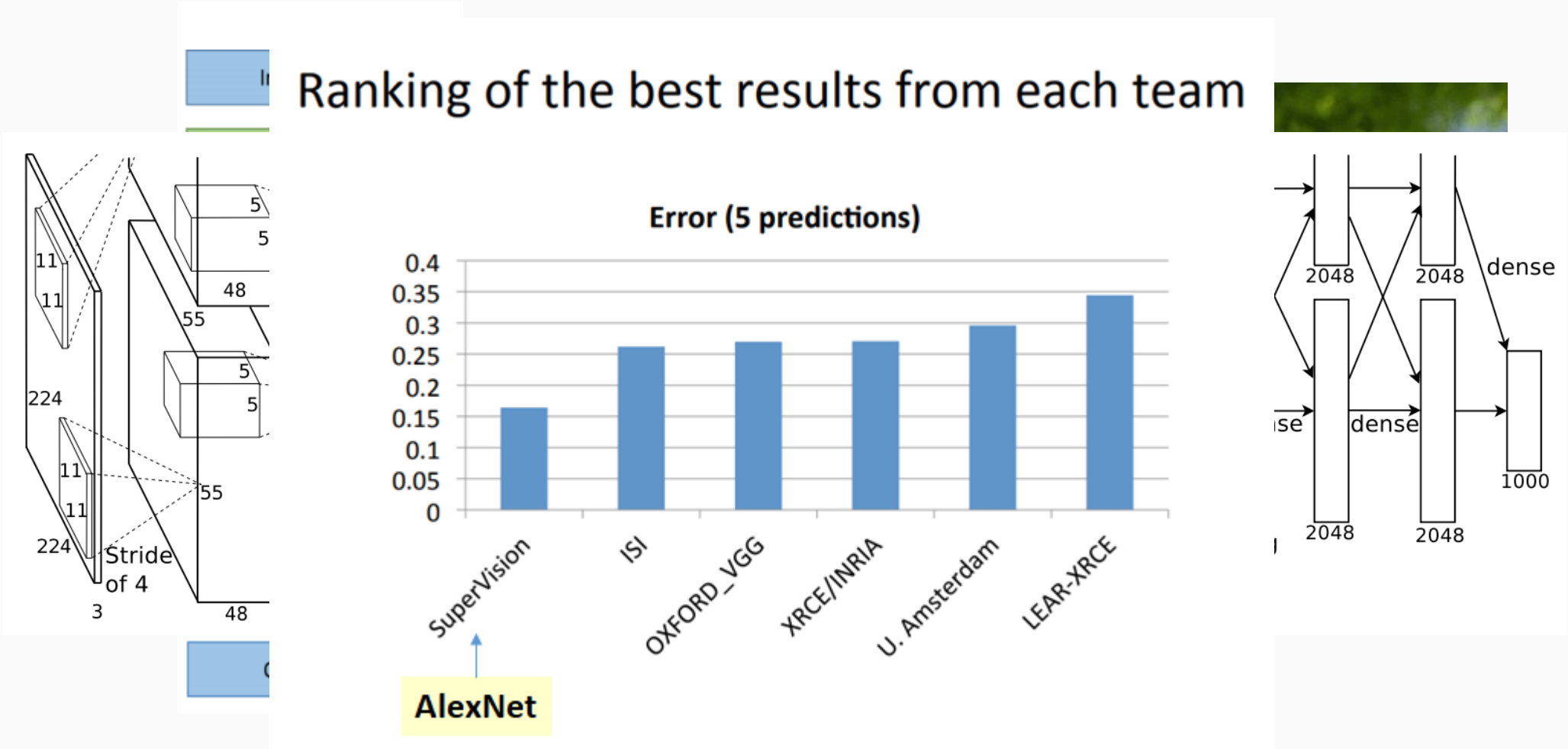
History of Neural Networks

LeNet (1998), LeNet5 (2010)

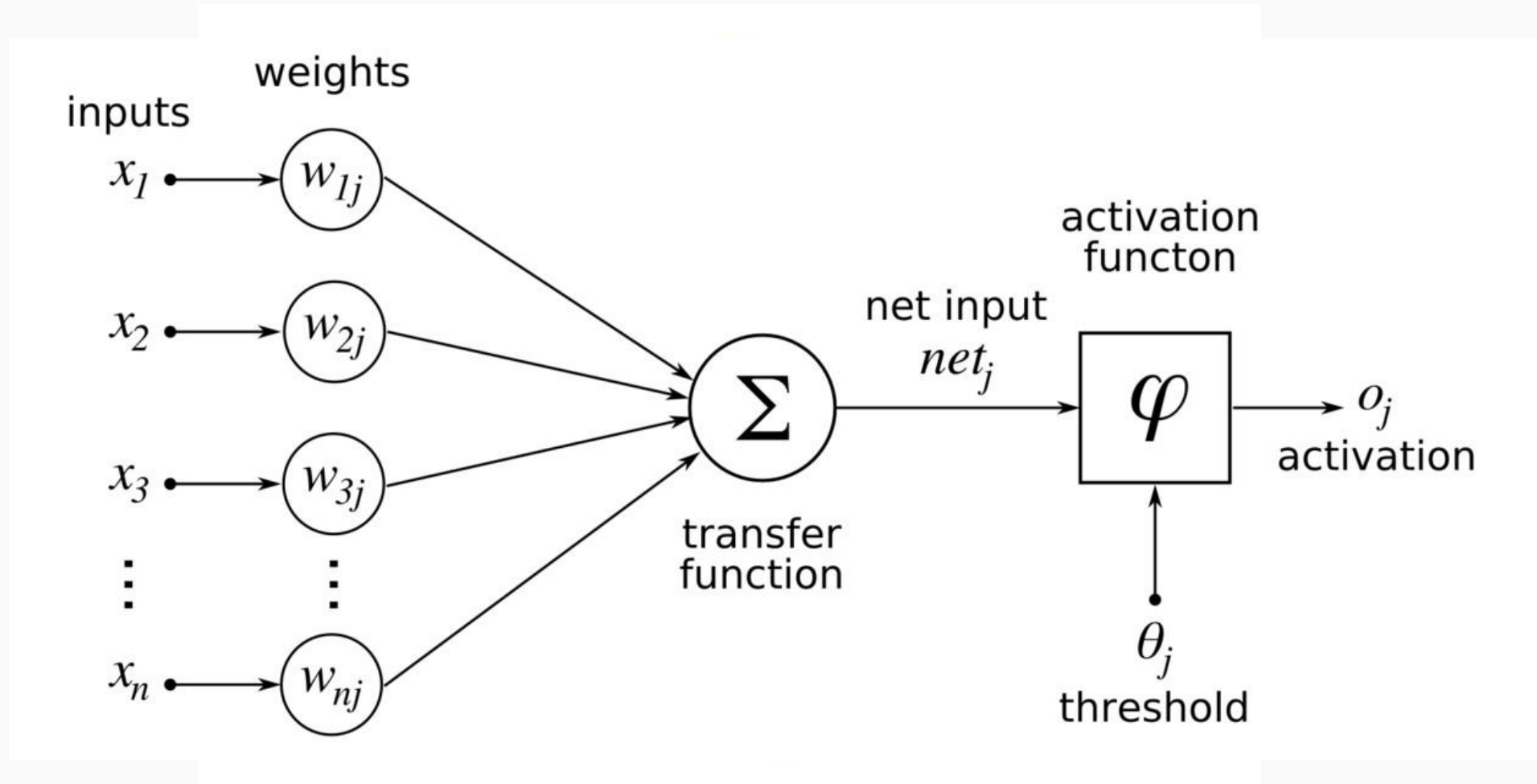


History of Neural Networks

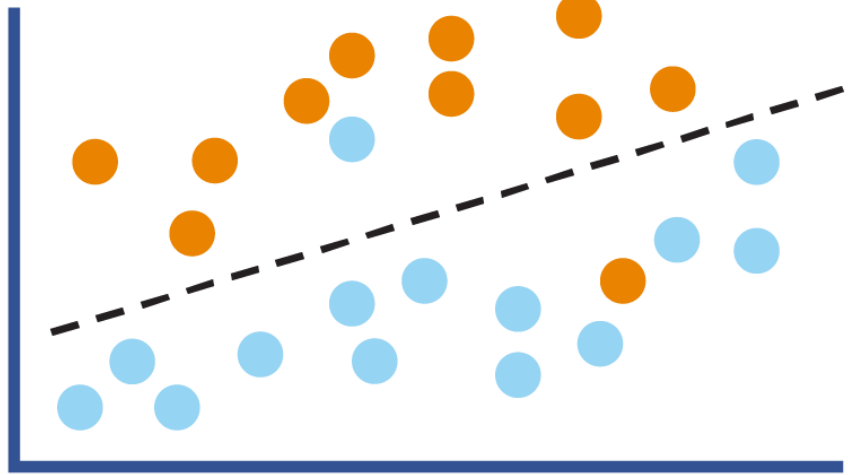
AlexNet (2012)



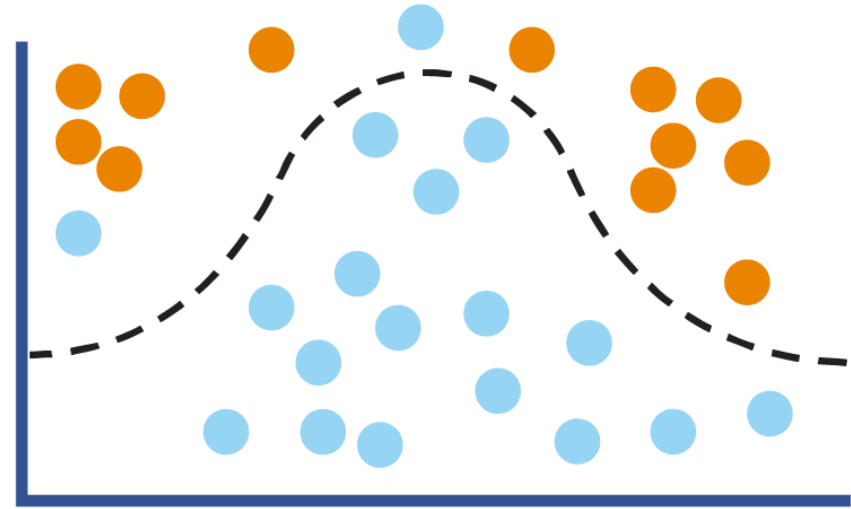
AlexNet (2012)



AlexNet (2012)



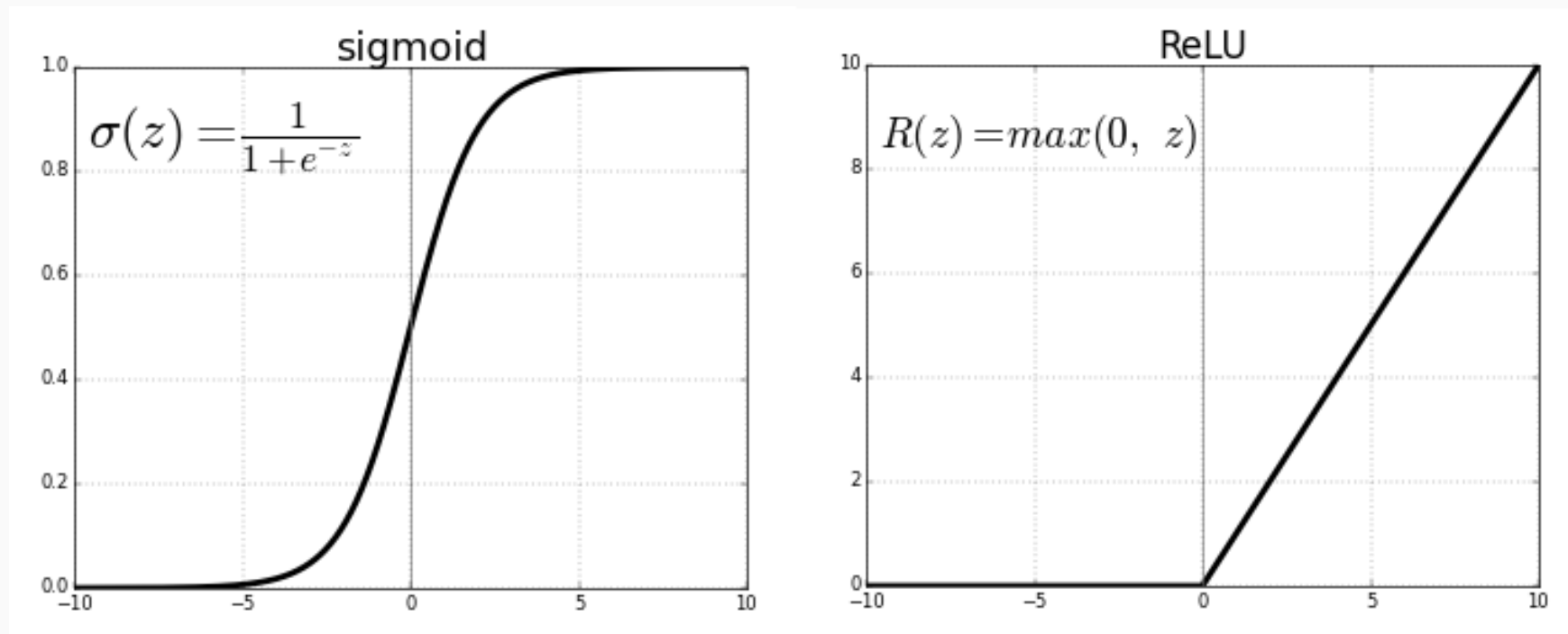
linear separation



non-linear separation

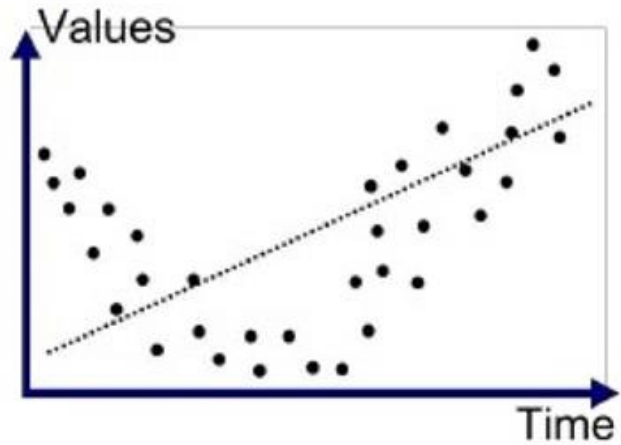
History of Neural Networks

AlexNet (2012)

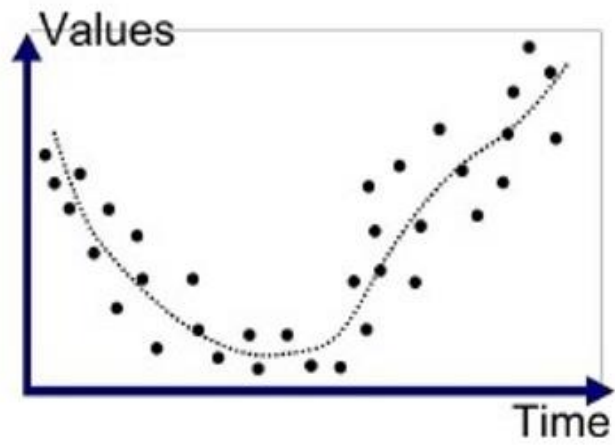


History of Neural Networks

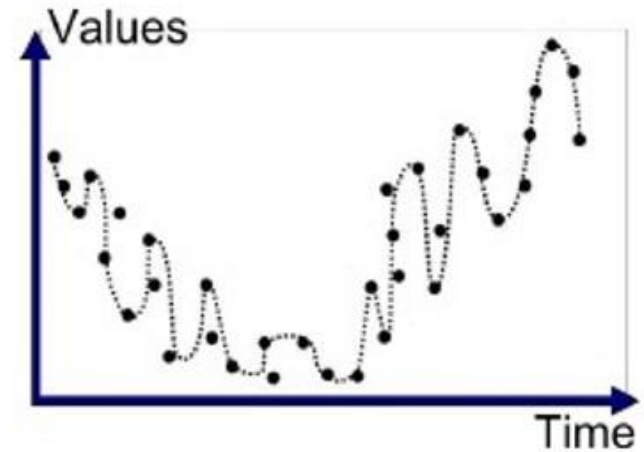
AlexNet (2012)



Underfitted

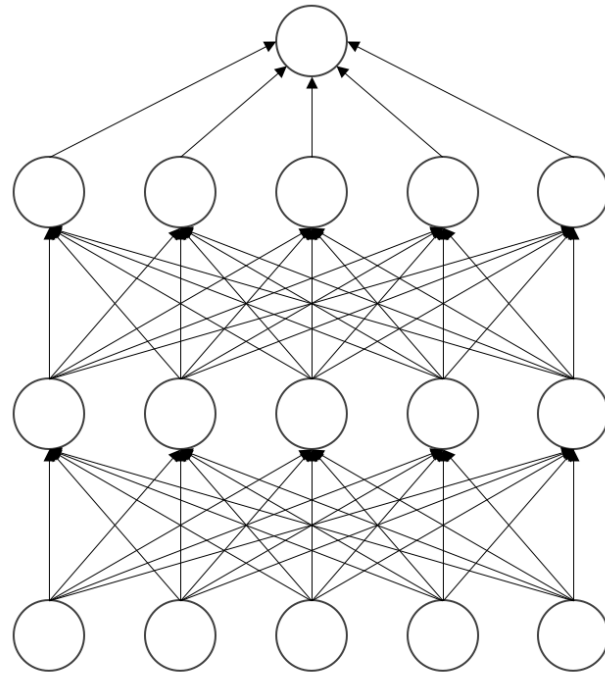


Good Fit/Robust

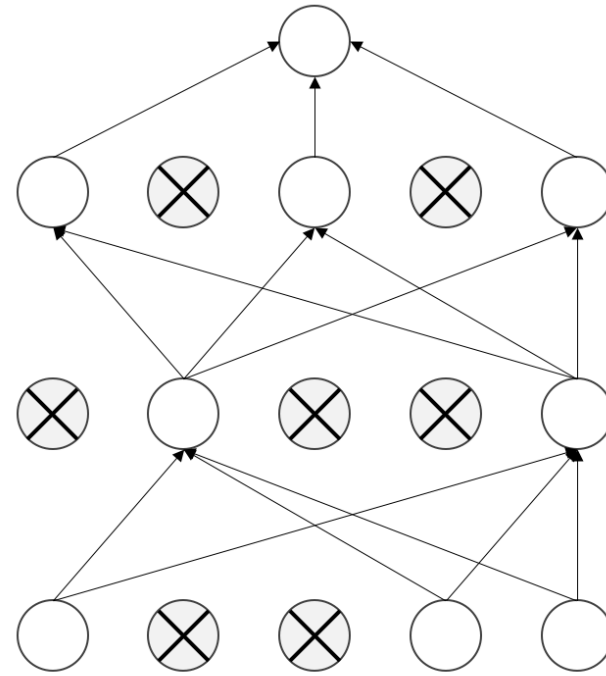


Overfitted

AlexNet (2012)



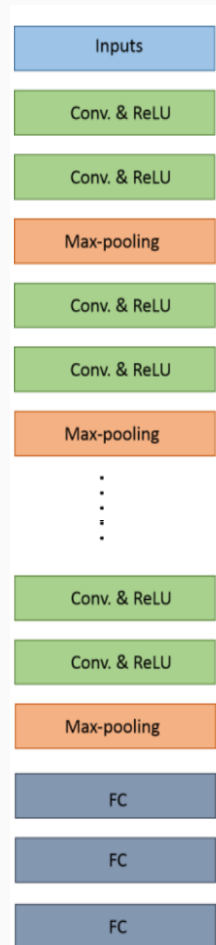
Standard Neural Net



After applying dropout

History of Neural Networks

VGGNet (2014)



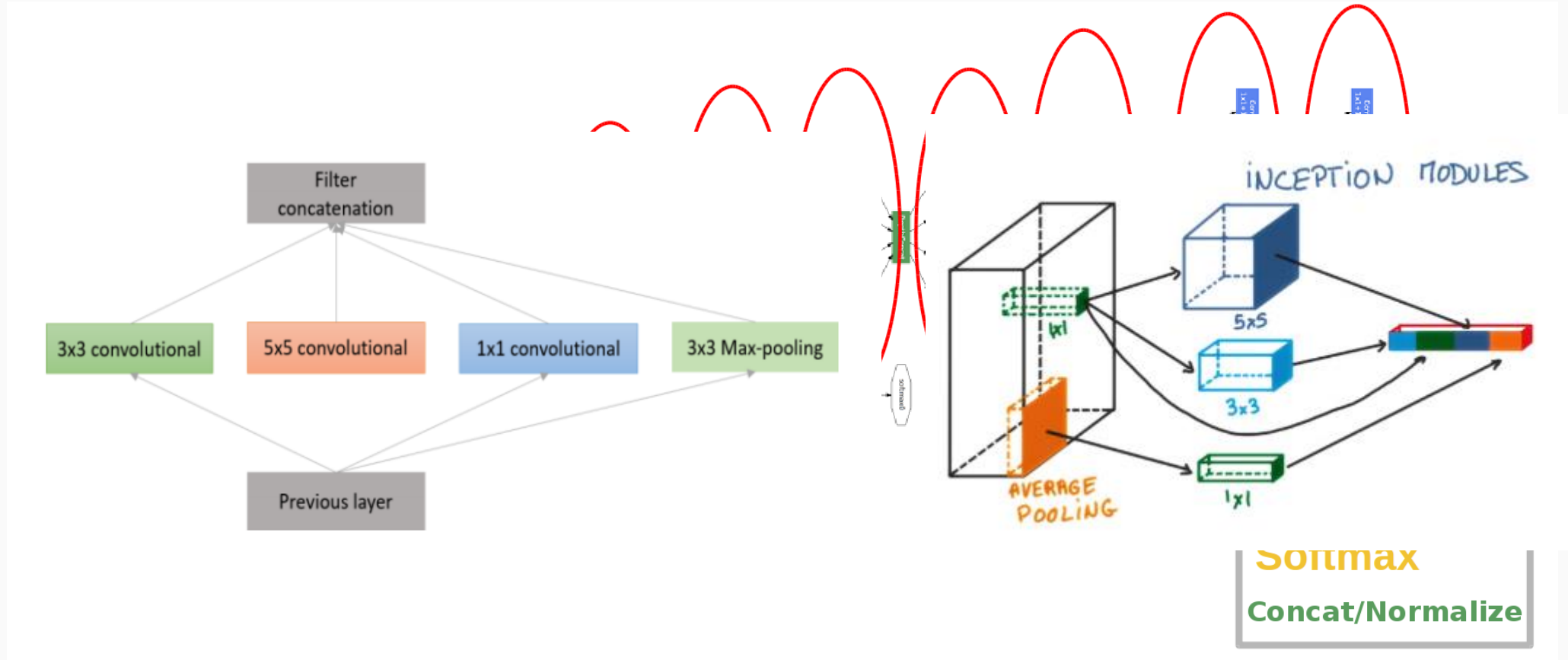
ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	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					

Table 2: Number of parameters (in millions).

Network	A,A-LRN	B	C	D	E
Number of parameters	133	133	134	138	144

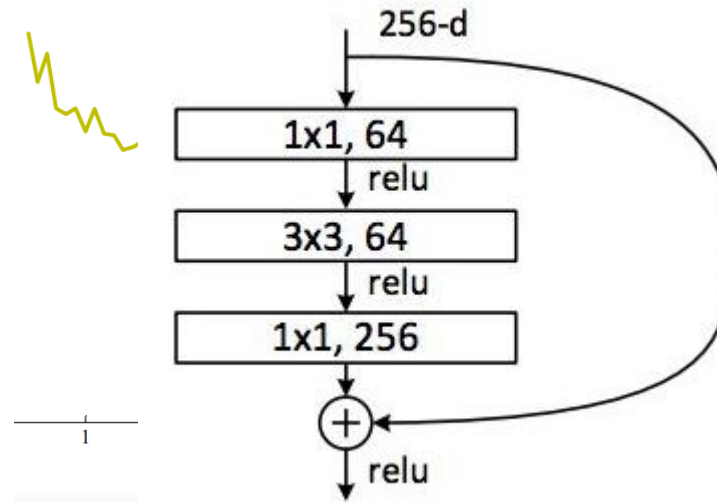
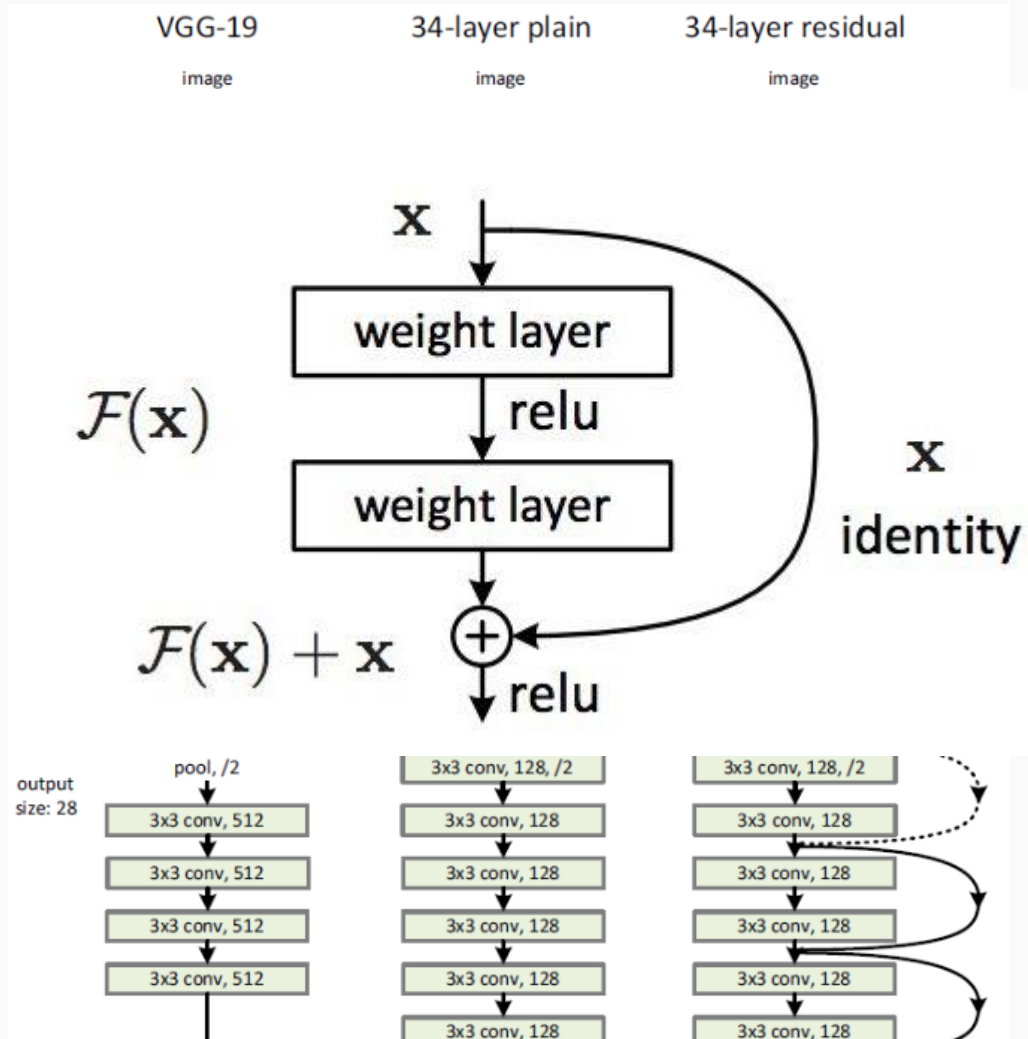
History of Neural Networks

GoogLeNet (2014)



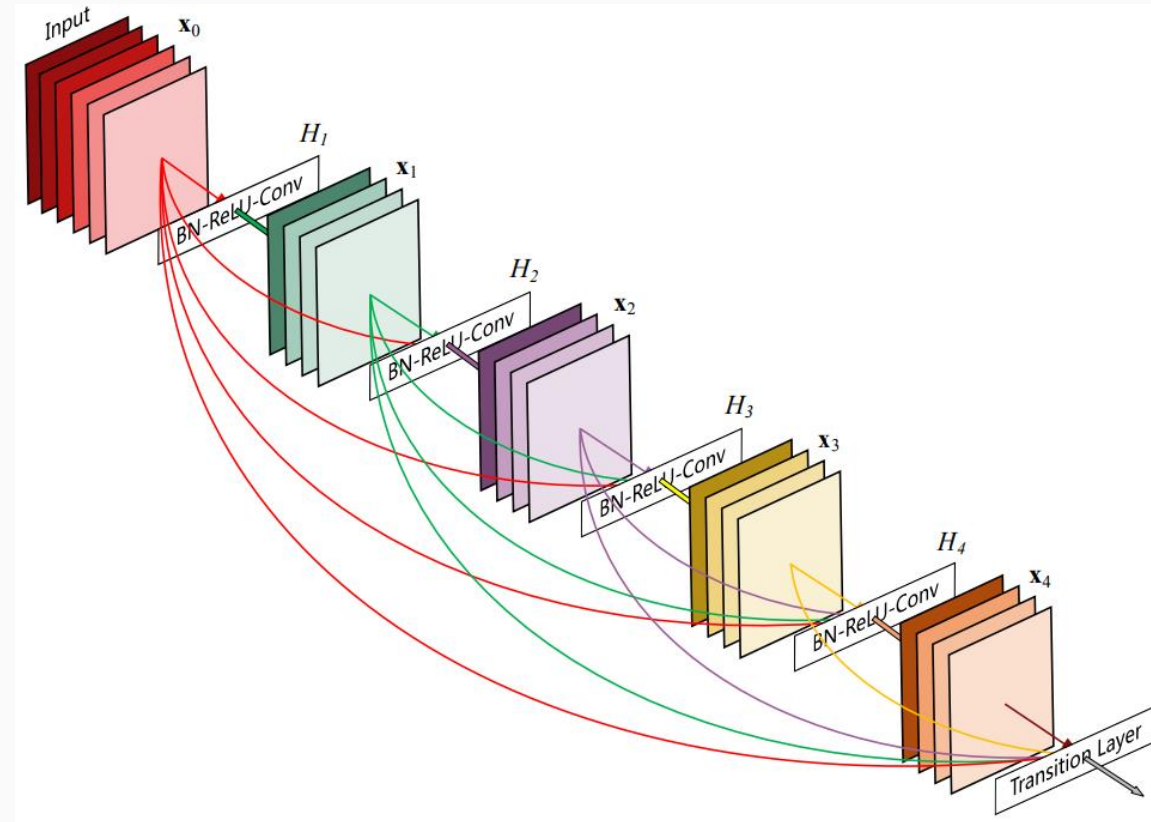
History of Neural Networks

ResNet (2015)



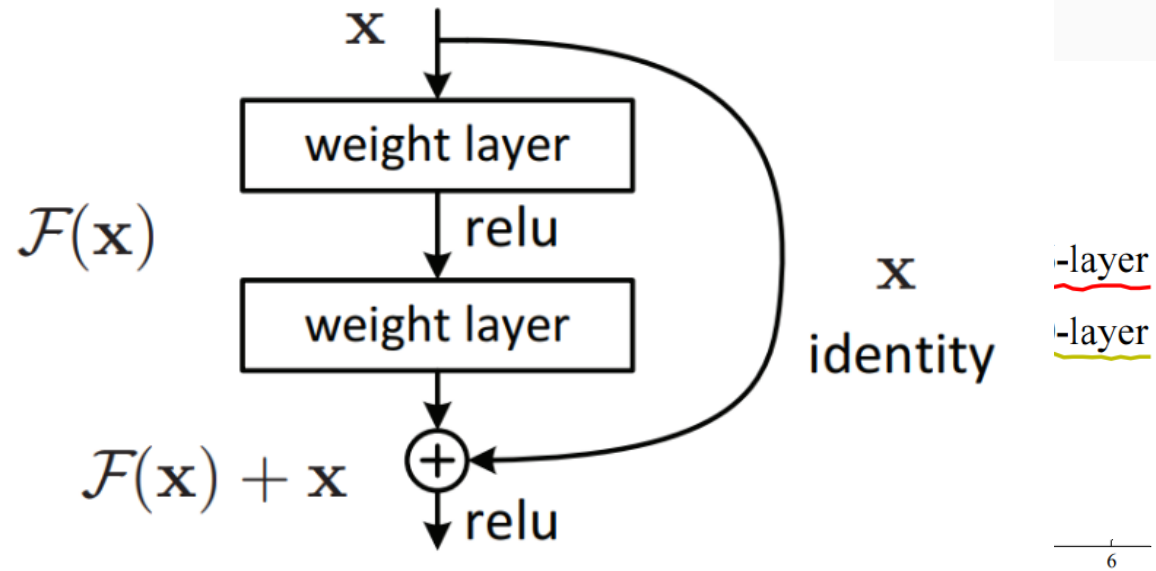
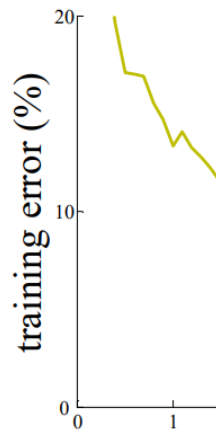
History of Neural Networks

DenseNet (2016)



ResNet

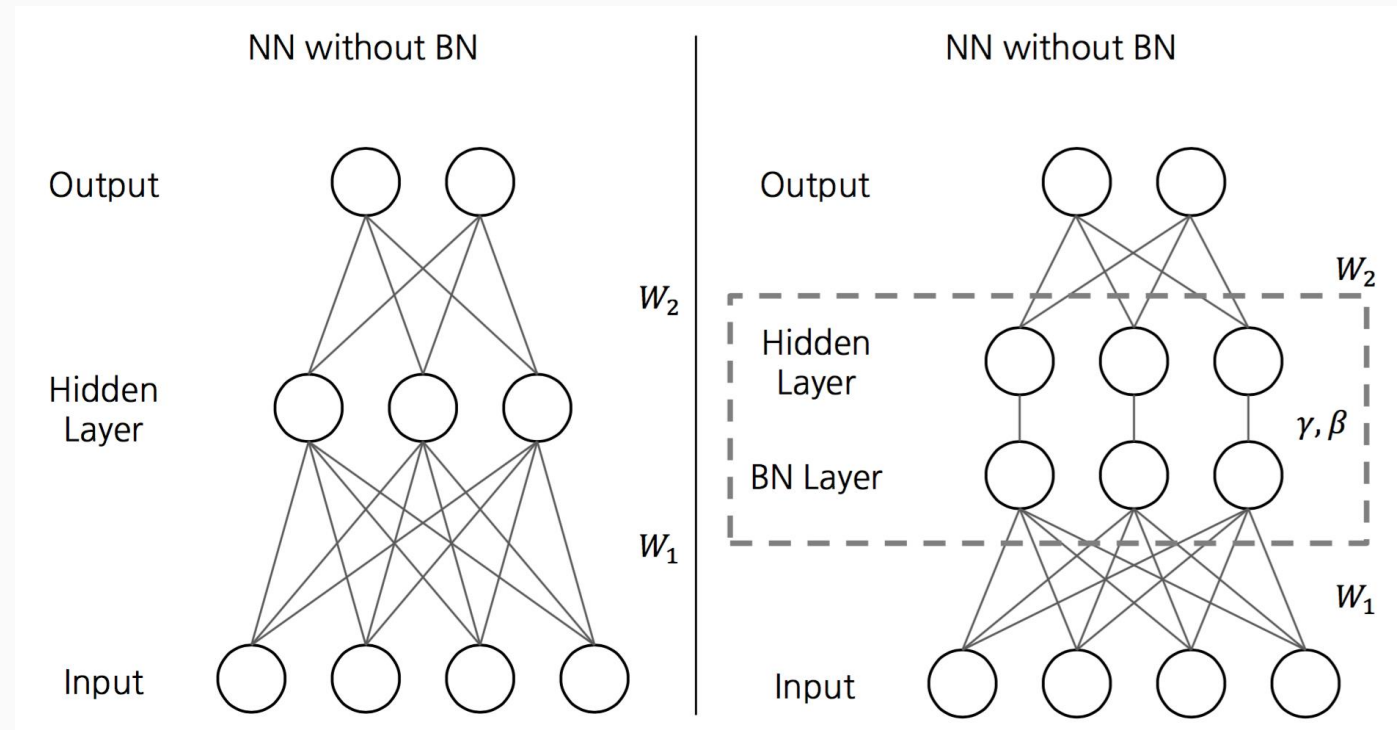
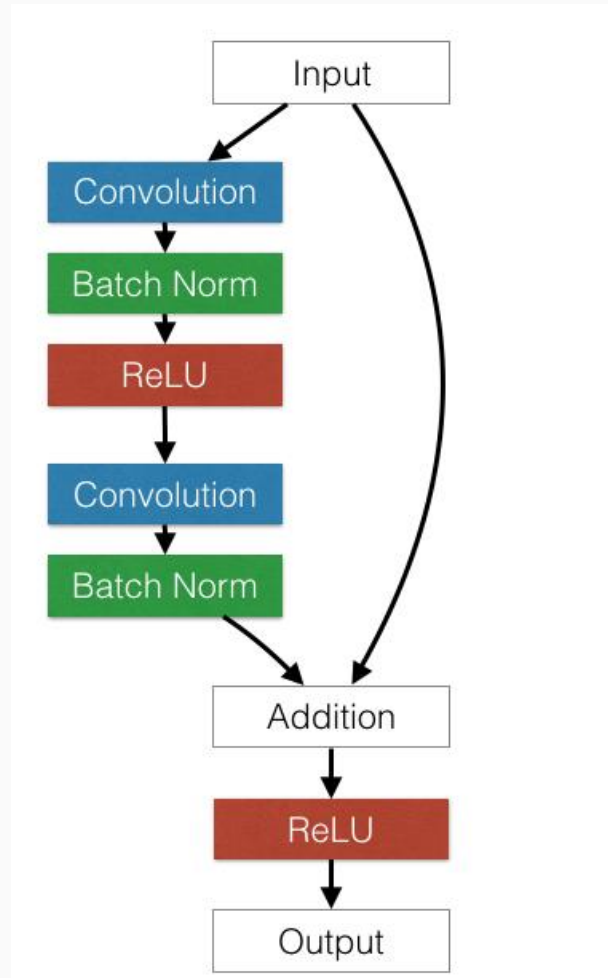
- 2015 ILSVRC 1st ranked / error – 3.6%
- Human error – 5 ~10 %
- Vanishing gradient problem
- Residual block



$$x_l = \mathcal{F}(x_{l-1}) + x_{l-1}$$

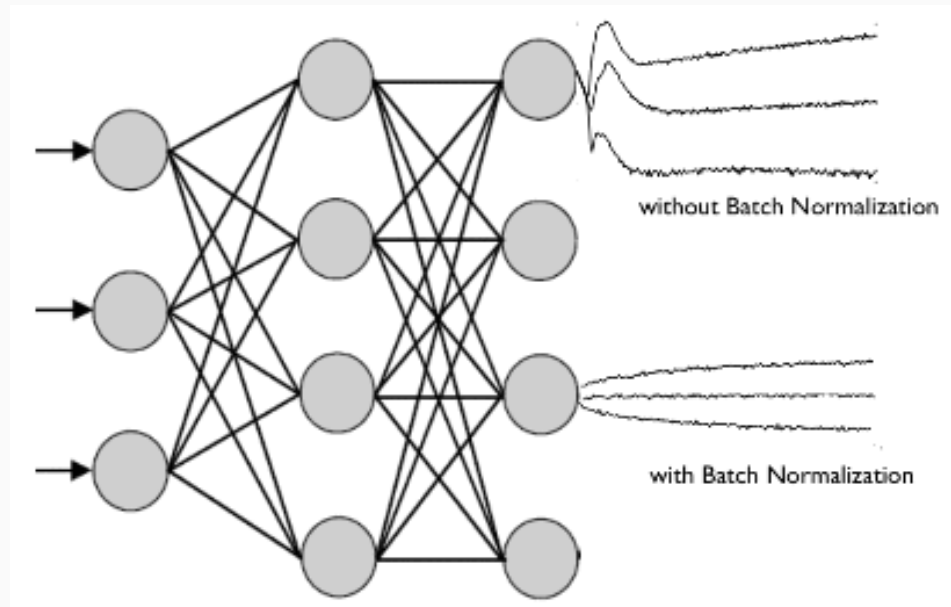
ResNet

- Batch Normalization



ResNet

- Batch Normalization



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

Parameters to be learned: γ, β

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

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

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

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

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

Algorithm 1: Batch Normalizing Transform, applied to activation x over a mini-batch.

ResNet

- Batch Normalization

