# Engineering Analytics & Machine Learning (ECSE202) Seminar 8


# Introduction to Deep Learening

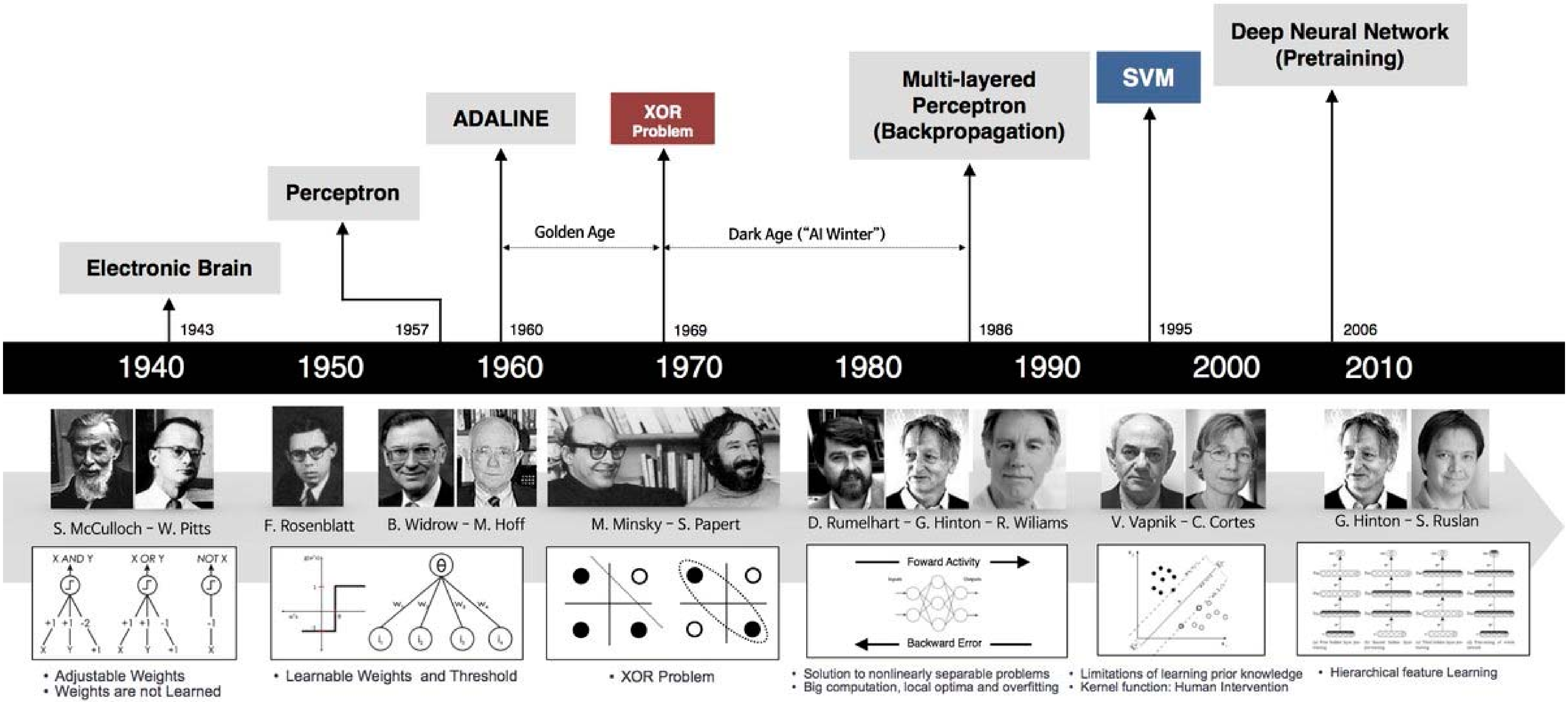
AY2018/2019 OCT SEMESTER



School of Engineering
TEMASEK POLYTECHNIC

# Brief history

# How we get here?




Souorce:https://twitter.com/deeplearn007/status/945381750494191616

# TED Talk by Fei Fei Li
# Deep Learning: ImageNet


https://www.youtube.com/watch?v=40riCqvRoMs

**Why you should listen**
Using algorithms built on machine learning methods such as neural network models, the Stanford Artificial Intelligence Lab led by Fei-Fei Li has created software capable of recognizing scenes in still photographs -- and accurately describe them using natural language.

Li's work with neural networks and computer vision (with Stanford's Vision Lab) marks a significant step forward for AI research, and could lead to applications ranging from more intuitive image searches to robots able to make autonomous decisions in unfamiliar situations.

# IMAGENET Large Scale Visual Recognition Challenge

## Year 2010

NEC-UIUC

Dense descriptor grid: HOG, LBP

↓

Coding: local coordinate, super-vector

↓

Pooling, SPM

↓

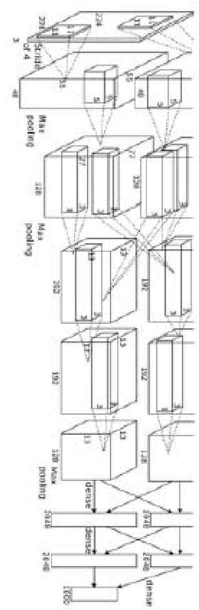
Linear SVM

[Lin CVPR 2011]



## Year 2012

SuperVision

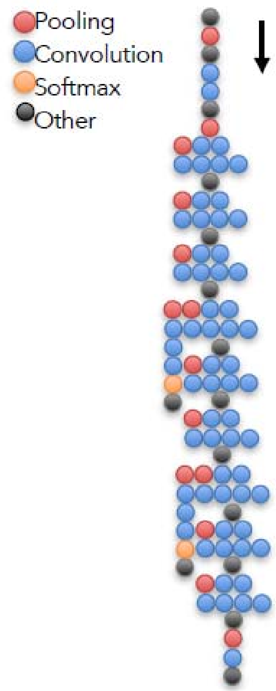


[Krizhevsky NIPS 2012]



## Year 2014

GoogLeNet




[Szegedy arxiv 2014]

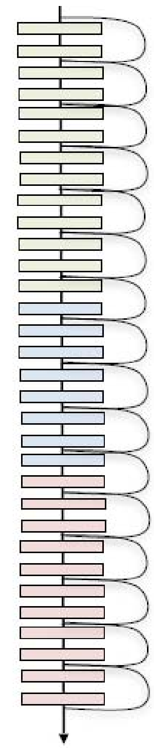
VGG

Image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
maxpool
conv-512
conv-512
maxpool
conv-512
conv-512
maxpool
fc-4096
fc-4096
fc-1000
softmax

[Simonyan arxiv 2014]

## Year 2015

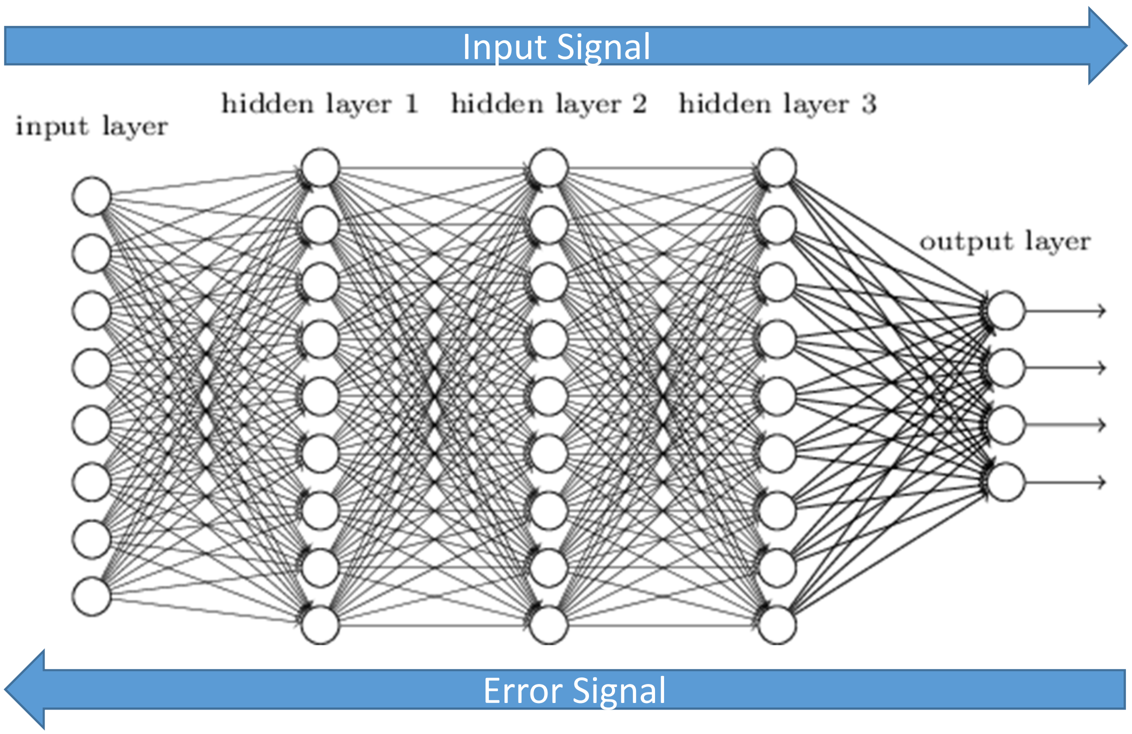
MSRA



[He ICCV 2015]

# Backpropagation

Feedfoward

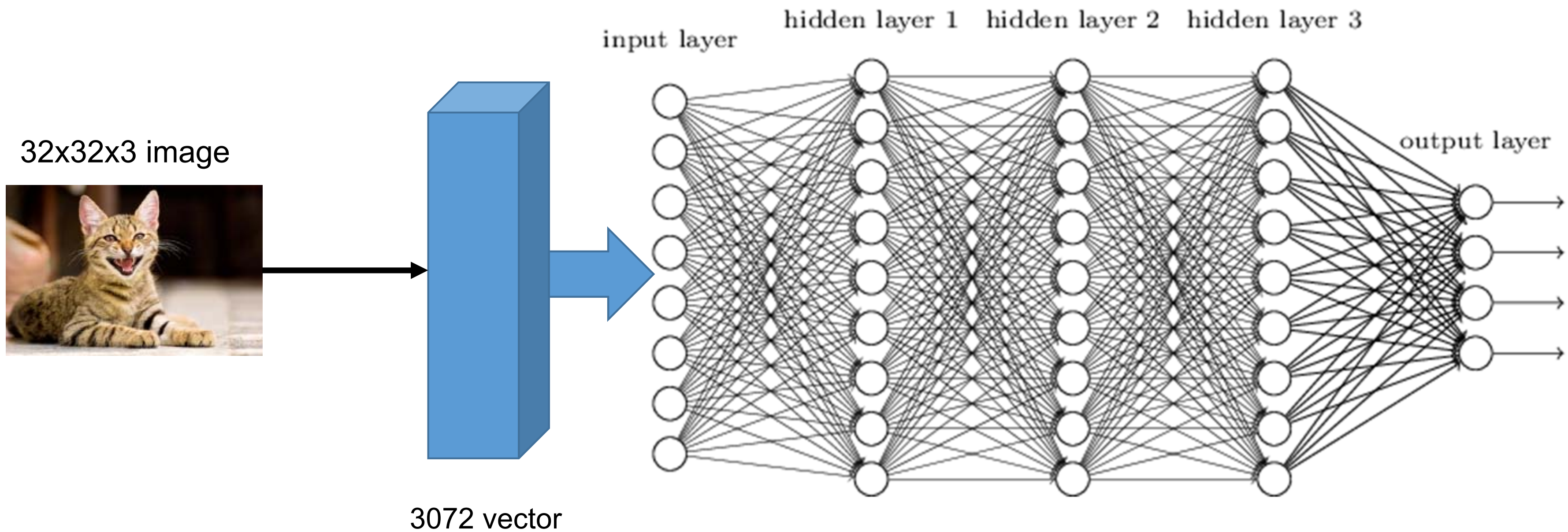



Backpropagation

Backpropagation was invented in the 1970s as a general optimization method for performing automatic differentiation of complex nested functions. However, it wasn't until 1986, with the publishing of a paper by Rumelhart, Hinton, and Williams, titled "Learning Representations by Back-Propagating Errors," that the importance of the algorithm was appreciated by the machine learning community at large.

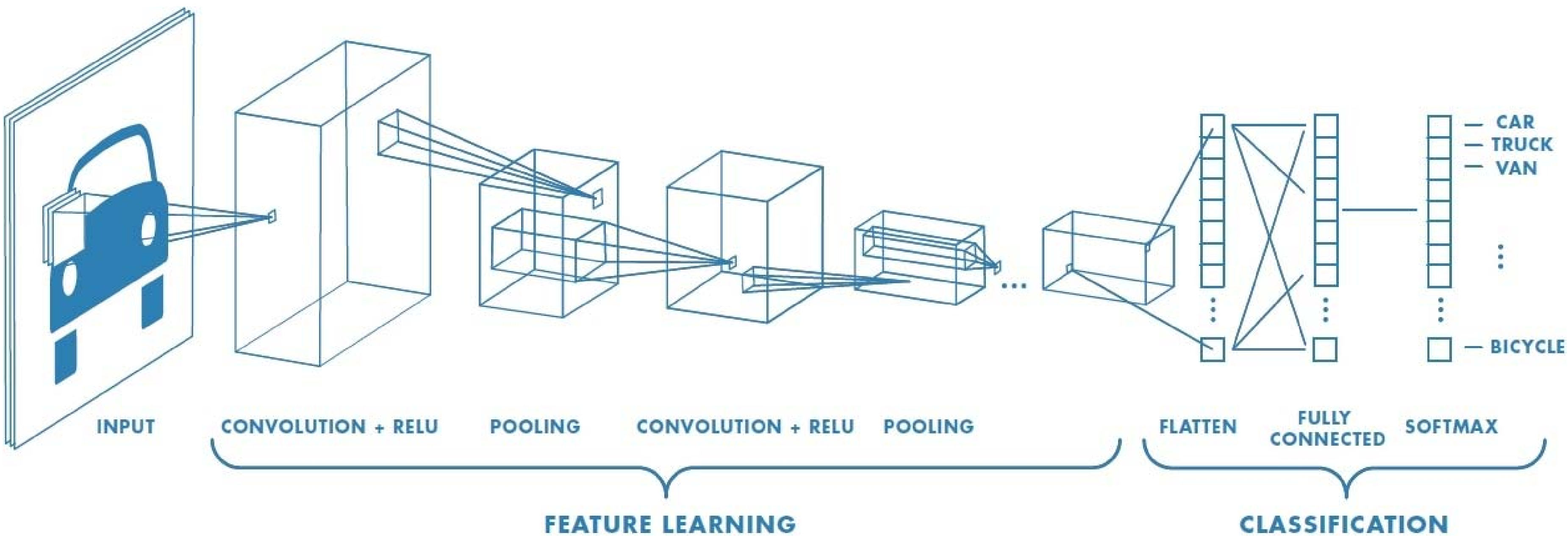
Source: https://brilliant.org/wiki/backpropagation/

# Image Classification with Fully connected Neural Networks




Spatial information of the input image is lost in the process

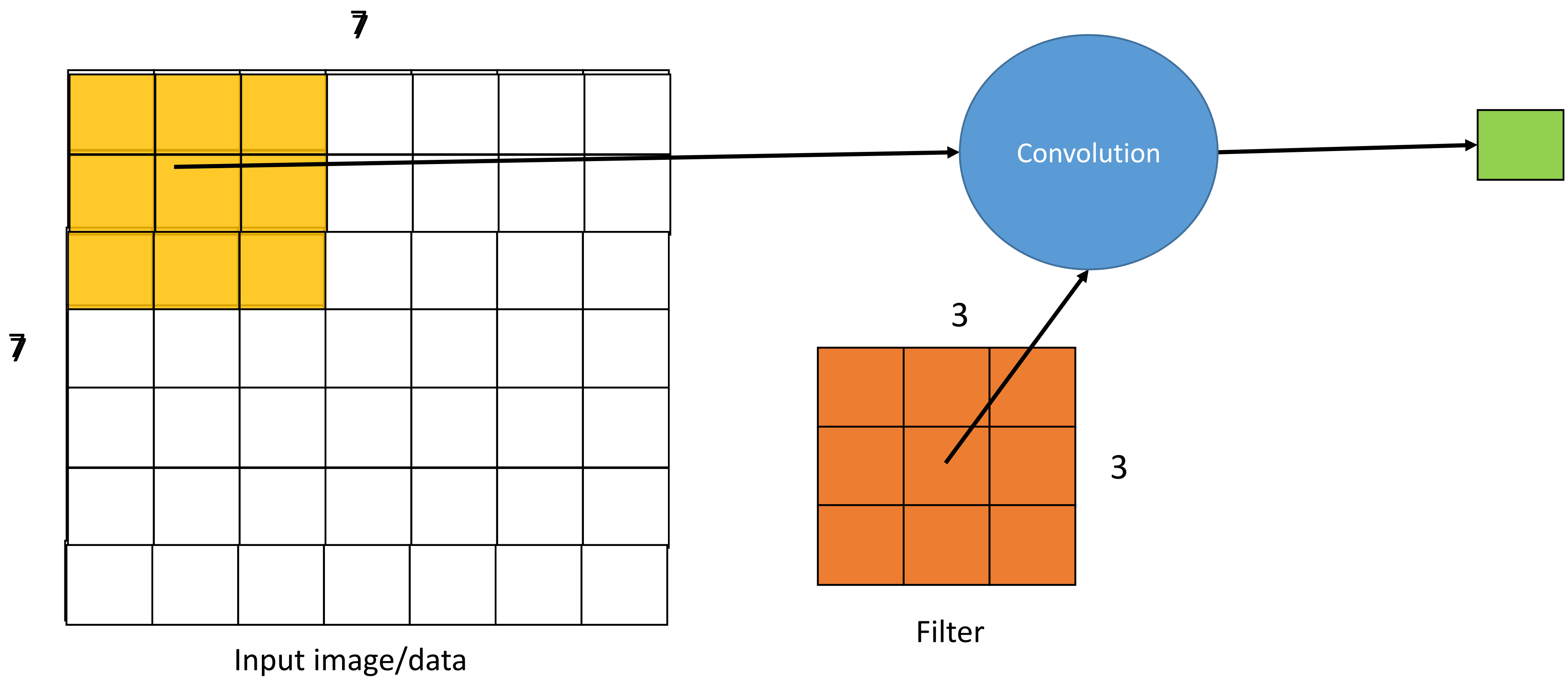
# Convolution Neural Network




Source: https://www.mathworks.com/discovery/convolutional-neural-network.html

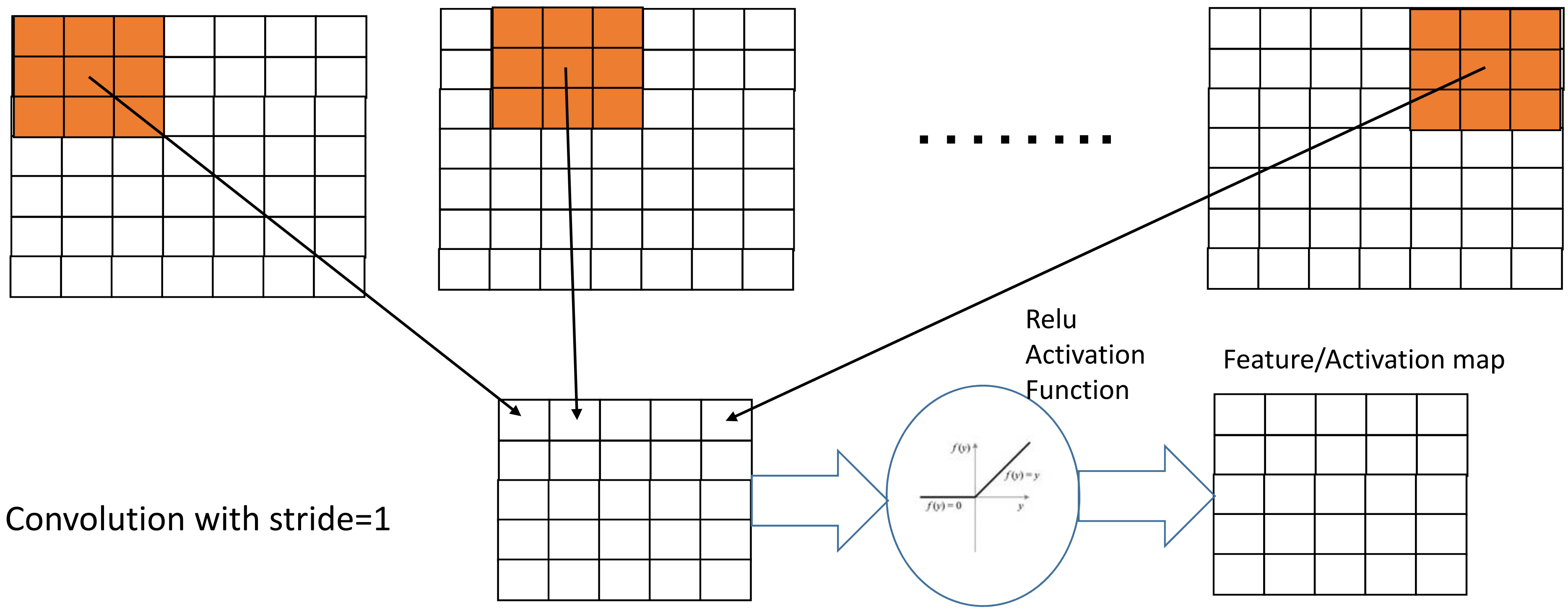
Spatial information about an image is lost in the process

# Convolution Neural Network



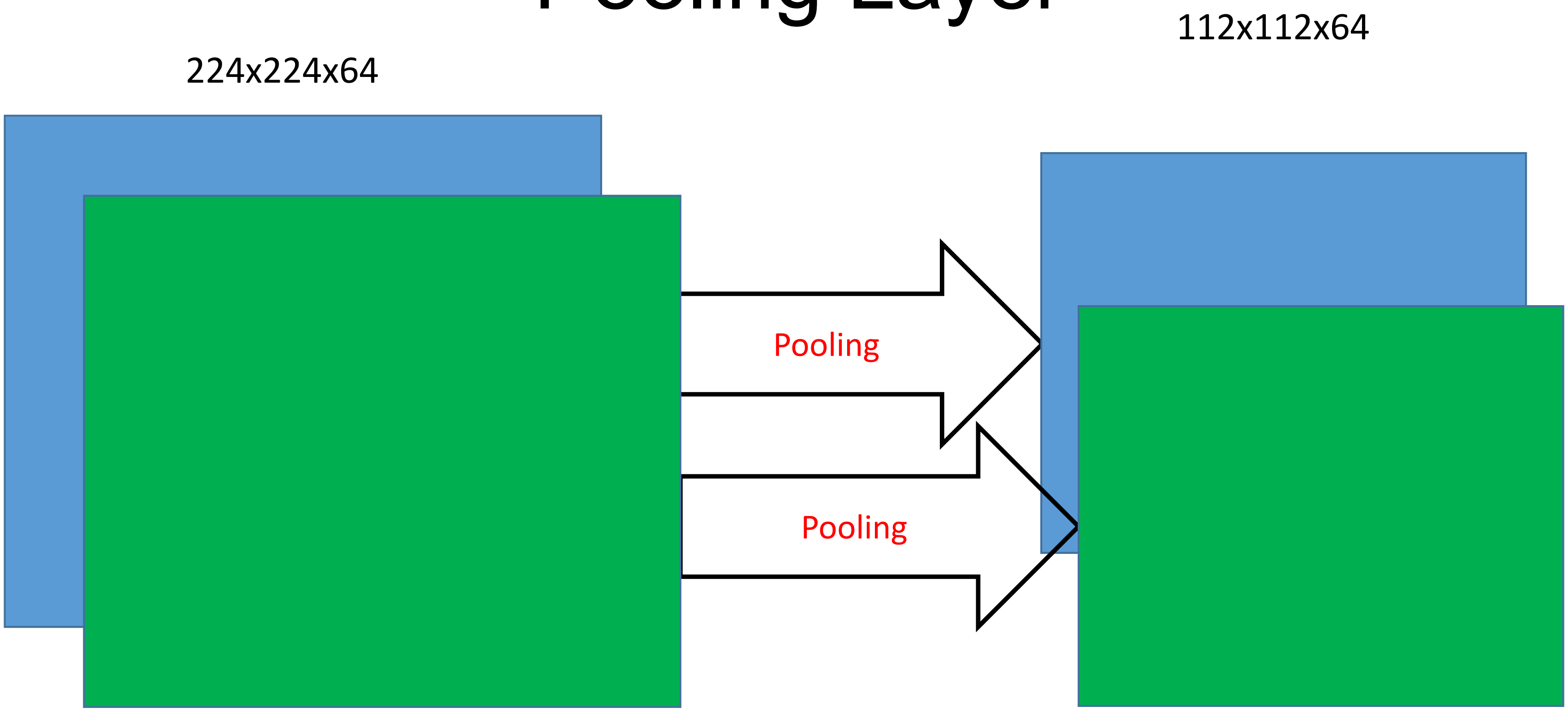

7x7 input image/data convolute with 3x3 filter with stride=1

# Convolution Neural Network



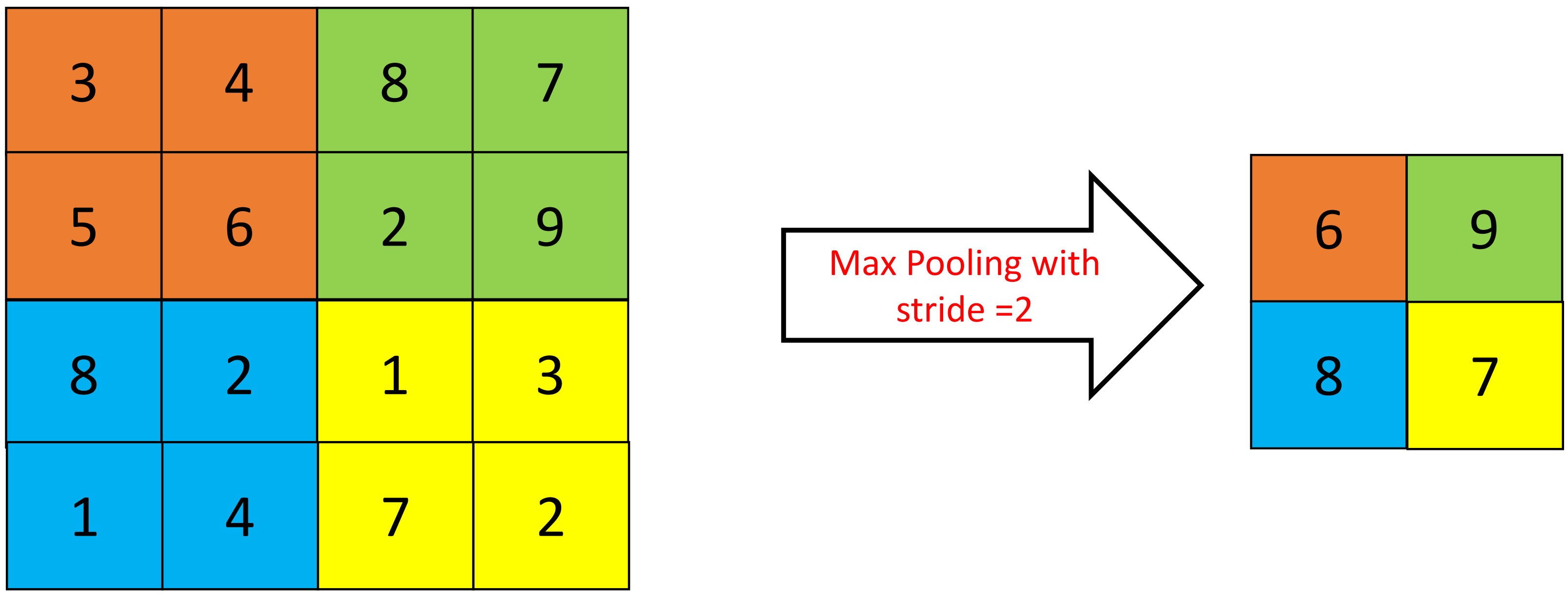

$$f[x,y]*g[x,y] \;=\; \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} f[n_1,n_2]\cdot g[x-n_1,y-n_2]$$

# Pooling Layer



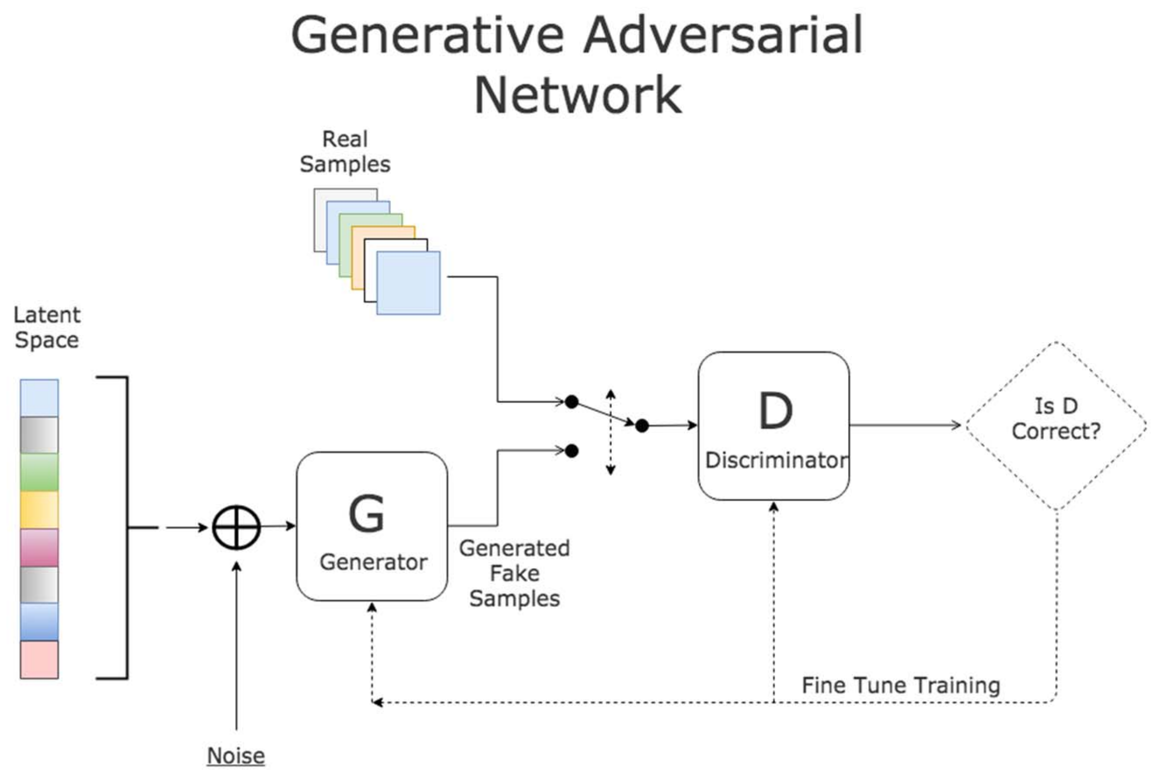

-makes the representation small and manageable
Operate over each feature map independently

# Max Pooling



# Generative Adversarial Networks (GANs)

# Generative Adversarial Networks (GANs)




- GAN is a relatively new Machine Learning architecture for neural networks pioneered by Ian Goodfellow and his colleagues at University of Montreal in 2014
- Consist of two competiting Neural Network the Generator and the Discriminator

Source: https://www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html

# GAN in Texture Generation or Super-Resolution

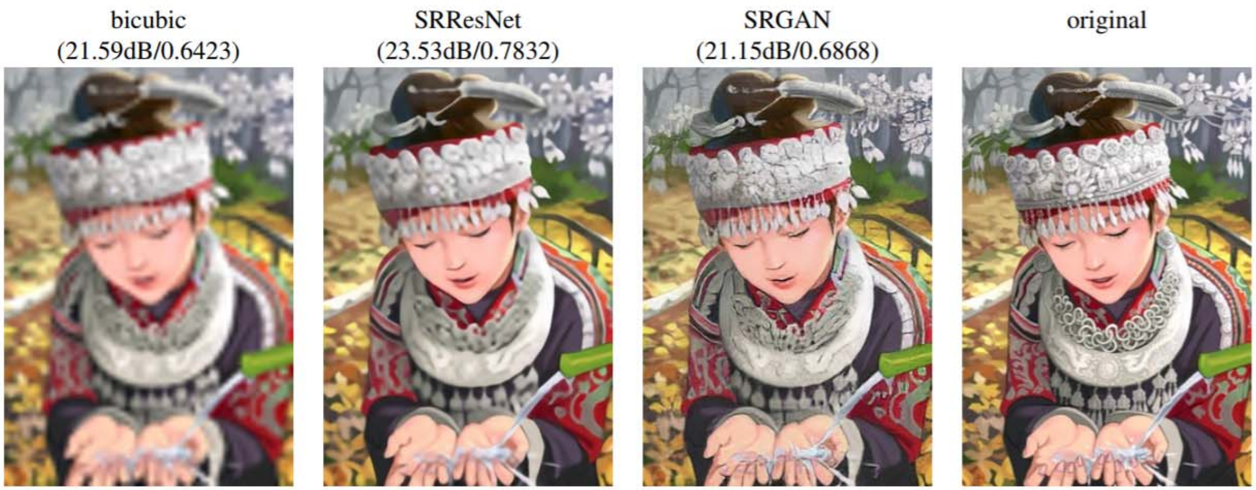


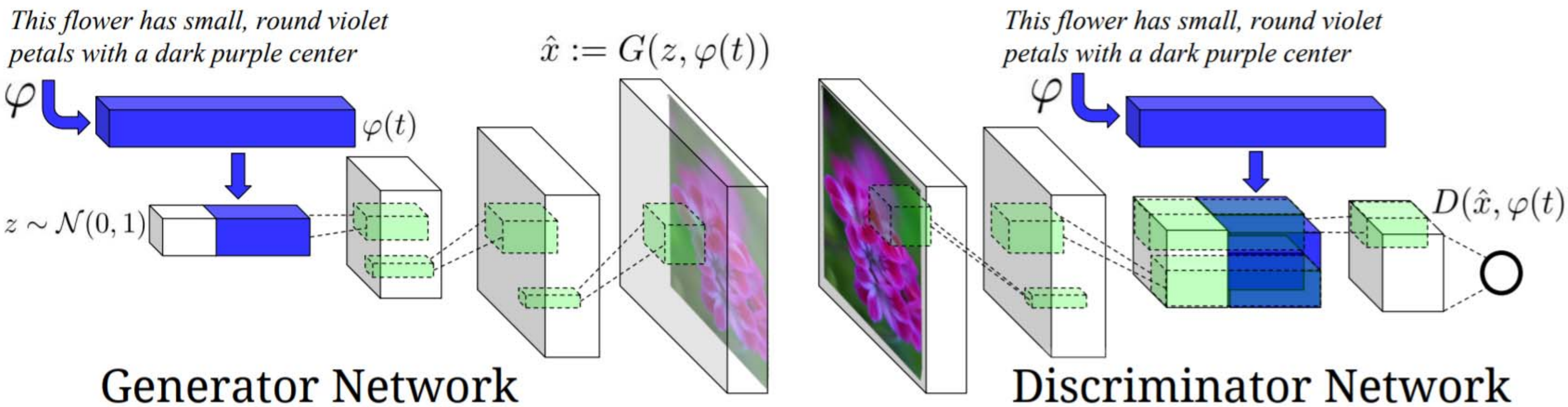

Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

The Twitter Cortex team uses GANs for super resolution.

Source: https://arxiv.org/pdf/1605.05396.pdf

arXiv paper on “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”

# GAN in Text to Image Synthesis




Sourcehttps://arxiv.org/pdf/1605.05396.pdf

Generative Adversarial Text-to-Image Synthesis Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee

# GAN in Text to Image Synthesis

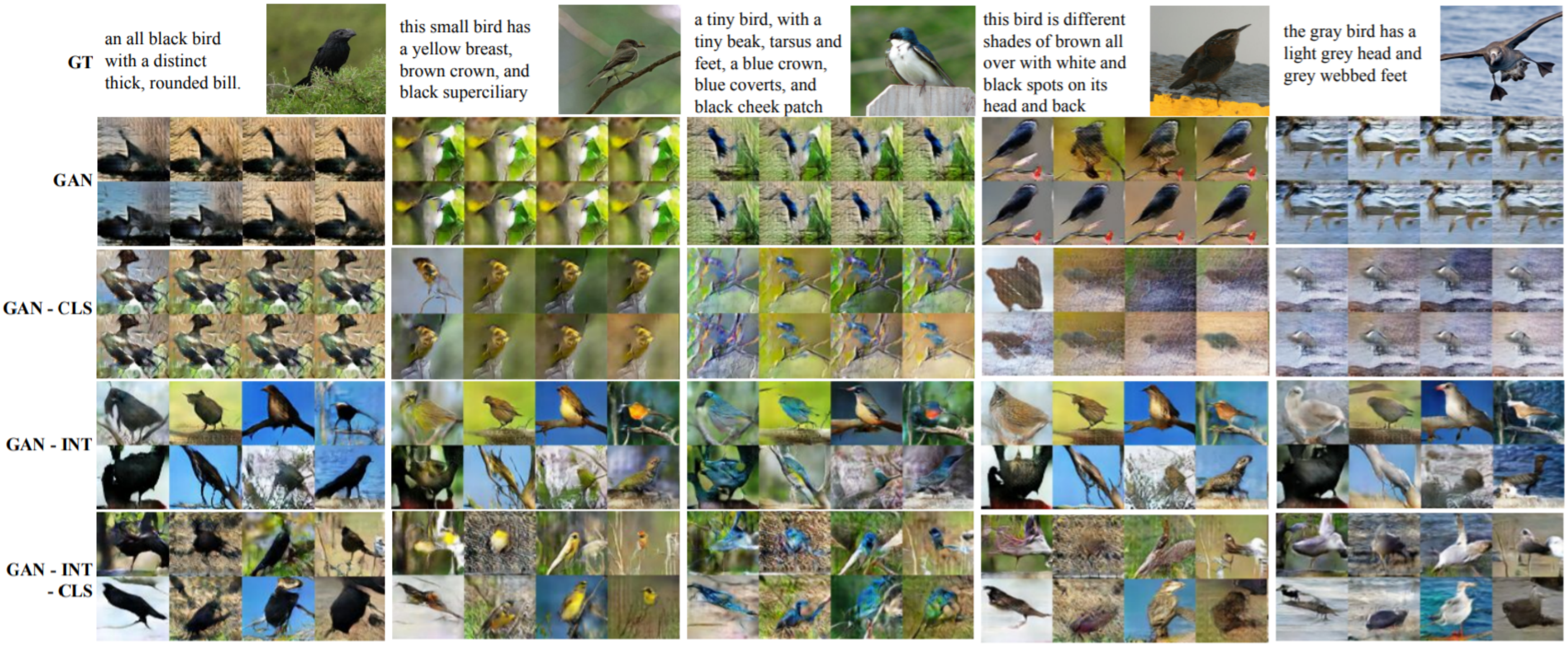



*Figure 3.* Zero-shot (i.e. conditioned on text from unseen test set categories) generated bird images using GAN, GAN-CLS, GAN-INT and GAN-INT-CLS. We found that interpolation regularizer was needed to reliably achieve visually-plausible results.

Sourcehttps://arxiv.org/pdf/1605.05396.pdf

# Learning Strategy

**Idea #1**: Choose hyperparameters that work best on the data

**BAD**: always works perfectly on training data

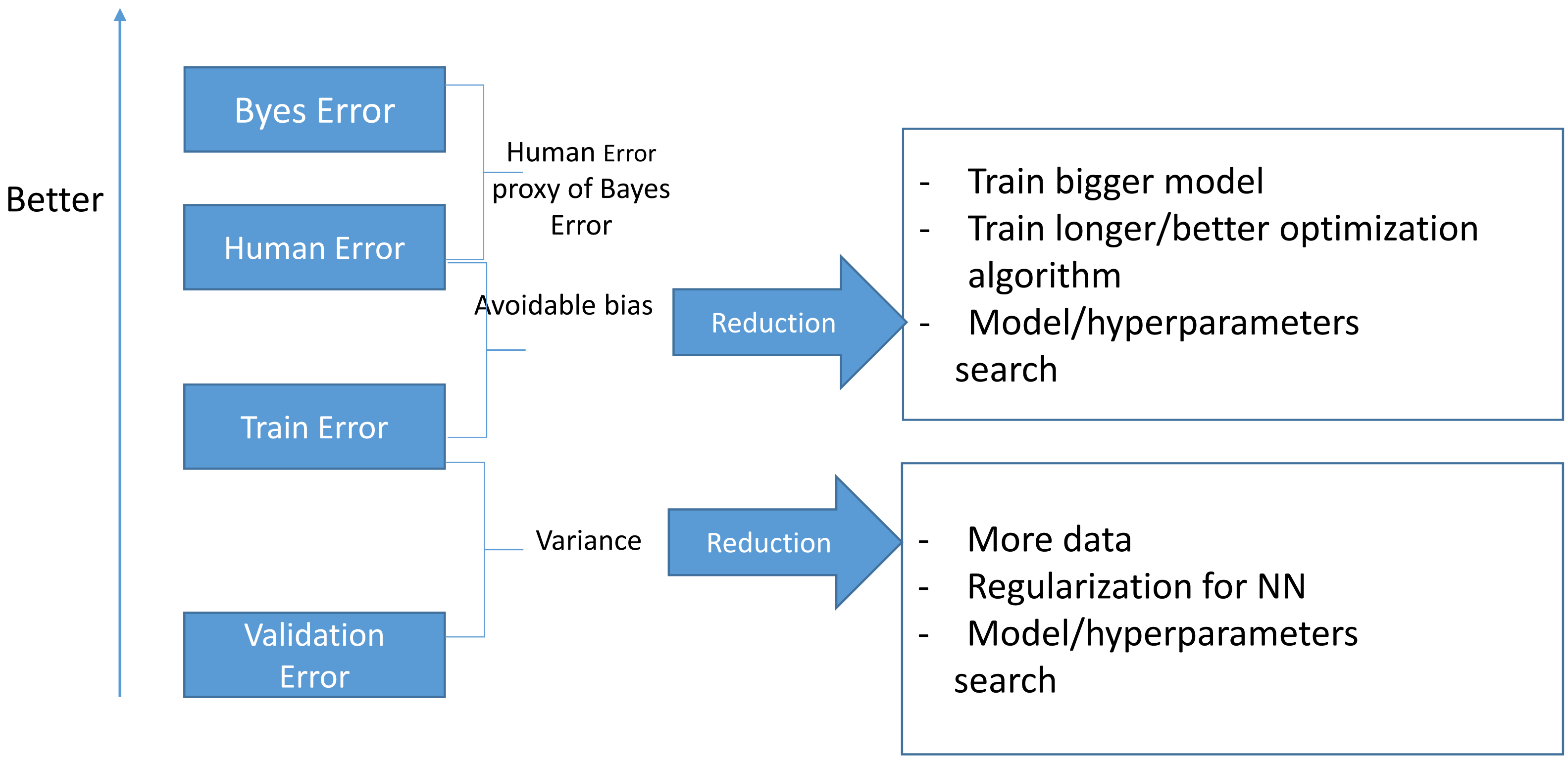

**Idea #2**: Split data into **train** and **test**, choose hyperparameters that work best on test data

**BAD**: No idea how algorithm will perform on new data

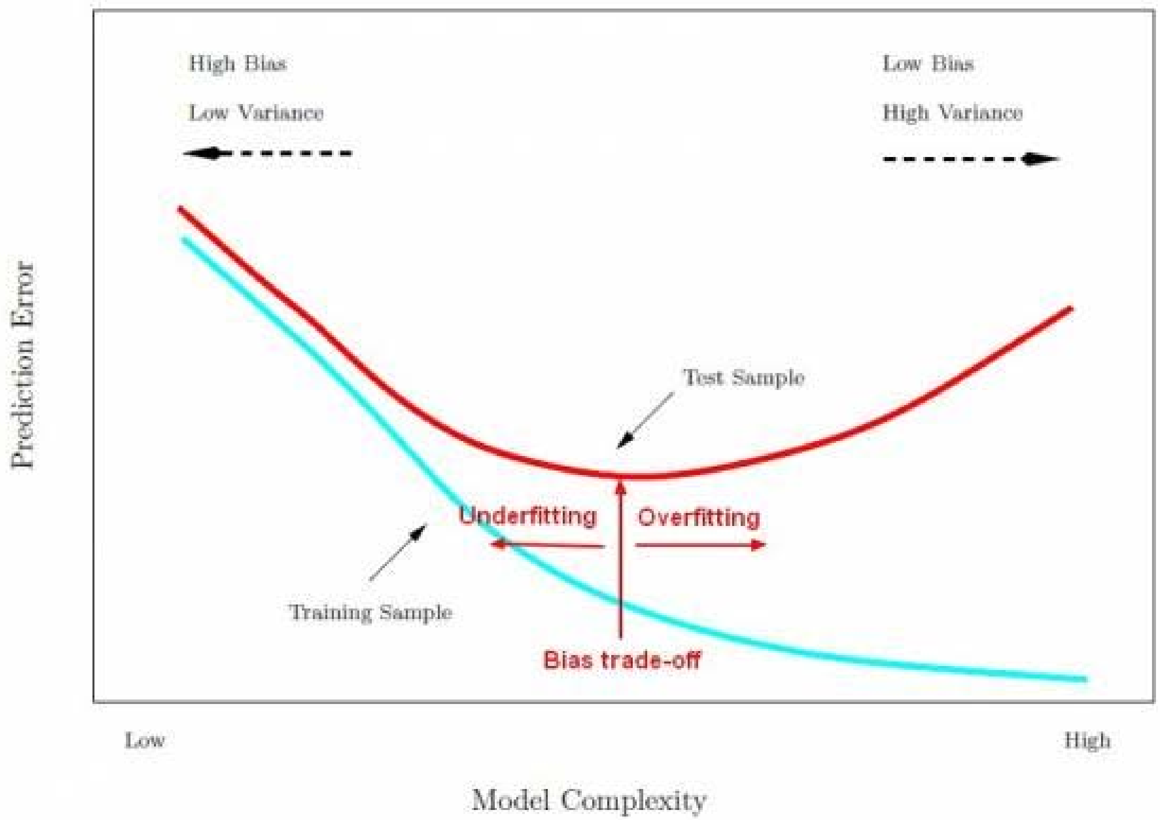

**Idea #3**: Split data into **train**, **val**, and **test**; choose hyperparameters on val and evaluate on test

**Better!**

# Strategy to Improve Prediction/Detection Result



# Overfitting and regularization



# L2 Regularization

$$\mathrm{E} = -\underbrace{\sum_{i=1}^{N}(t_i \log y_i) + (1 - t_i)\log(1 - y_i))}_{\text{Cross Entropy Cost function}} + \underbrace{\frac{\beta}{2n}\sum w^2}_{\text{L2 Regularization}}$$

Regularization parameter

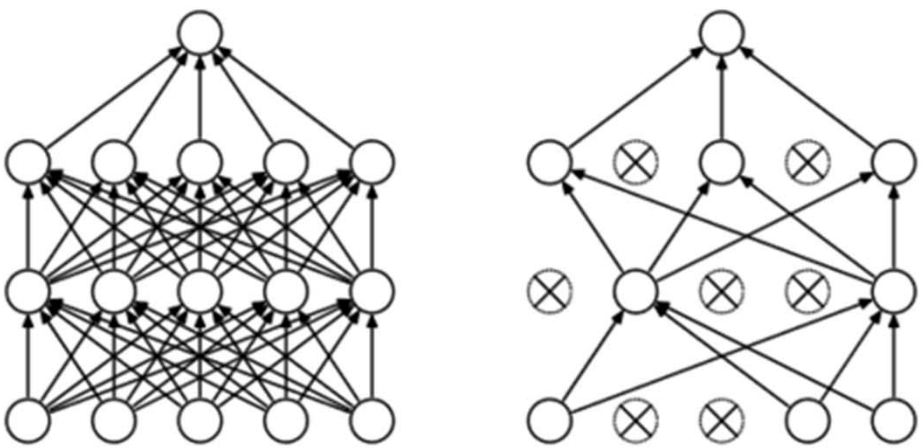
$$\frac{\beta}{2n}\sum w^2$$

Sum of square of
weights in the network

Make such that the
network prefers to learn
small weights

n is the size of training
set

# Dropout

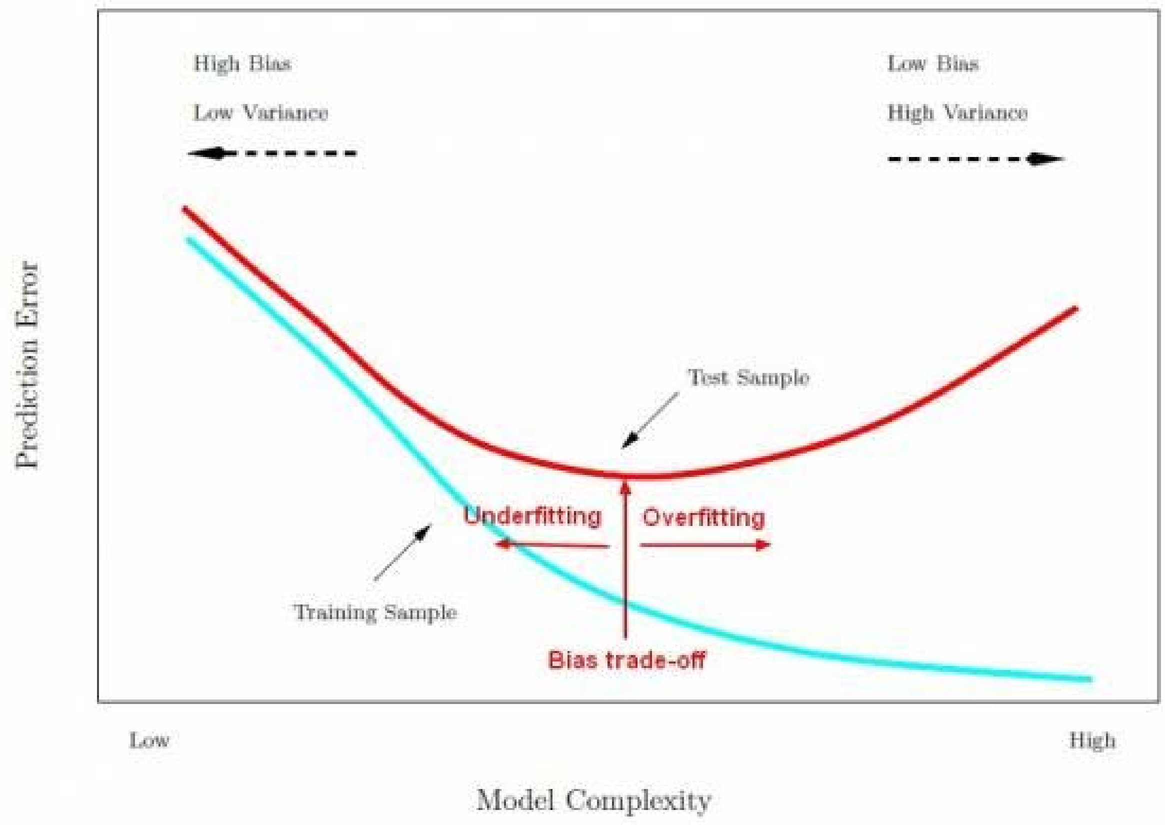


(a) Standard Neural Net (b) After applying dropout.

Sources: https://www.commonlounge.com/discussion/694fd08c36994186a48d122e511f29d5

- Dropout essential mean each neuron is drop (does not exist) at random with fixed probability of 1-p and kept with probability p.
- The value of p maybe different for different layer in the same network
- p=0.5 for hidden layer and p=1 for input layer (no dropout) shown to provide good result
- No neuron is drop during evaluation and prediction
- The output of each neuron is multiplied by p so that the input to the next layer has the same expected value.

# Early Stopping



# Learning Rate

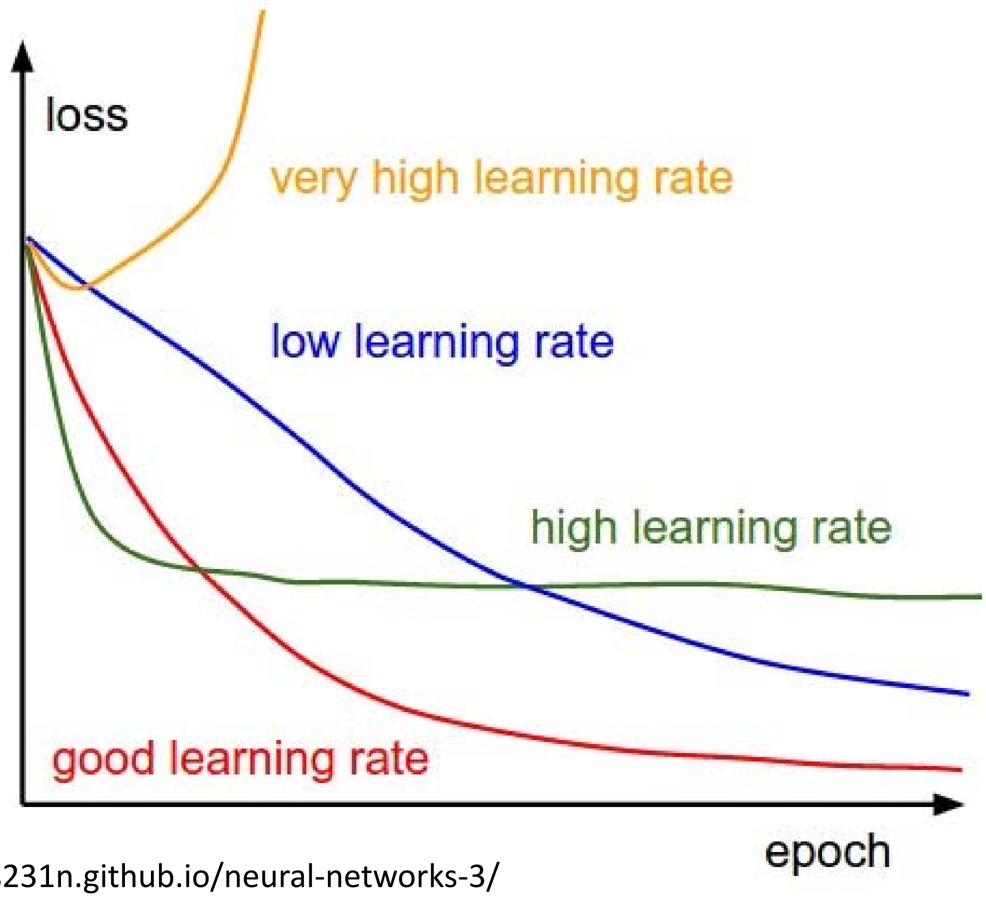



Image Scource: http://cs231n.github.io/neural-networks-3/