TÜRKİYE
YAPAY ZEKA
ZİRVESİ

İTÜ

SiMiT Lab
simitlab.itu.edu.tr

İTÜ ile Yapay Zeka

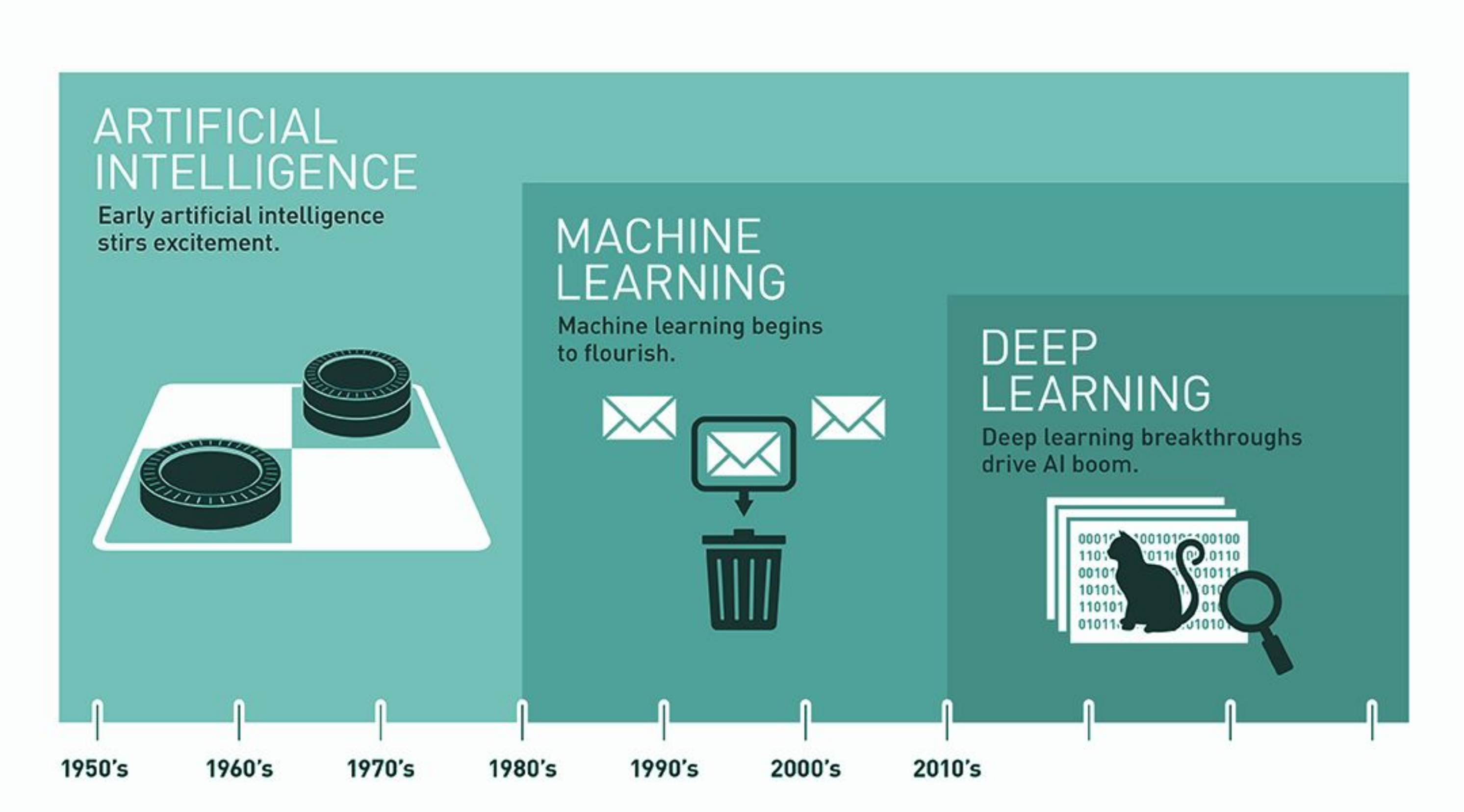
İrem Eyiokur, Doğucan Yaman, Anıl Genç, Omid Abdollahi Aghdam

{eyiokur16, yamand16, genca16, abdollahi15}@itu.edu.tr

Smart Interaction, Mobile Intelligence, and Multimedia Technologies Lab (SiMiT Lab)

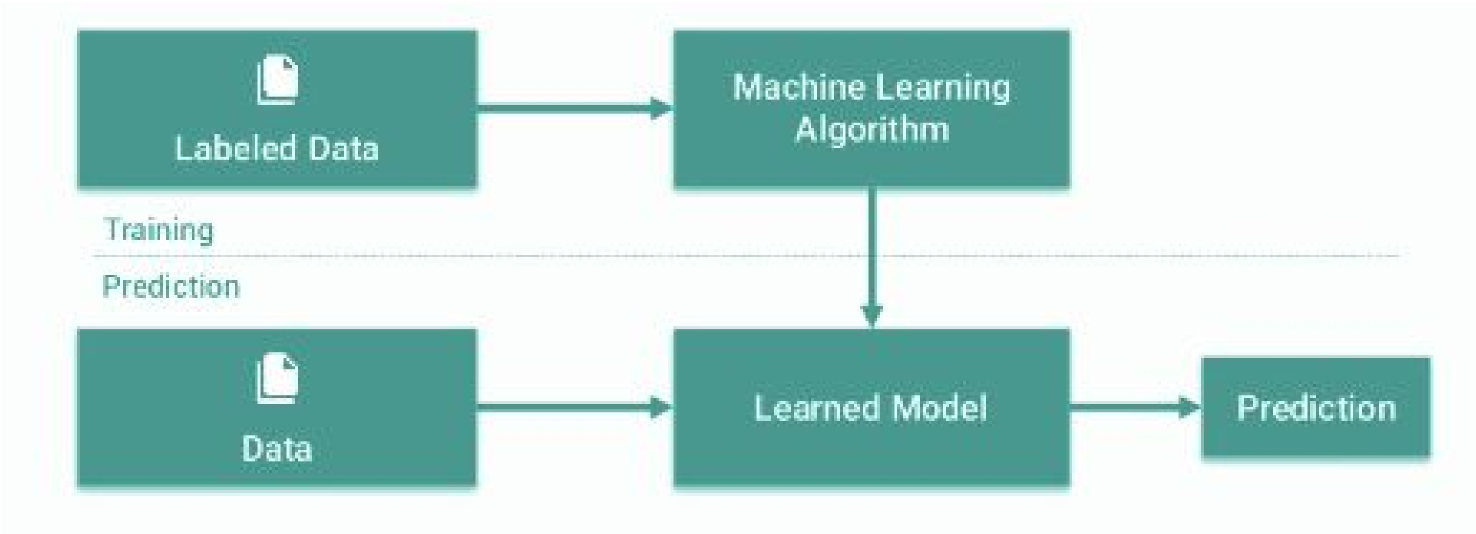
Lab Director : Doç. Dr. Hazım Kemal EKENEL

What is Deep Learning ?



blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

Machine Learning

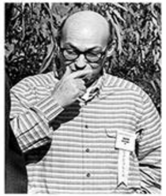


slideshare.net/LuMa921/deep-learning-a-visual-introduction

A BRIEF HISTORY OF DEEP LEARNING

1958
Cornell psychologist Frank Rosenblatt unveils the Perceptron, a single-layer neural network on a room-size computer.

1969
AI giant Marvin Minsky of MIT cowrites a book casting doubt on the viability of neural networks. They fall out of favor.



1986
Neural nets pioneer Geoffrey Hinton and others find a way to train multilayer neural networks to correct mistakes. A flurry of activity ensues.

1989
French researcher Yann LeCun, then at Bell Labs, begins foundational work on a type of neural net that becomes crucial for image recognition.

FREDERIC LEWIS—ARCHIVE PHOTOS/GETTY IMAGES, LEFT: ANN E. YOW-DYSON—GETTY IMAGES

1991
German researchers Sepp Hochreiter and Jürgen Schmidhuber pioneer a neural net with memory features, which eventually proves superior for natural-language processing.

1997
IBM's Deep Blue beats **world champion Garry Kasparov** (right) in chess using traditional AI techniques.

STAN HONDA—AFP/GETTY IMAGES

Mid-1990s
Neural nets fall into disfavor again, eclipsed by other machine-learning techniques.

2007
Fei-Fei Li founds ImageNet and begins assembling a database of 14 million labeled images that can be used for machine-learning research.

2011
Microsoft introduces neural nets into its speech-recognition features.

2011
IBM's Watson beats two champions at Jeopardy using traditional AI techniques.

CARLOS CHAVARRIA—THE NEW YORK TIMES/REDUX PICTURES

2012
JUNE
Google Brain publishes the "cat experiment." A neural net, shown 10 million unlabeled YouTube images, has trained itself to recognize cats.

AUGUST
Google introduces neural nets into its speech-recognition features.

OCTOBER
A neural net designed by two of Hinton's students wins the annual ImageNet contest by a wide margin.

2013
MAY
Google improves photo search using neural nets.

JIM WILSON—THE NEW YORK TIMES/REDUX PICTURES

2014
JANUARY
Google acquires DeepMind, a startup specializing in combining deep learning and reinforcement learning, for $600 million.

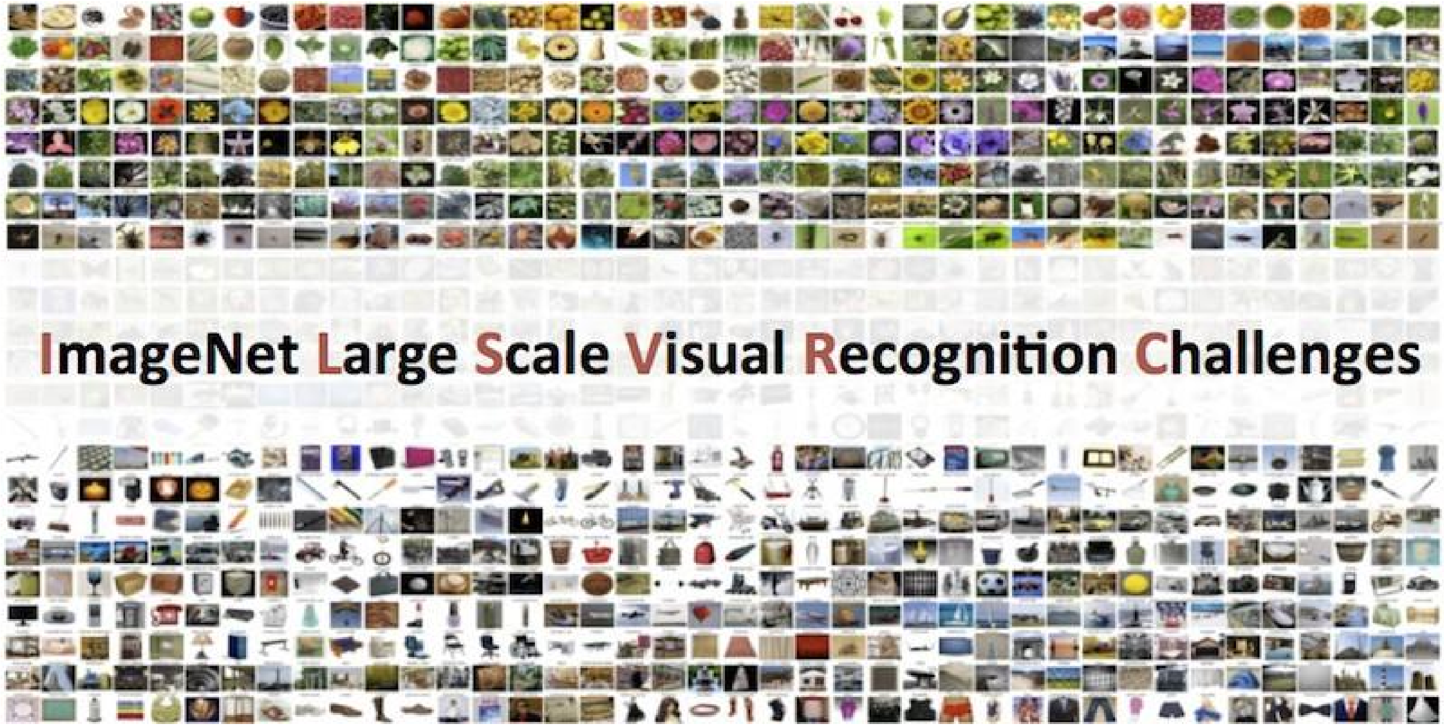
2015
DECEMBER
A team from Microsoft, using neural nets, outperforms a human on the ImageNet challenge.

2016
MARCH
DeepMind's AlphaGo, using deep learning, defeats world champion **Lee Sedol** in the Chinese game of go, four games to one.

LEE JIN-MAN—AP PHOTO

Source: - fortune.com

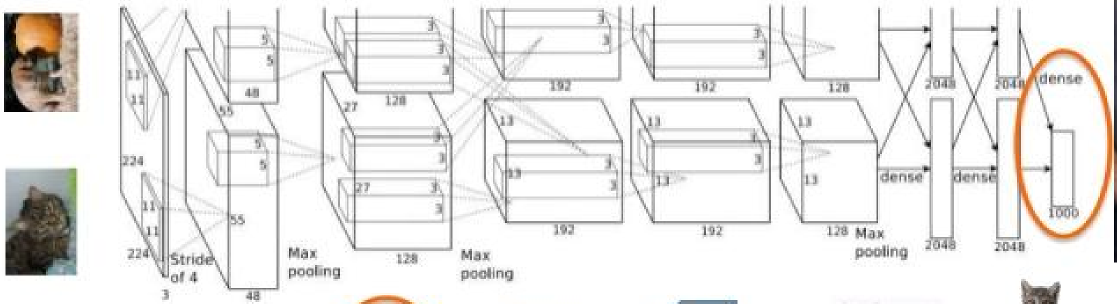
A Brief History of Deep Learning

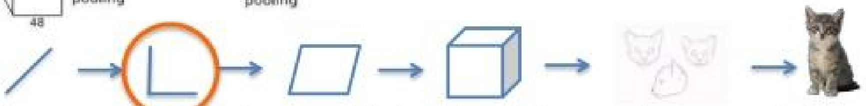


medium.com/chris-nielsen/making-my-own-deep-learning-image-recogniser-d707435a788e

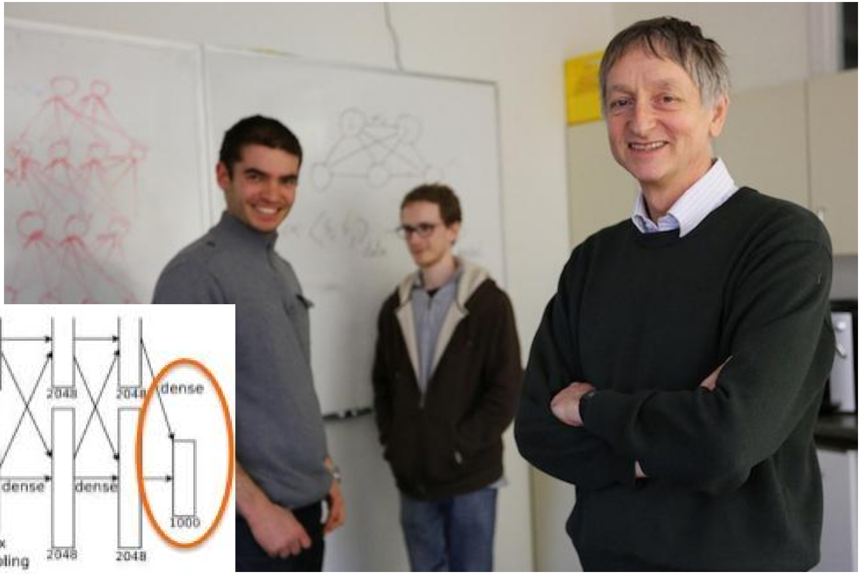
ILSVRC 2012 Winner: AlexNet(Krizhevsky et al.)

AlexNet CNN Architecture





When AlexNet is processing an image, this is what is happening at each layer.



lya Sutskever (left), Alex Krizhevsky (middle), Geoffrey Hinton (right)

University of Toronto

www.slideshare.net/AlexanderKorbonits/deep-learning-with-python-pydata-seattle-2015

Deep Learning Allstars

Geoffrey Hinton: University of Toronto & Google

Yann LeCun: New York University & Facebook

Andrew Ng: Stanford

Yoshua Bengio: University of Montreal

Jürgen Schmidhuber: Swiss AI Lab & NNAISENSE

www.slideshare.net/LuMa921/deep-learning-a-visual-introduction

Deep Learning Allstars

Deep Learning is an algorithm which has no theoretical limitations of what it can learn; the more data you give and the more computational time you provide, the better it is.

-Geoffrey Hinton (Prof. @ Toronto University - Resercher @ Google)

Deep Learning Allstars

Artificial Intelligence is the new electricity.

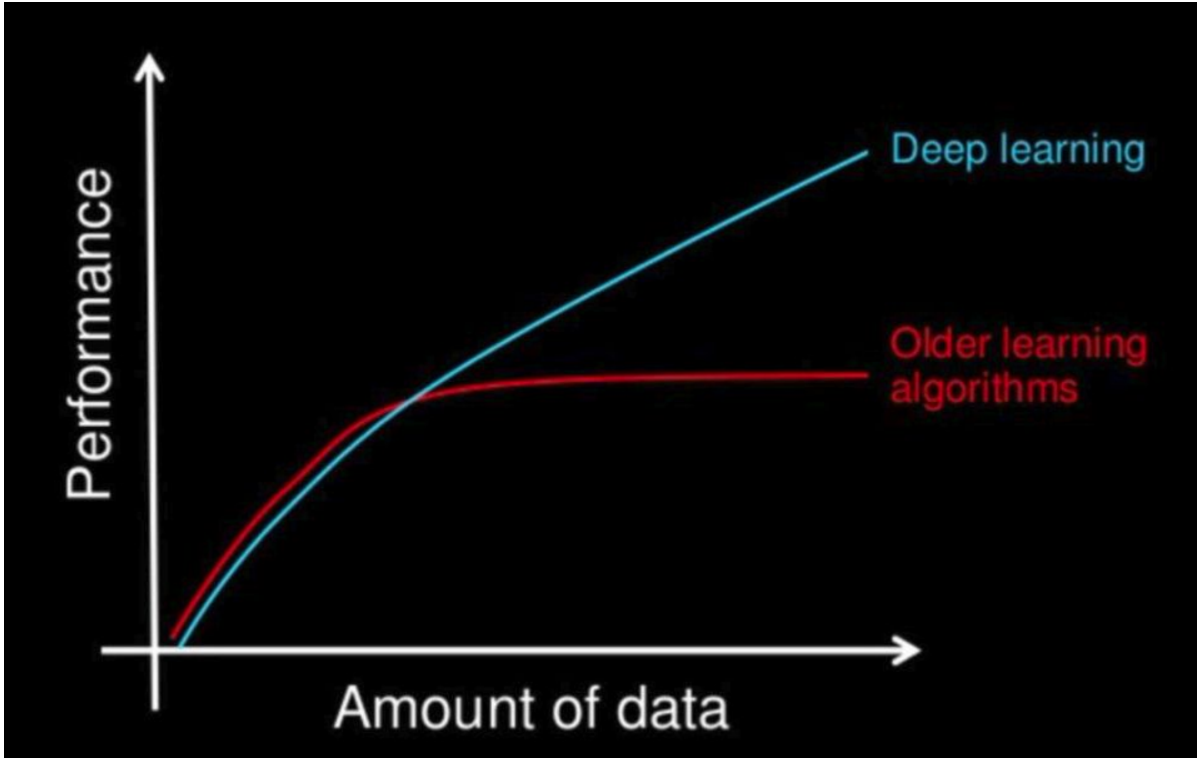
I have worked all my life in Machine Learning, and I've never seen one algorithm knock over benchmarks like Deep Learning.

-Andrew Ng (Prof. @ Stanford University)

Why Deep Learning ?

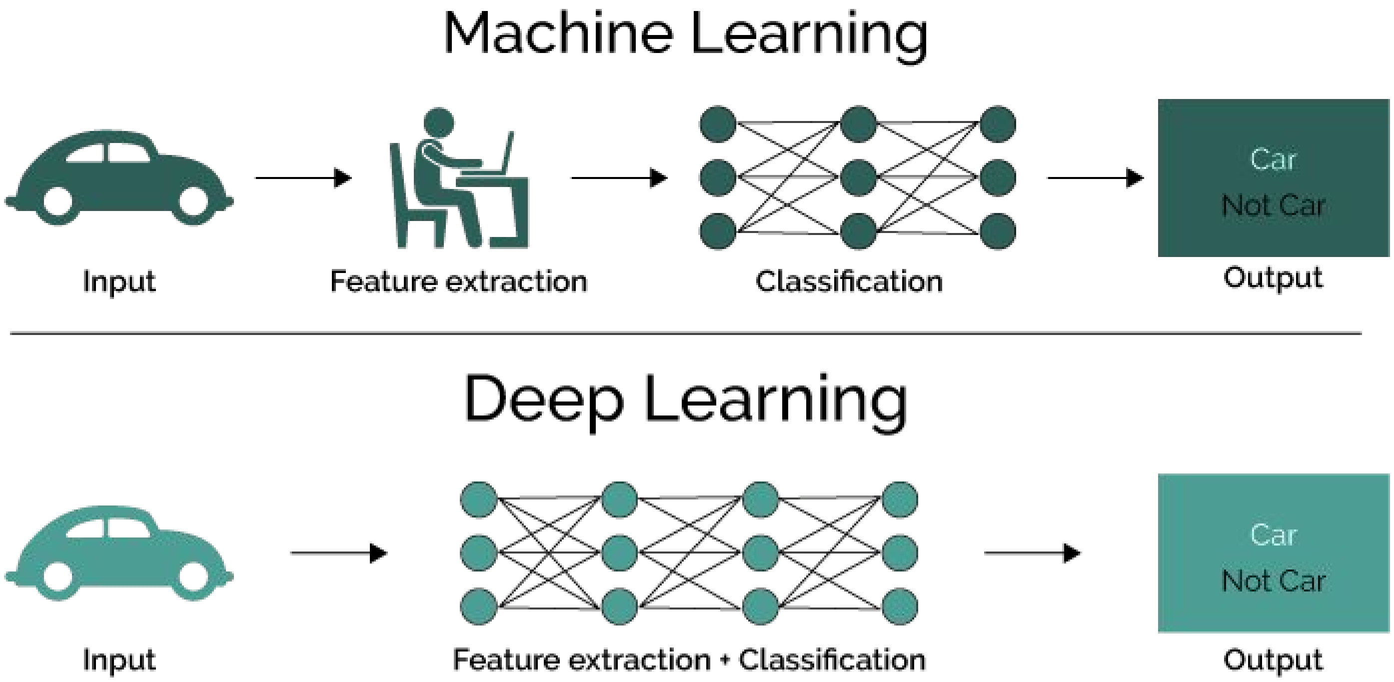
www.slideshare.net/LuMa921/deep-learning-a-visual-introduction

Why Deep Learning ?



slideshare.net/ExtractConf

Difference Between ML and DL?



/www.xenonstack.com/blog/data-science/log-analytics-with-deep-learning-and-machine-learning

Challenges of Deep Learning

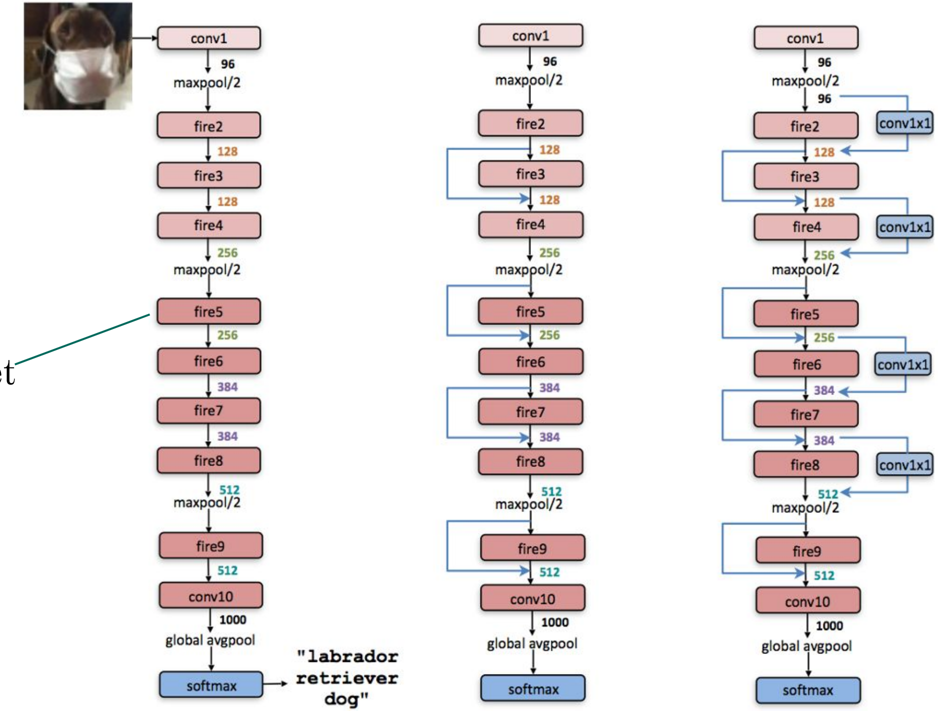
The First Challenge: Model Size

The Second Challenge: Speed

The Third Challenge: Energy Efficiency

Challenges of Deep Learning

SquzezeNet
Models



[SqueezeNet, Iandola]

The First Challenge: Model Size

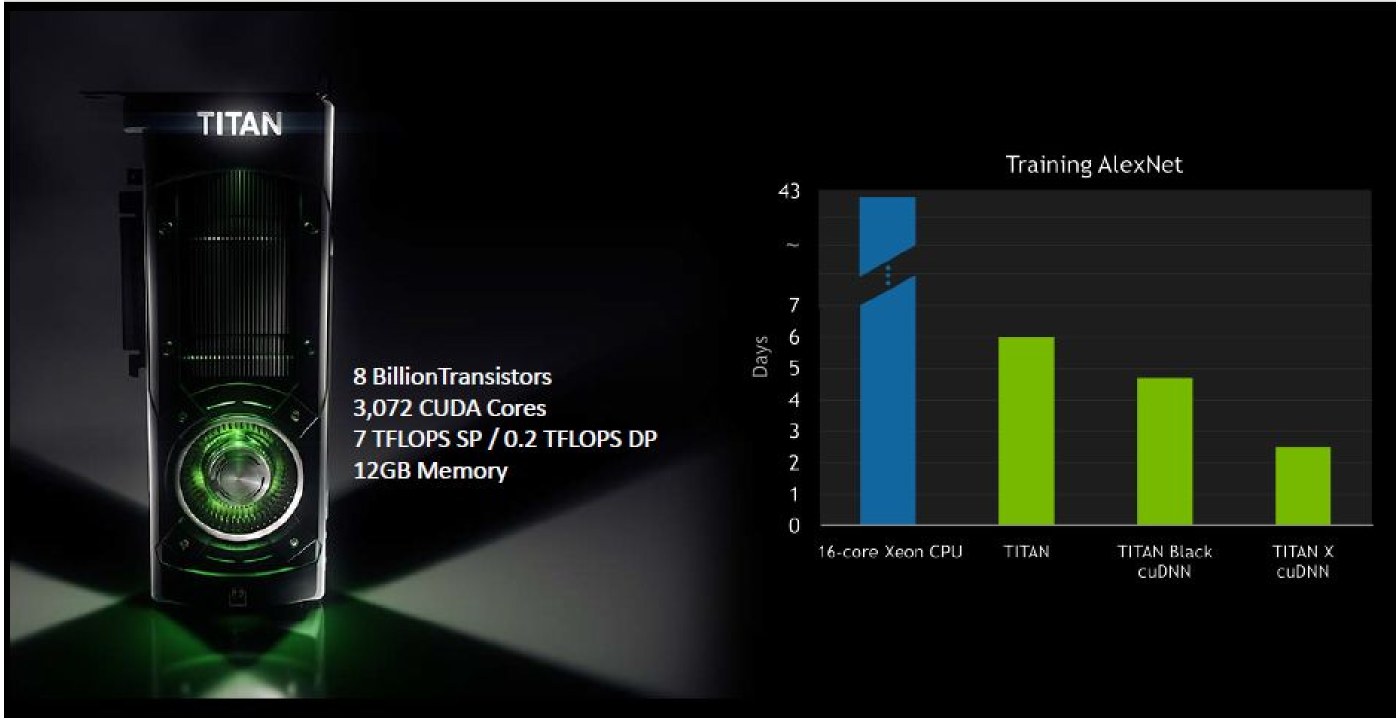
Challenges of Deep Learning

CNN architecture	Compression Approach	Data Type	Original → Compressed Model Size	Reduction in Model Size vs. AlexNet	Top-1 ImageNet Accuracy	Top-5 ImageNet Accuracy
AlexNet	None (baseline)	32 bit	240MB	1x	57.2%	80.3%
AlexNet	SVD [5]	32 bit	240MB → 48MB	5x	56.0%	79.4%
AlexNet	Network Pruning [11]	32 bit	240MB → 27MB	9x	57.2%	80.3%
AlexNet	Deep Compression [10]	5-8 bit	240MB → 6.9MB	35x	57.2%	80.3%
SqueezeNet (ours)	None	32 bit	4.8MB	**50x**	57.5%	80.3%
SqueezeNet (ours)	Deep Compression	8 bit	4.8MB → 0.66MB	**363x**	57.5%	80.3%
SqueezeNet (ours)	Deep Compression	6 bit	4.8MB → 0.47MB	**510x**	57.5%	80.3%

[SqueezeNet, Iandola]

The First Challenge: Model Size

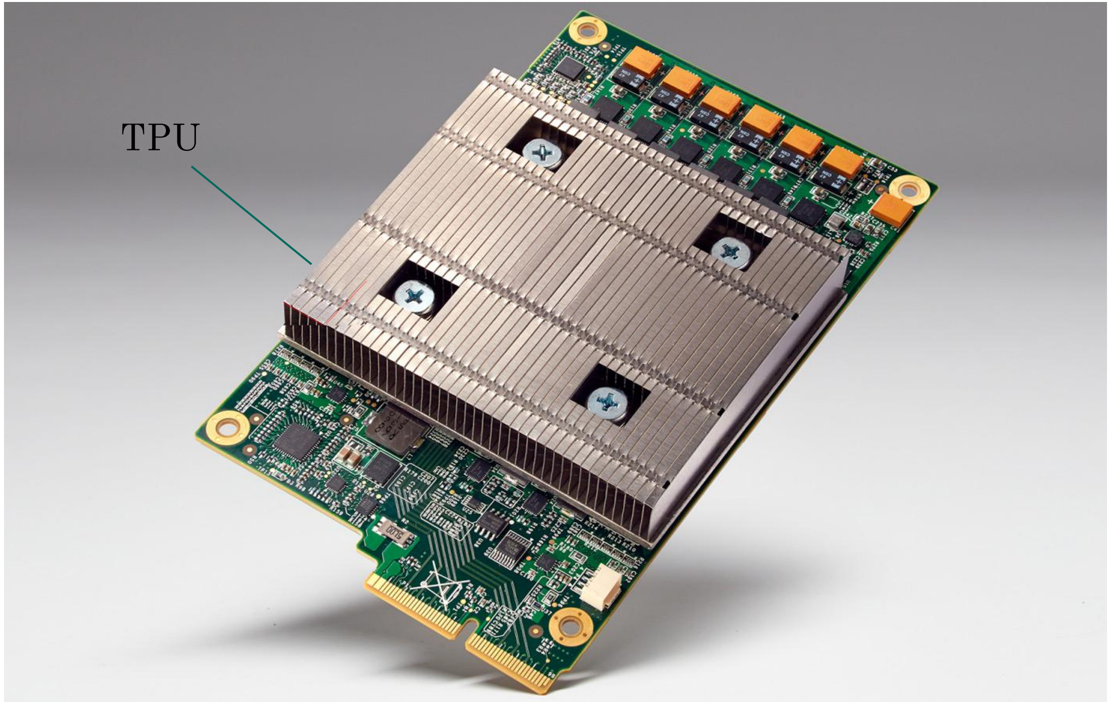
Challenges of Deep Learning



The Second Challenge: Speed

slideshare.net/YusufKurt/derin-renme-nedir-akademik-biliim-2016-sunumu

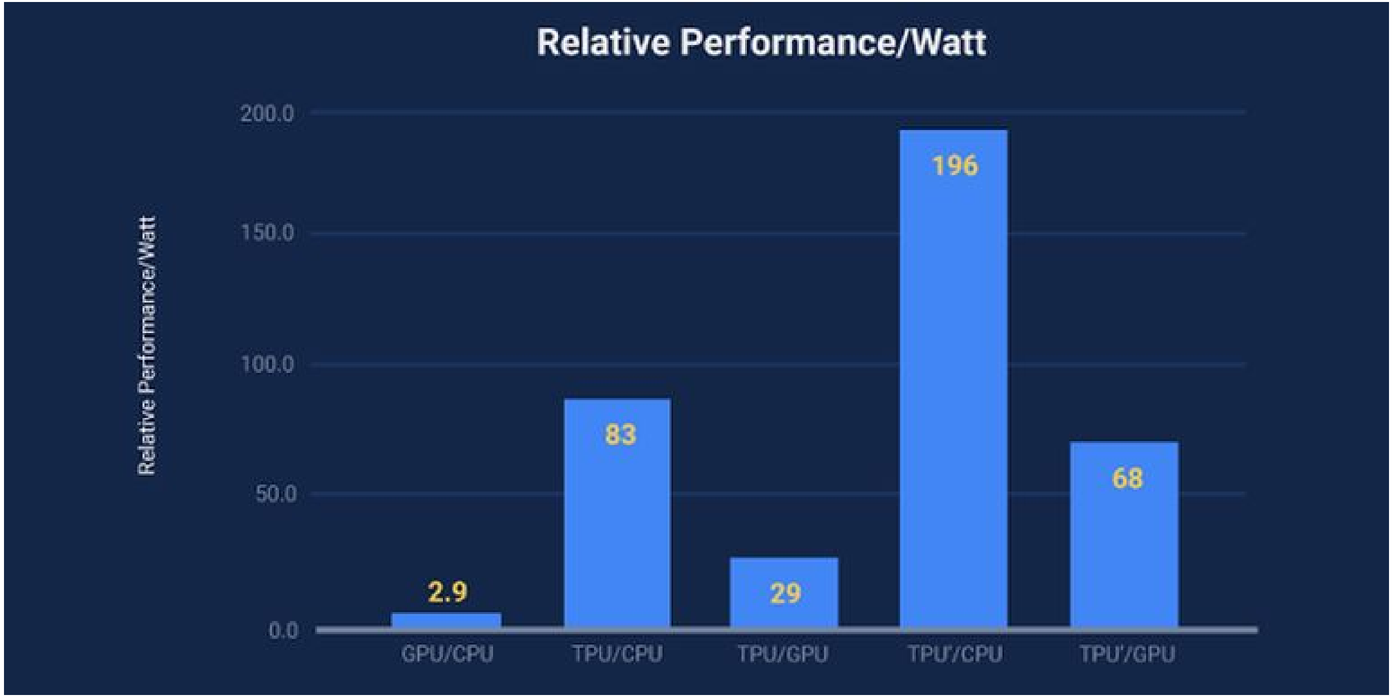
Challenges of Deep Learning



The Second Challenge: Energy Efficiency

cloudplatform.googleblog.com/2016/05/Google-supercharges-machine-learning-tasks-with-custom-chip.html

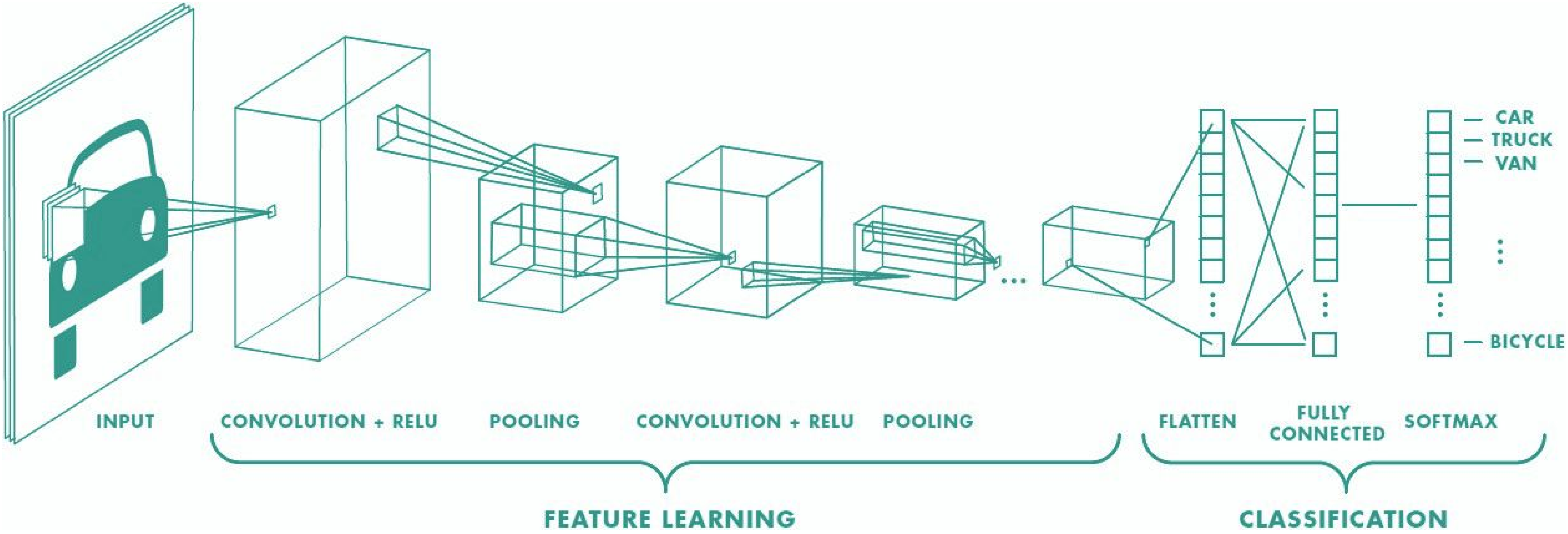
Challenges of Deep Learning



The Second Challenge: Energy Efficiency

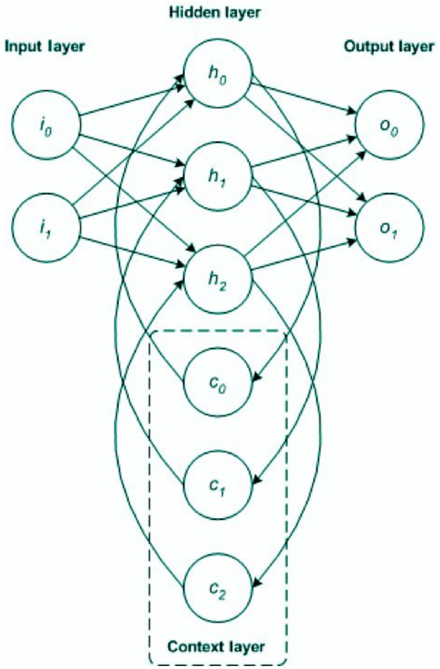
cloudplatform.googleblog.com/2017/04/quantifying-the-performance-of-the-TPU-our-first-machine-learning-chip.html

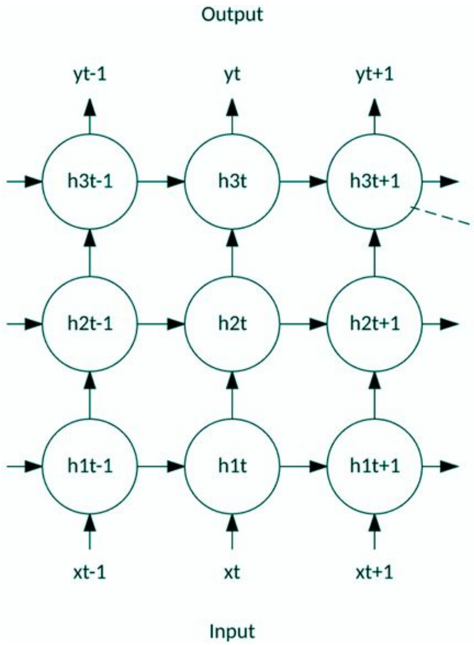
Deep Learning Models - CNN Model

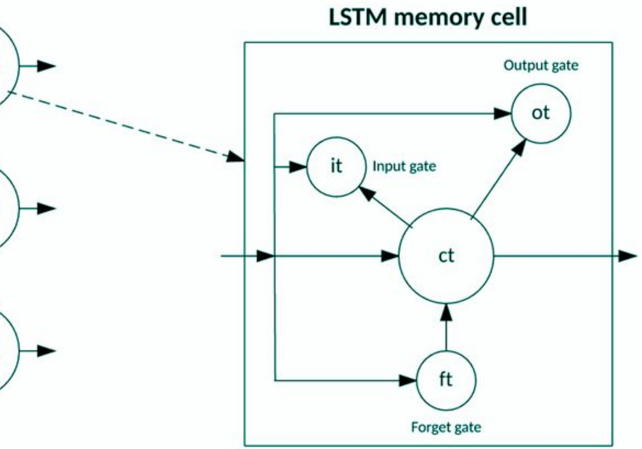


https://www.mathworks.com/content/mathworks/www/en/discovery/convolutional-neural-network.html

Deep Learning Models - RNN & LSTM



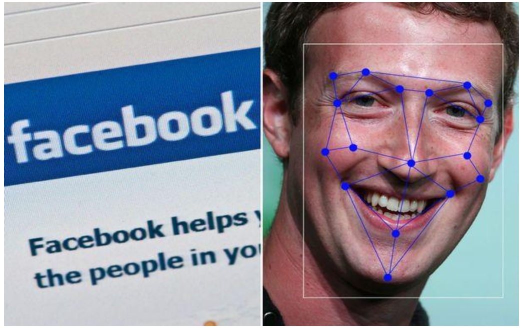


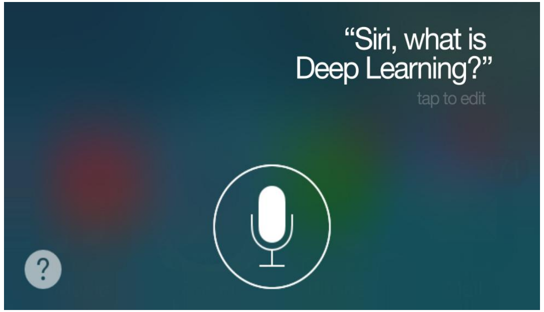


Usage Areas of DL

- Speech Recognition => RNN, LSTM
- Handwritten Recognition => RNN, LSTM
- Natural Language Processing => CNN, LSTM
- Image Captioning => LSTM
- Image Recognition/Segmentation/Detection => CNN
- Video Analysis => CNN

Usage Areas of DL

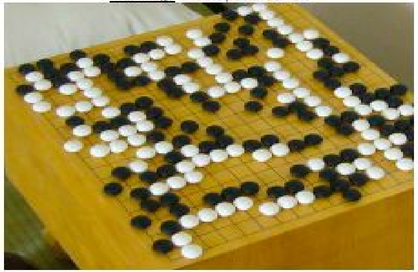




Usage Areas of DL

Self-Driving Car

Machine Translation

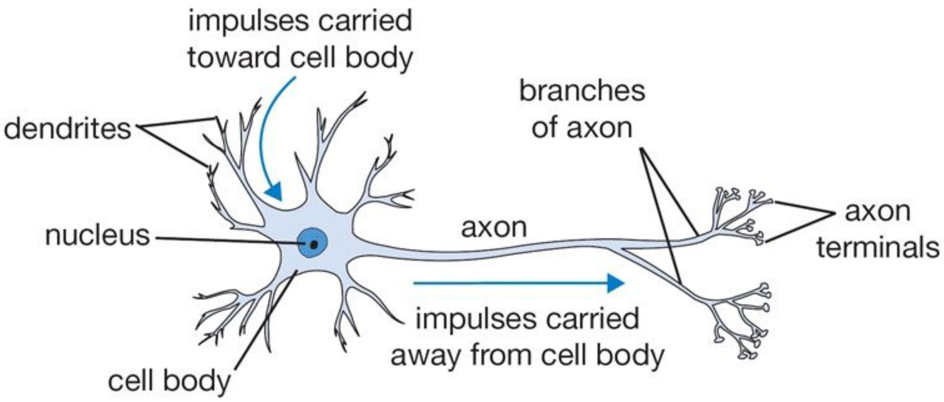


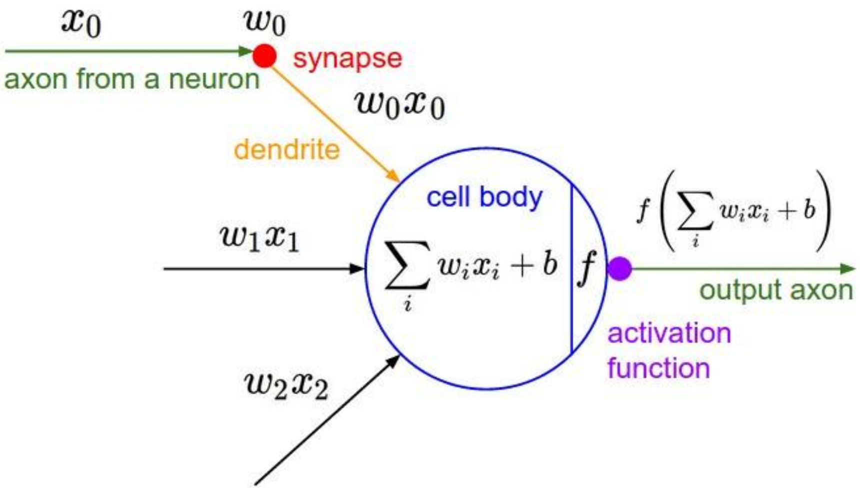
AlphaGo

Smart Robots

Frameworks for Deep Learning

Neural Network

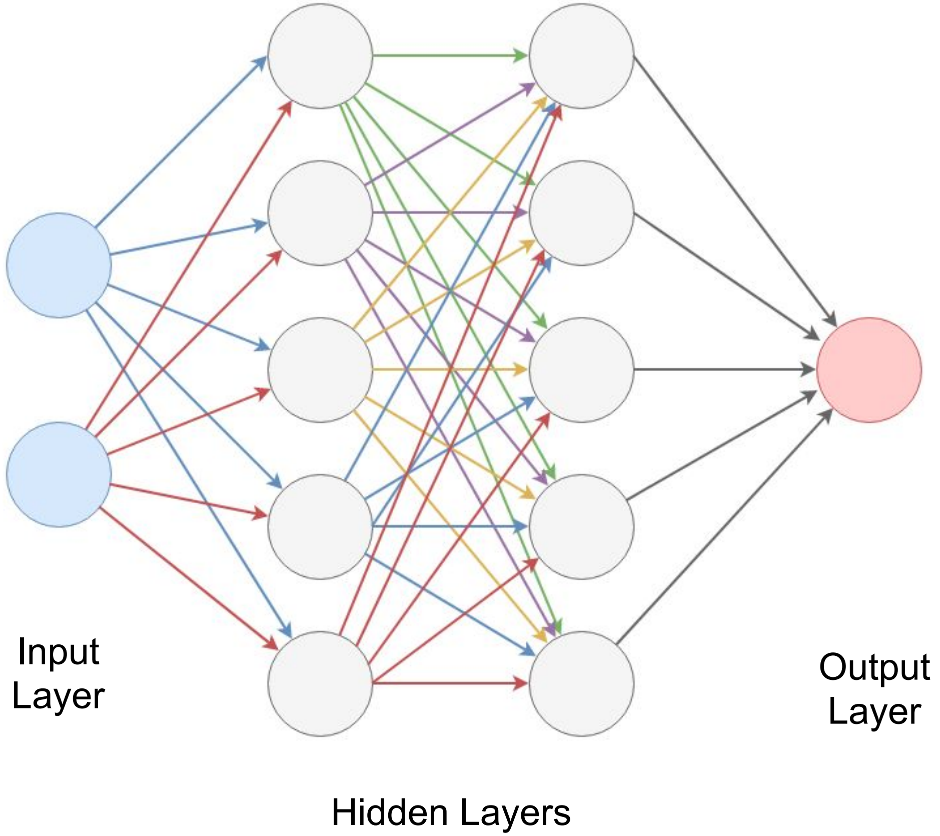




biological neuron (left) and its mathematical model (right).

cs231n.github.io/neural-networks-1/

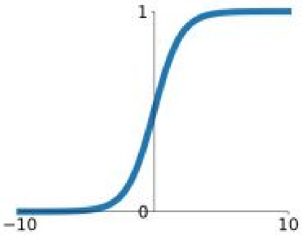
Neural Network - Fully Connected Network



NN Models - Activation Function

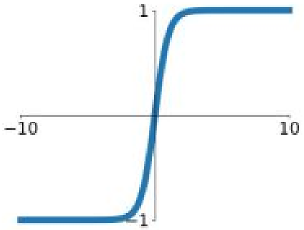
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



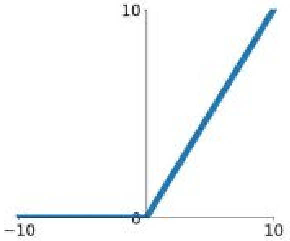
tanh

$\tanh(x)$



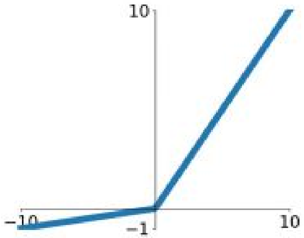
ReLU

$\max(0, x)$



Leaky ReLU

$\max(0.1x, x)$

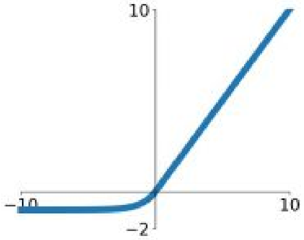


Maxout

$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



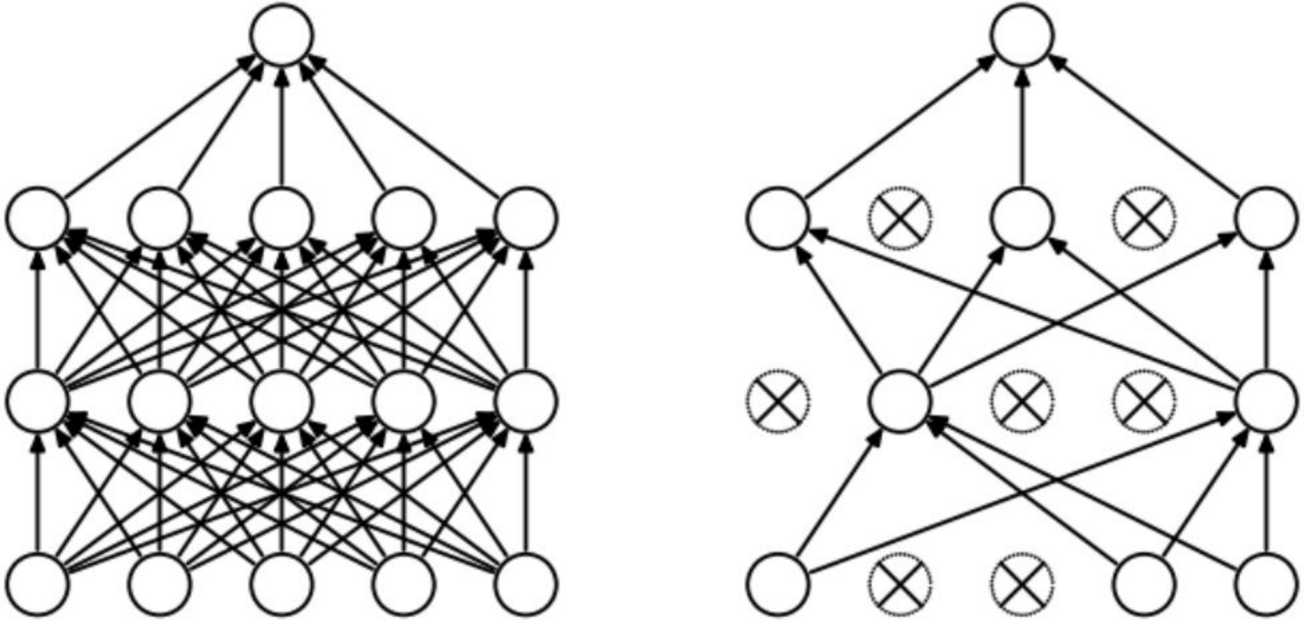
http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture6.pdf

NN Model - Regularization

$$L = \frac{1}{N}\sum_{i=1}^{N}\sum_{i \neq j} max(0, f(x_i;W)_j - f(x_i;W)_{y_i} + 1) + \lambda R(W)$$

- ➔ L1 Regularization, $\sum_k \sum_l |W_{k,l}|$
- ➔ L2 Regularization, $\sum_k \sum_l W_{k,l}^2$
- ➔ Max Norm Regularization
- ➔ L1+L2 Regularization
- ➔ Dropout
- ➔ Data Augmentation

NN Model - Dropout

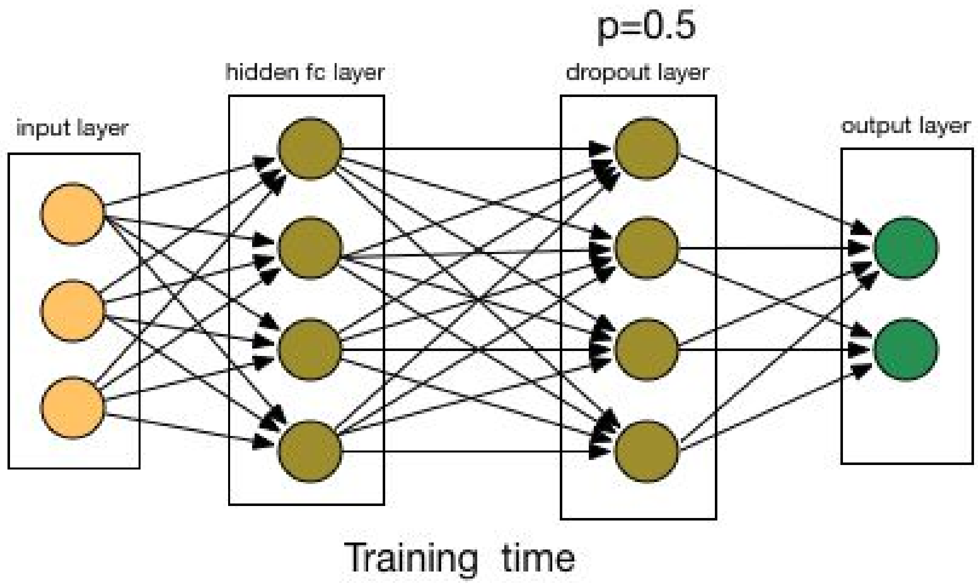


(a) Standard Neural Net

(b) After applying dropout.

Srivastava, N., Hinton G., et. al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", JMLR, 2014

NN Model - Dropout



https://chatbotslife.com/regularization-in-deep-learning-f649a45d6e0

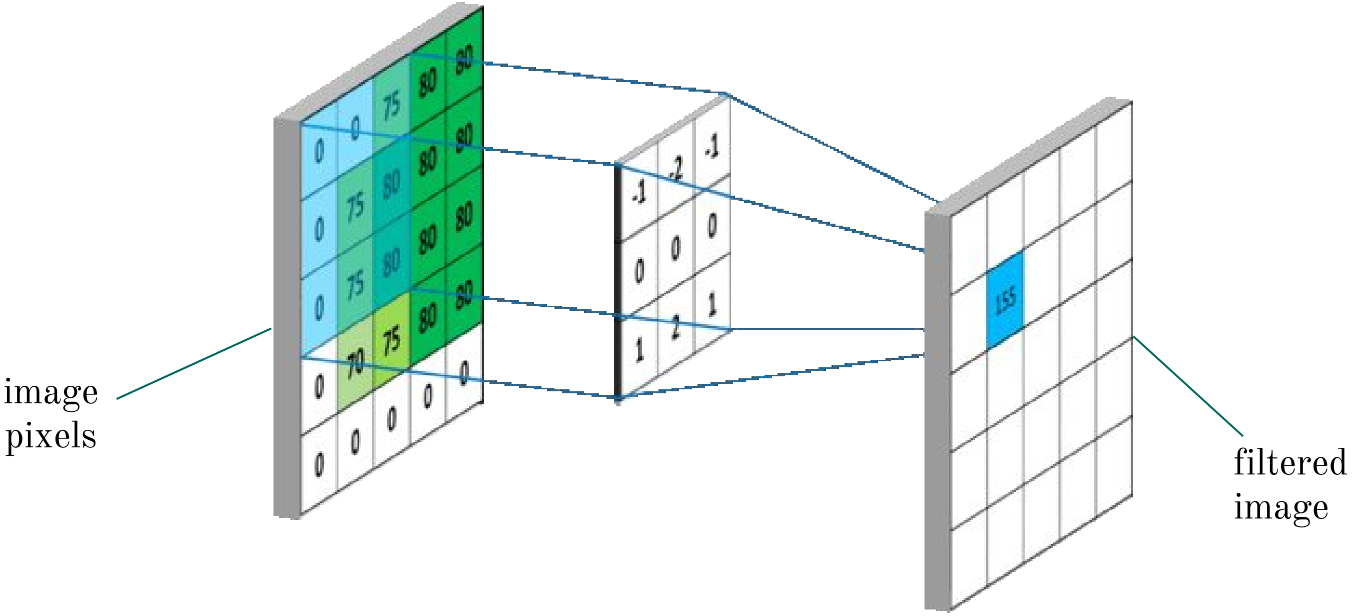
NN Model - Fully Connected Layers

Linear Classifier $\longrightarrow f = Wx$

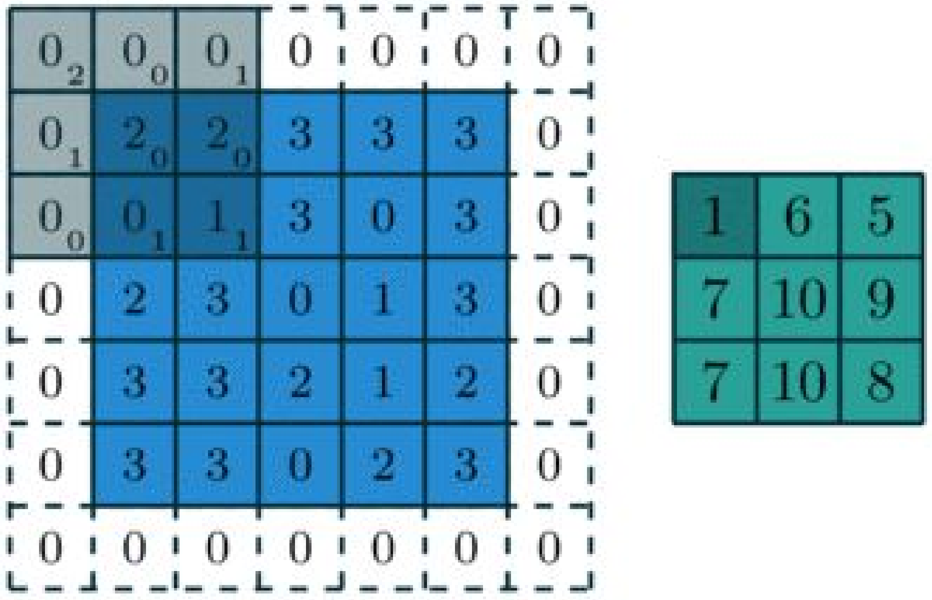
2-layer Neural Network $\longrightarrow f = W_2 max(0, W_1 x)$

3-layer Neural Network $\longrightarrow f = W_3 max(0, W_2 max(0, W_1 x))$

CNN Model - Convolution



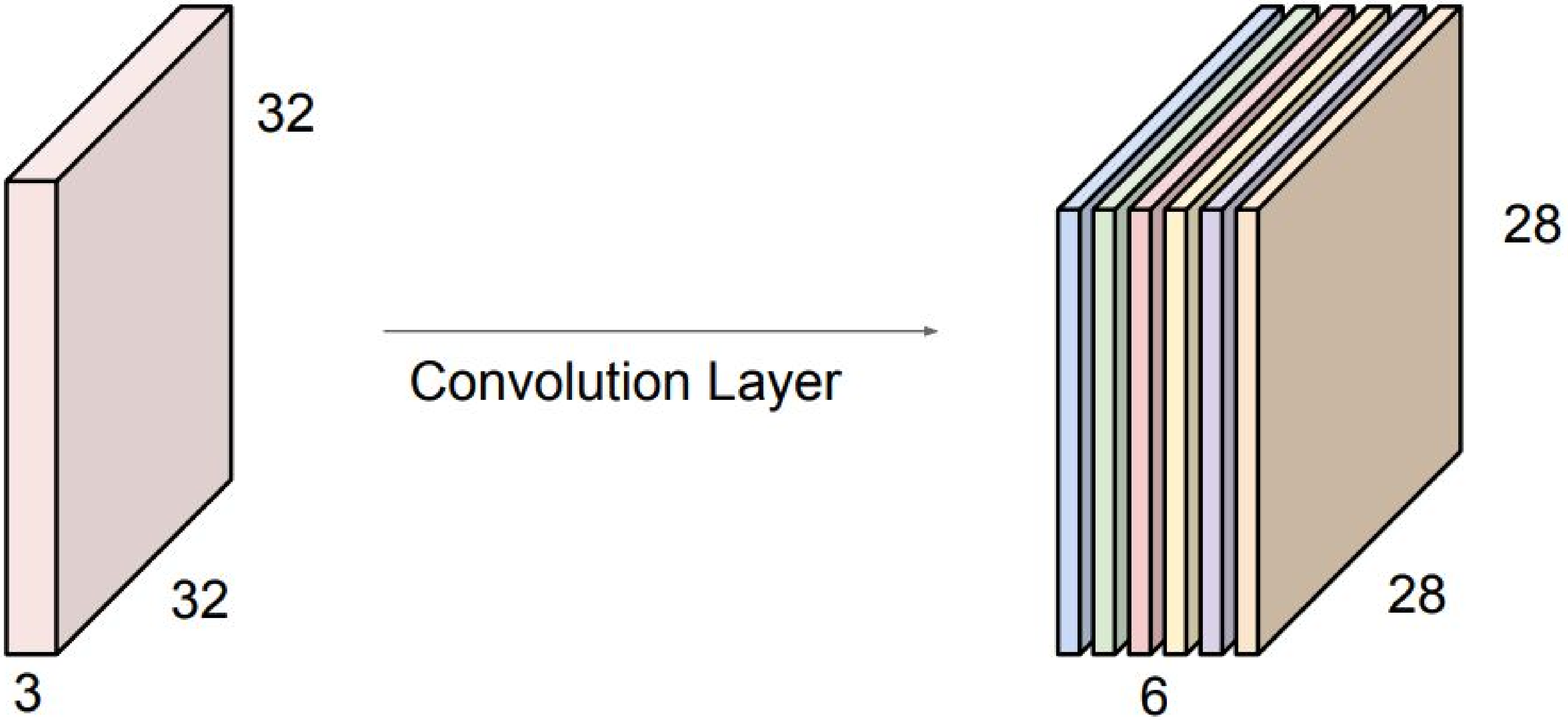
CNN Model - Convolution



Convolution operation with zero padding

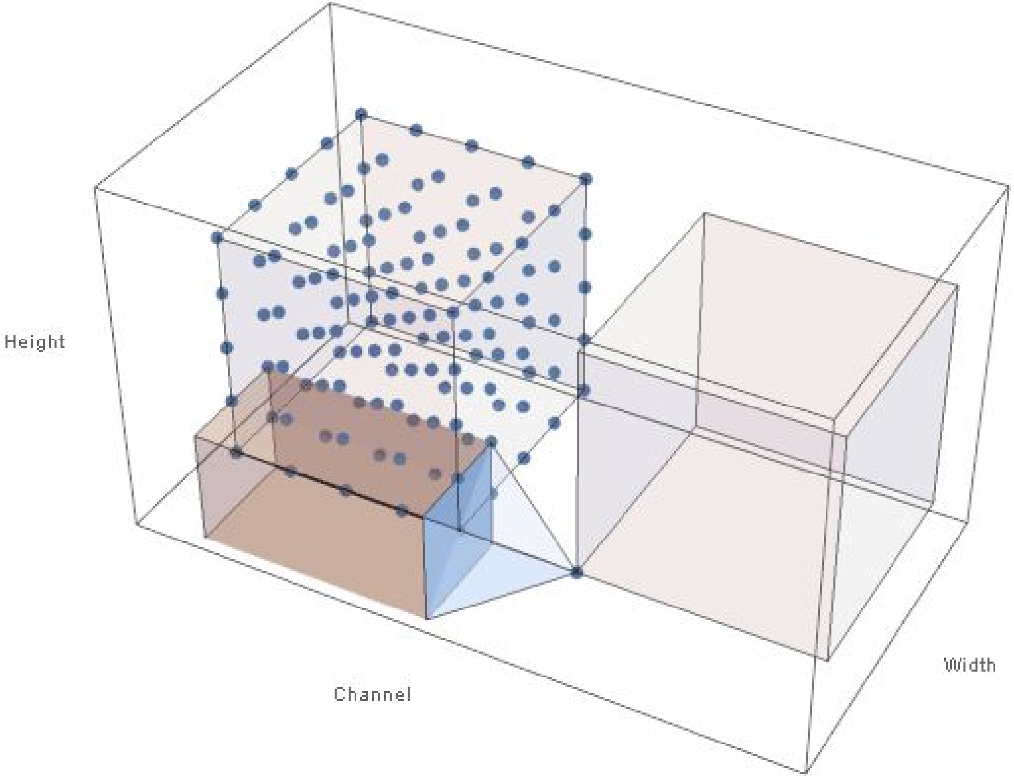
deeplearning.net/software/theano/tutorial/conv_arithmetic.html

CNN Model - Convolutional Layer



http://cs231n.stanford.edu/slides/2017/

CNN Model - Convolutional Layer



nicholastsmith.wordpress.com/2017/04/01/image-classification-using-convolutional-neural-networks-in-tensorflow/

CNN Model - Convolutional Layer

(MxM) : Original image size

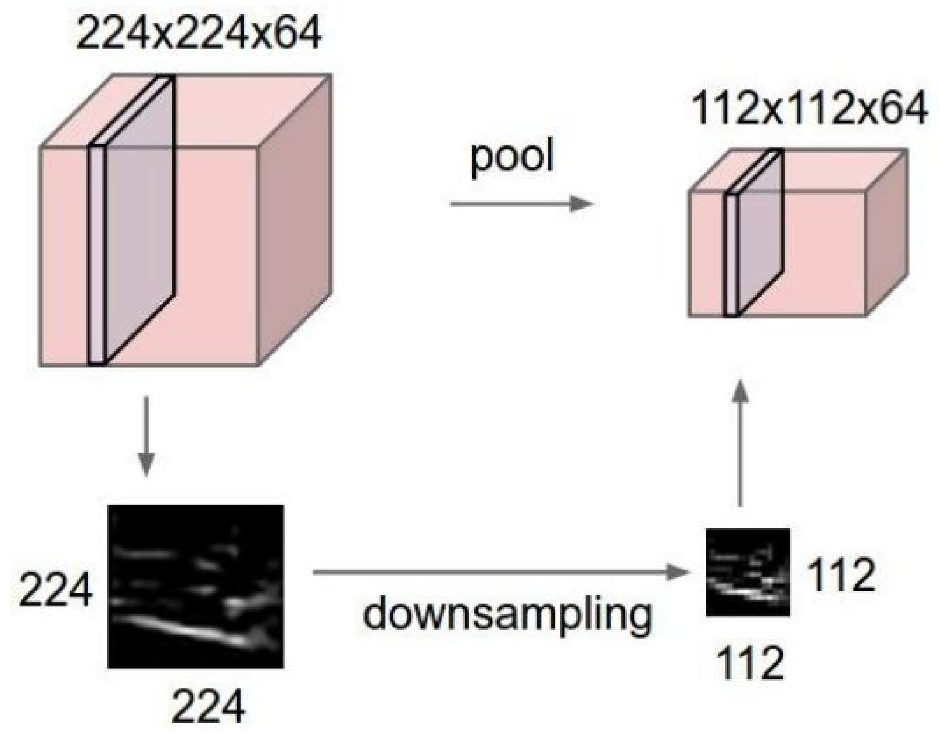
(KxK) : Filter size

(NxN) : Output image size

s : stride, step size of the filter

$$N = (M - K)/s + 1$$

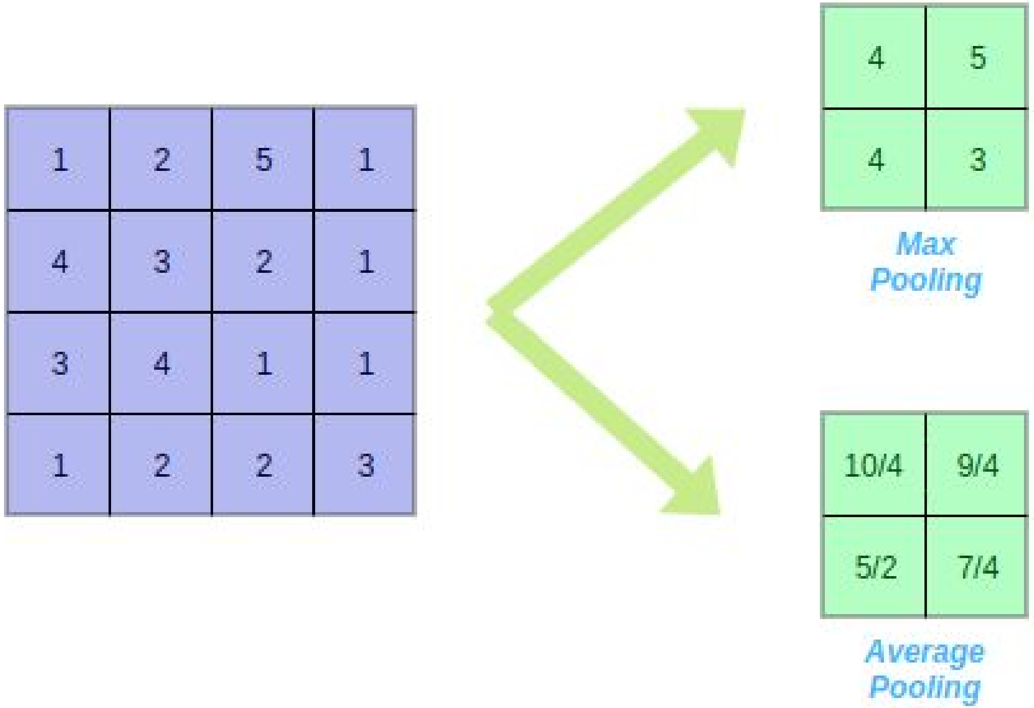
CNN Model - Pooling Layer



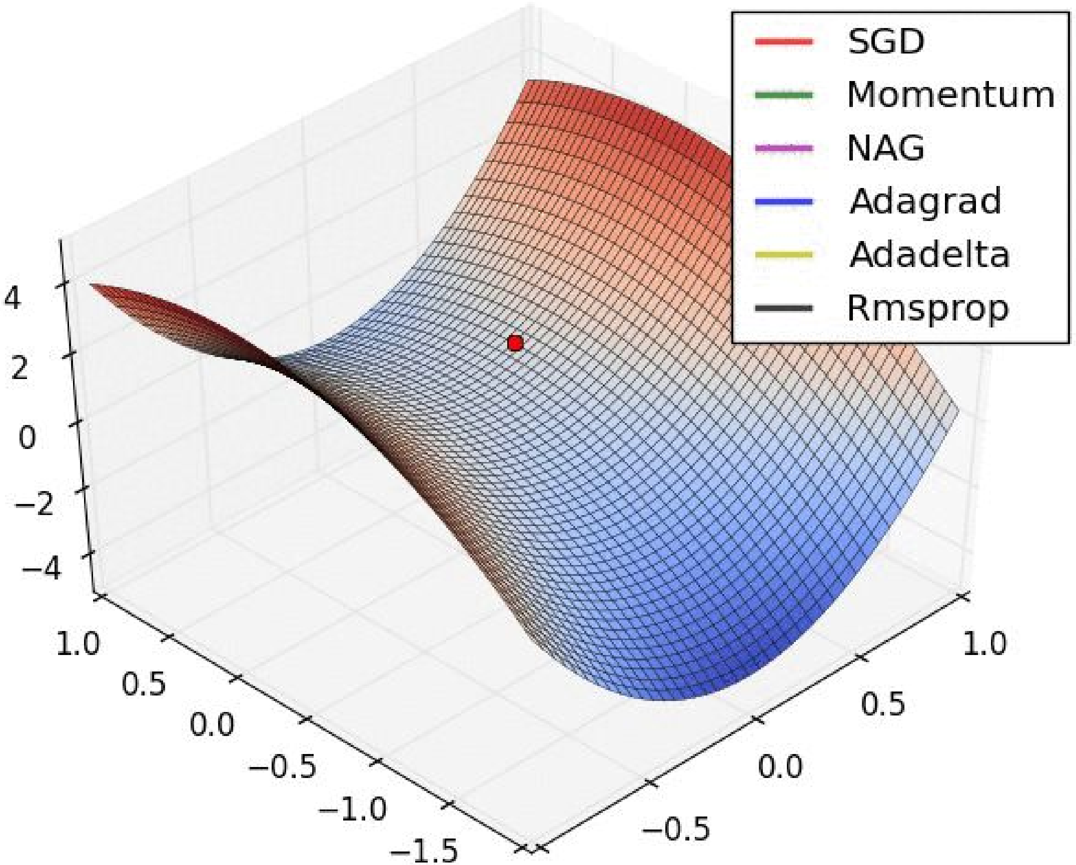
http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture5.pdf

CNN Model - Pooling Layer

Stride: 2



Optimization Methods



Transfer Learning

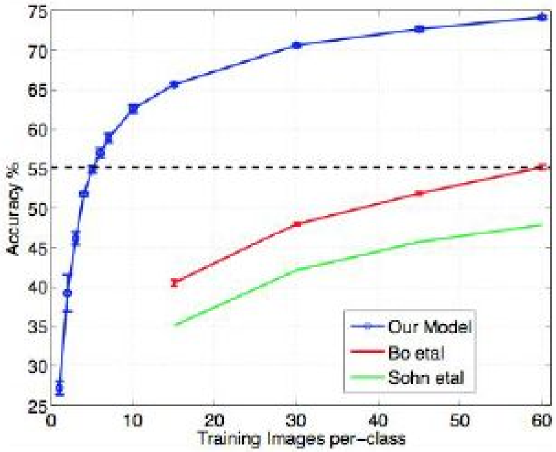
When to Fine-tune ?

A good first step

- More robust optimization - good initialization helps
- Needs less data
- Faster learning

State-of-the-art results in

- recognition
- detection
- segmentation



[Zeiler-Fergus]

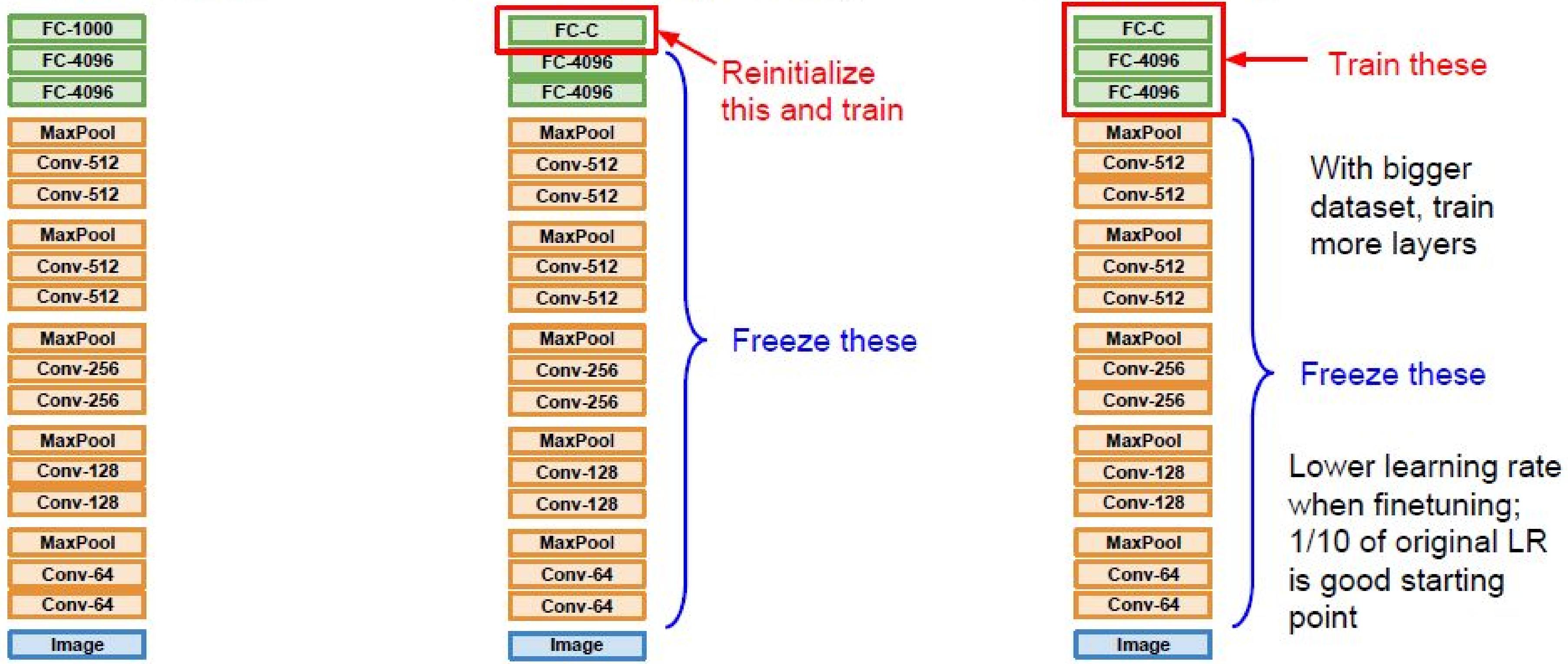
Transfer Learning with CNNs

Donahue et al, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition", ICML 2014
Razavian et al, "CNN Features Off-the-Shelf: An Astounding Baseline for Recognition", CVPR Workshops 2014

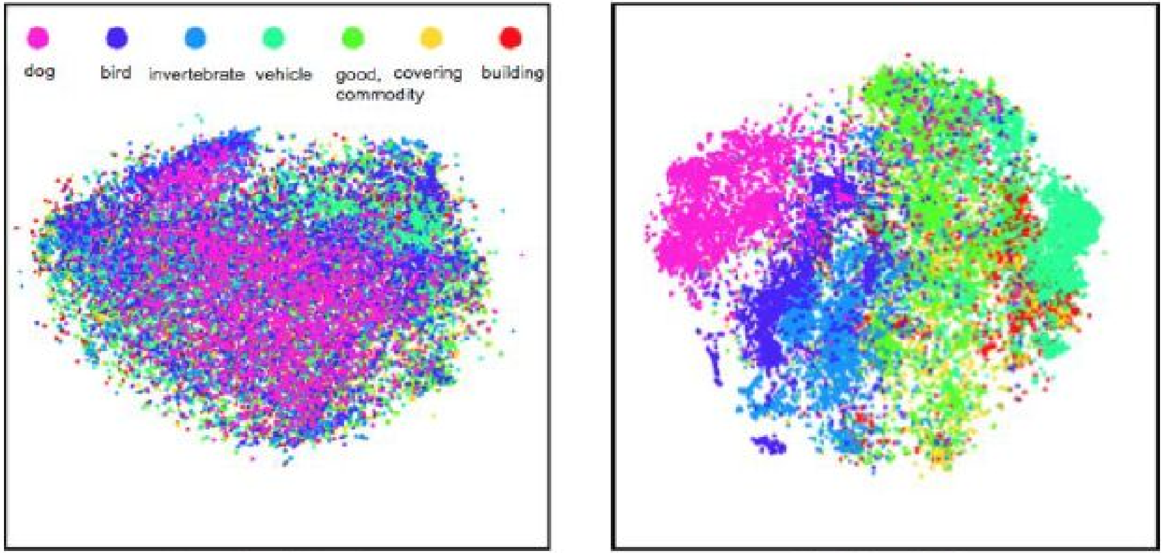
1. Train on Imagenet

2. Small Dataset (C classes)

3. Bigger dataset



Transfer Learning - Feature Extraction



Low-level: $Pool_1$ High-level: FC_6

Classes separate in the deep representations and transfer to many tasks.
[DeCAF] [Zeiler-Fergus]

Take Home Messages

- First try fine-tuning before attempting to build and train a deep neural network from scratch.
- Generally use ReLU as activation function.
- Use regularization to produce more generable models.
- Transfer from as closer domain as possible.

Katıldığınız için Teşekkür Ederiz

İrem Eyiokur, Doğucan Yaman, Anıl Genç, Omid Abdollahi Aghdam

{eyiokur16, yamand16, genca16, abdollahi15}@itu.edu.tr