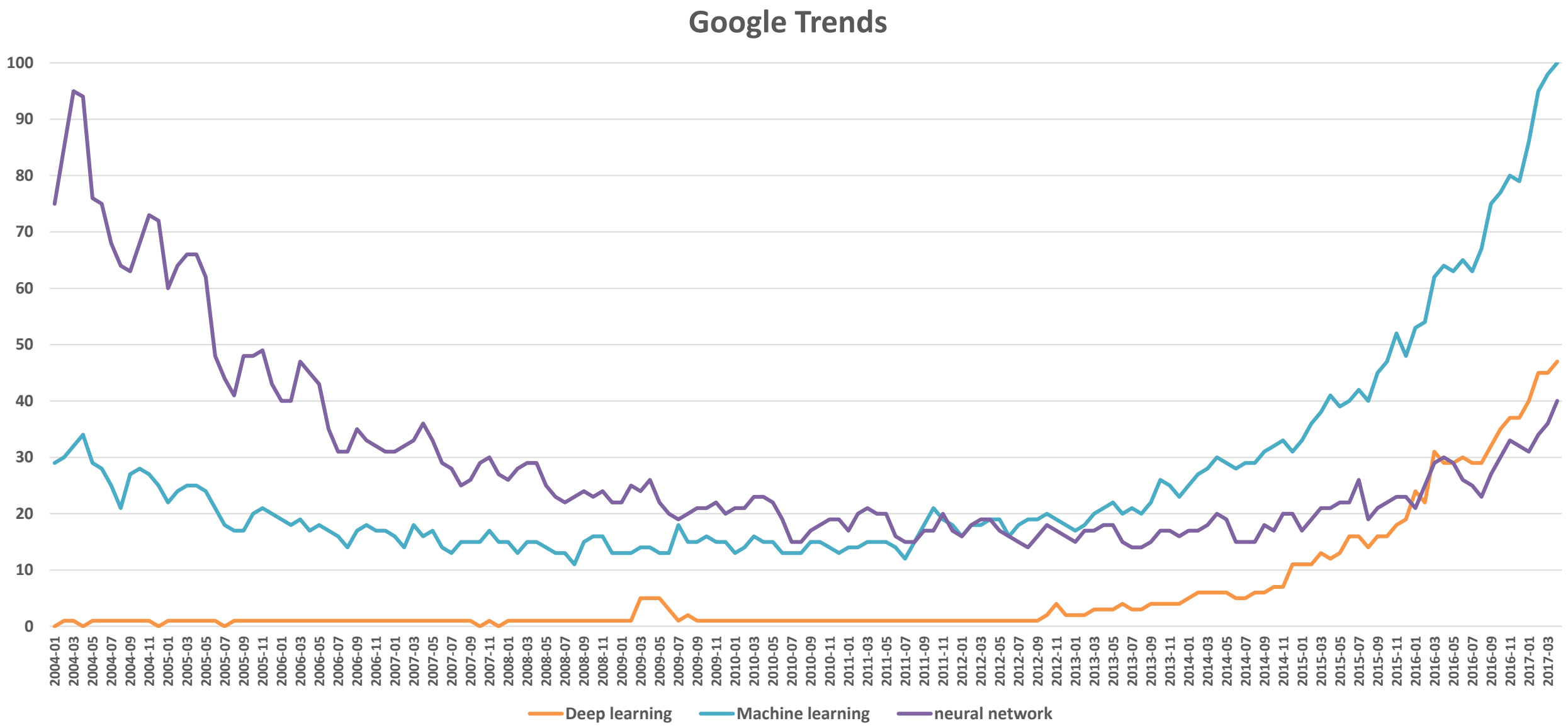
Advanced Topics in Computer Science Lecture 3

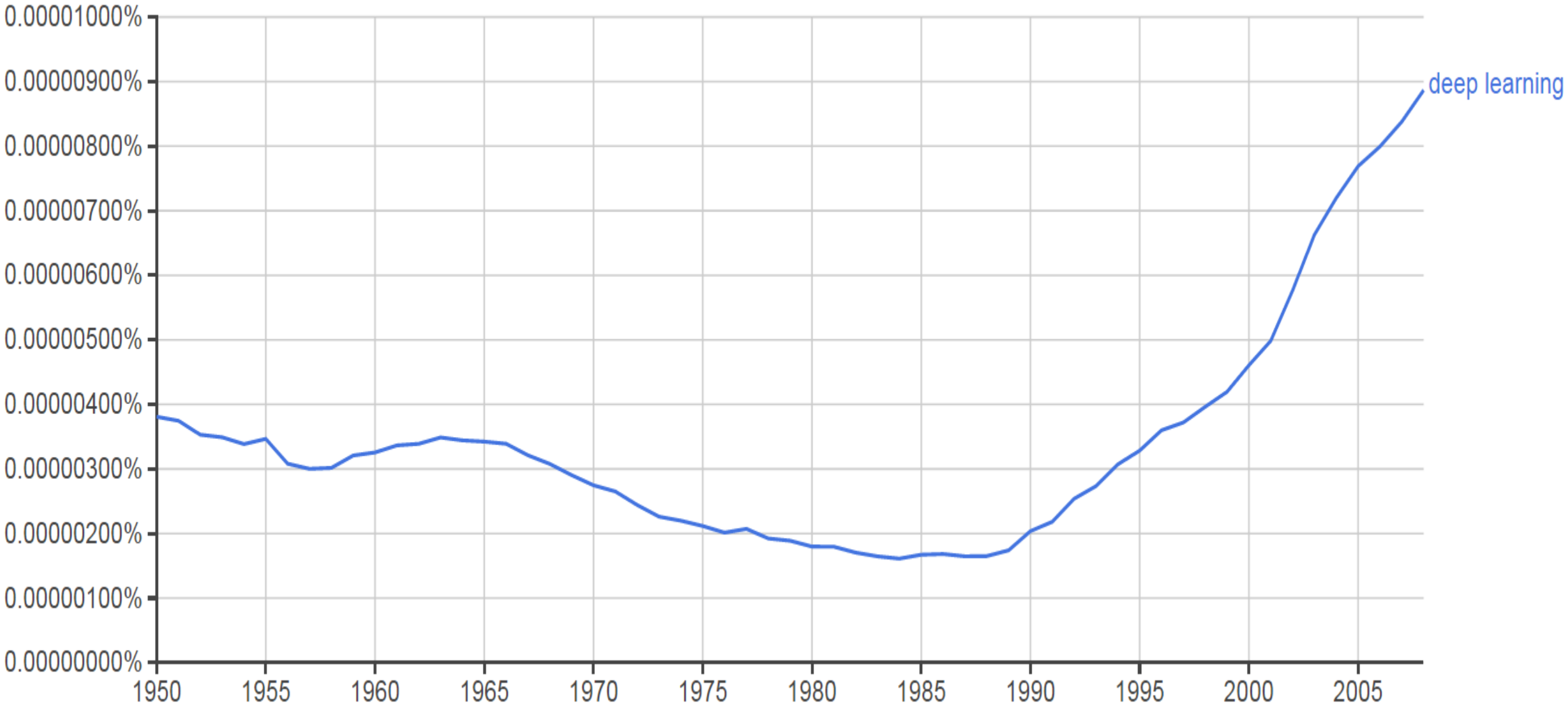
Convolutional Neural Networks

Dr Mohamed Loey
Lecturer, Faculty of Computers and Information Benha University
Egypt

Google Trends



Google NGRAM



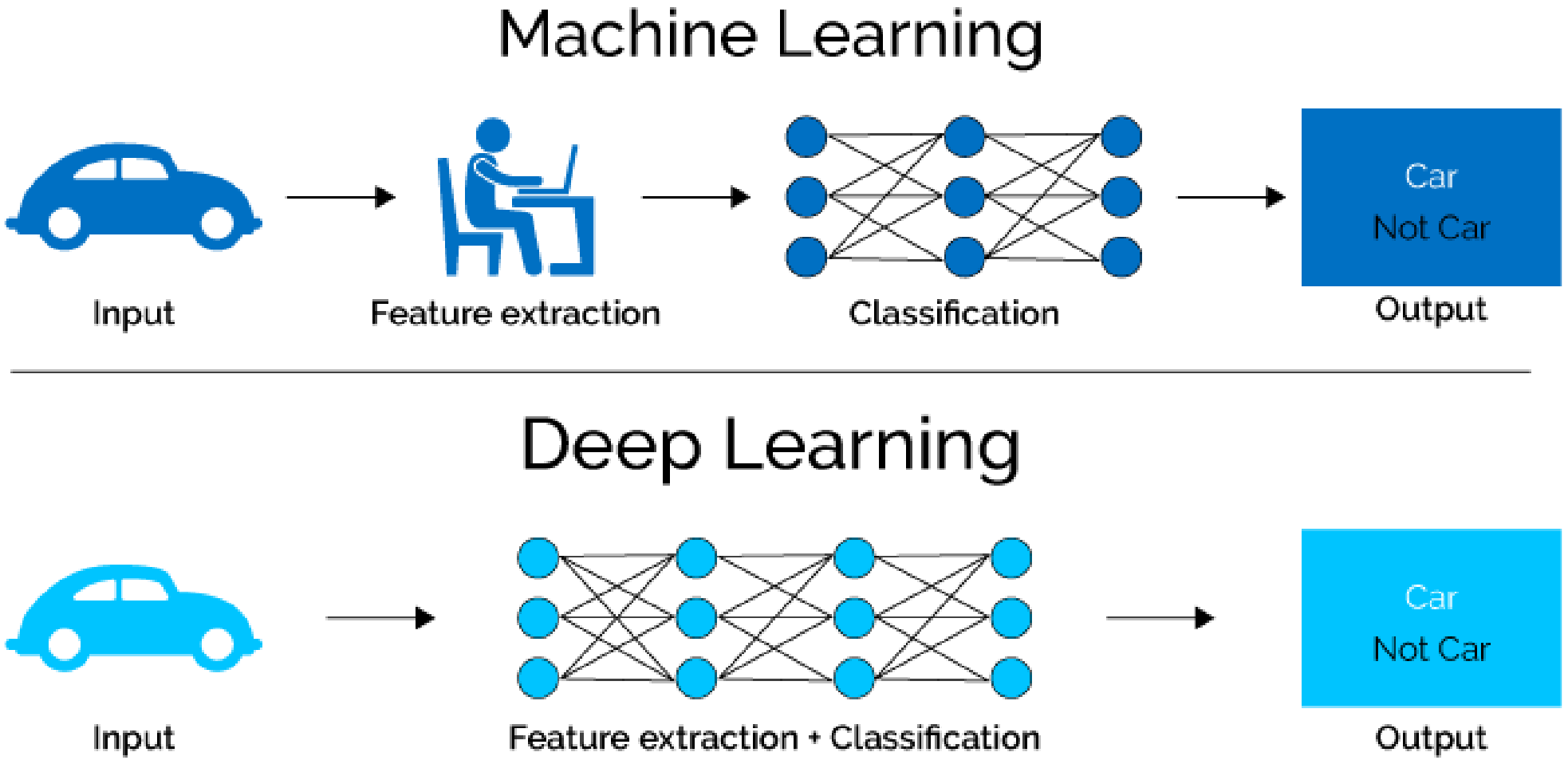
Google Queries

Queries	Top	Rising
deep machine learning	100	
deep learning network	85	
google deep learning	75	
deep learning tutorial	75	
deep neural networks	60	
neural networks	60	
deep neural network	60	

What is Deep Learning?

- **Part of the machine learning** field of learning representations of data.
- **hierarchy of multiple layers** that mimic the neural networks of our brain
- **If** you provide the system **tons of information**, it begins to understand it and respond in useful ways.

Deep vs Machine Learning



Why we needs Deep Learning?

- SuperIntelligent Devices
- Best Solution for
  - image recognition
  - speech recognition
  - natural language processing
  - Big Data

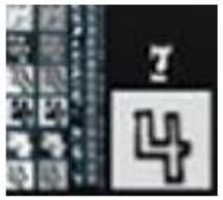
A brief History

1958 Perceptron

1969
Perceptron criticized

1974 Backpropagation

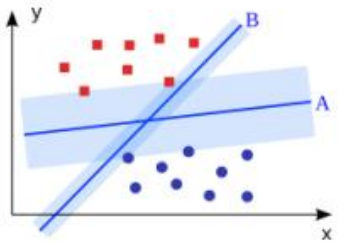
awkward silence (AI Winter)



Convolution Neural Networks for Handwritten Recognition

1998

1995
SVM reigns



2006
Restricted Boltzmann Machine

Google Brain Project on 16k Cores

2012

2012
AlexNet wins ImageNet

IMAGENET

Superstar Researchers

Geoffrey Hinton: University of Toronto & Google

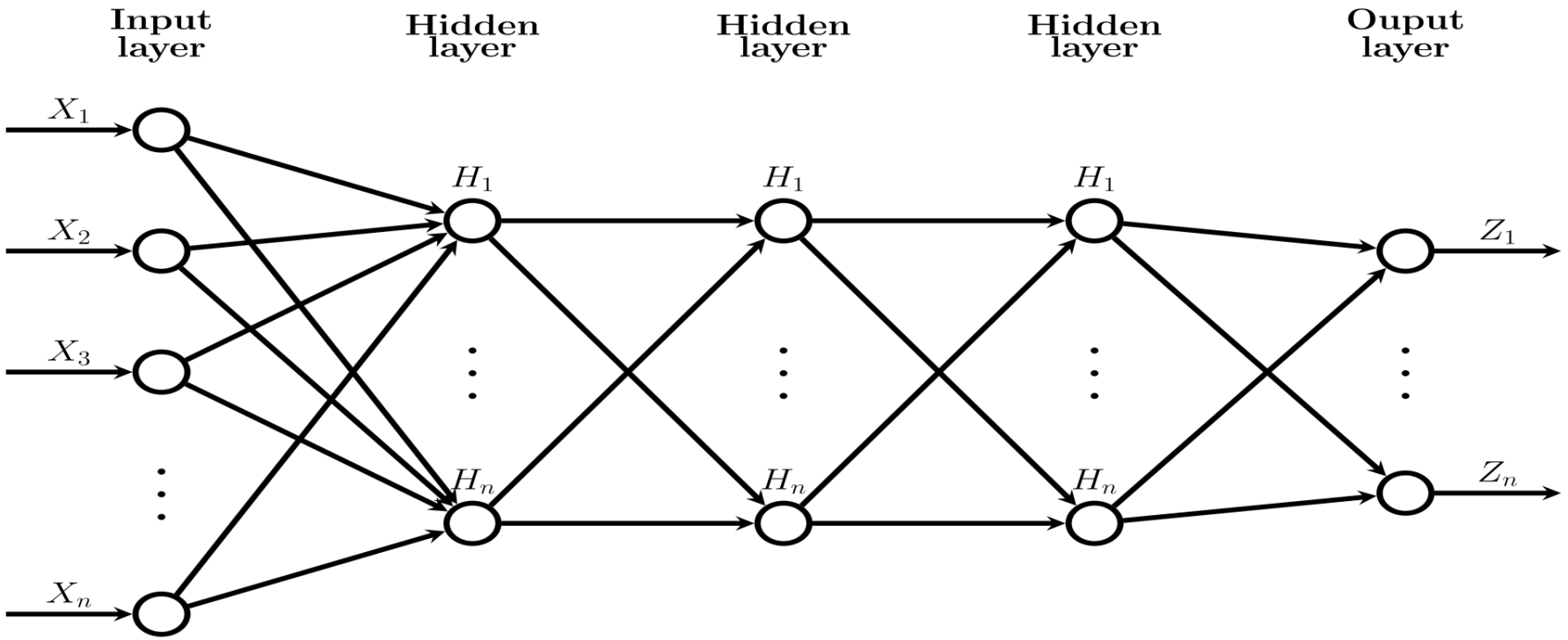
Yann LeCun: New York University & Facebook

Andrew Ng: Stanford & Baidu

Superstar Companies

Deep Learning

- Deep learning (DL) is a hierarchical structure network which through simulates the human brain's structure to extract the internal and external input data's features



Deep Learning Requirements

- Large data set with good quality
- Measurable and describable goals
- Enough computing power
- Neural Network (Brain of Human)

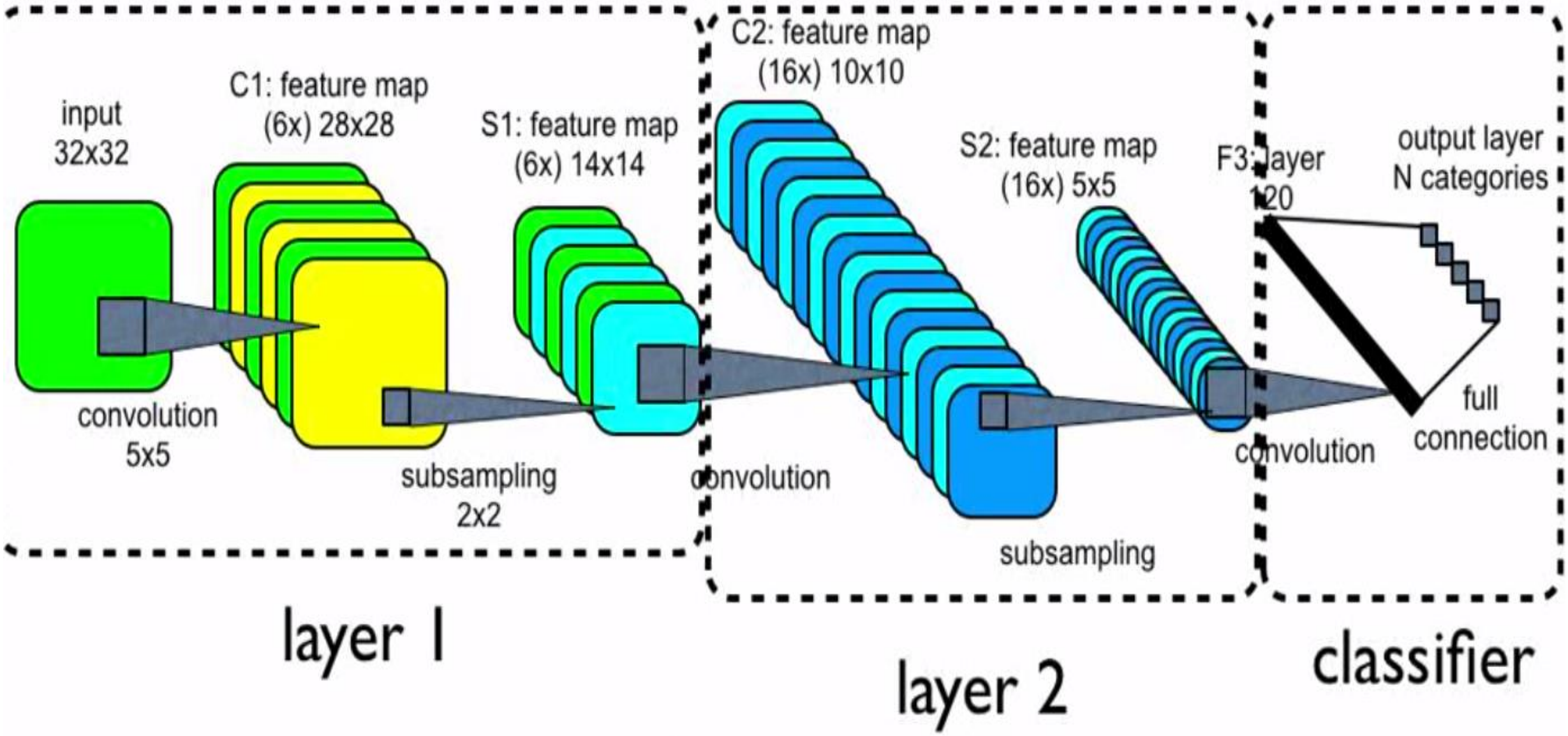
Deep Learning Architectures

- Deep Neural Networks
- Deep Belief Networks
- Convolutional Neural Networks
- Deep Boltzmann Machines
- Deep Stacking Networks

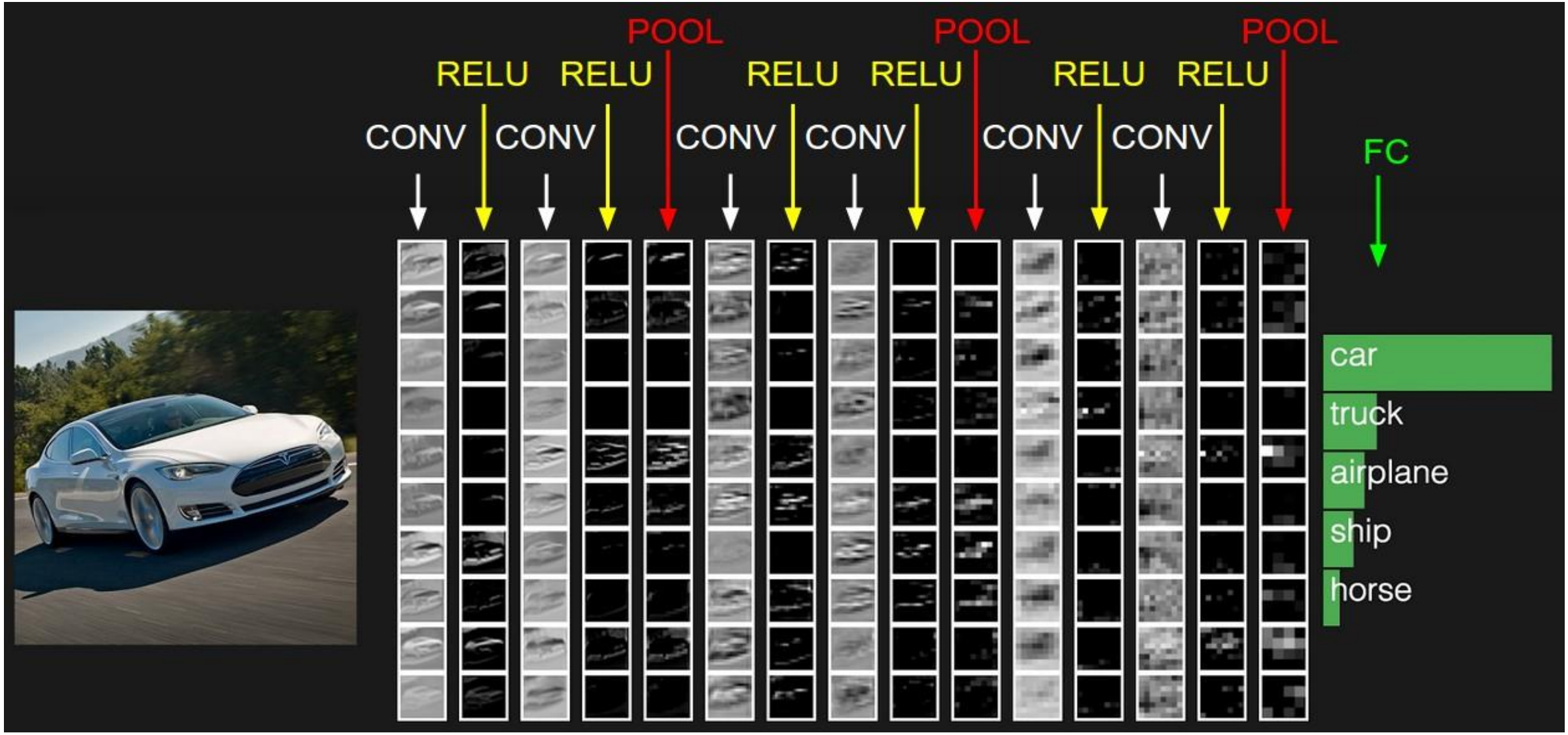
Convolution Neural Network

- **Convolution Neural Networks (CNN)** is supervised learning and a family of multi-layer neural networks particularly designed for use on two dimensional data, such as images and videos.
- A **CNN** consists of a number of layers:
  - Convolutional layers.
  - Pooling Layers.
  - Fully-Connected Layers.

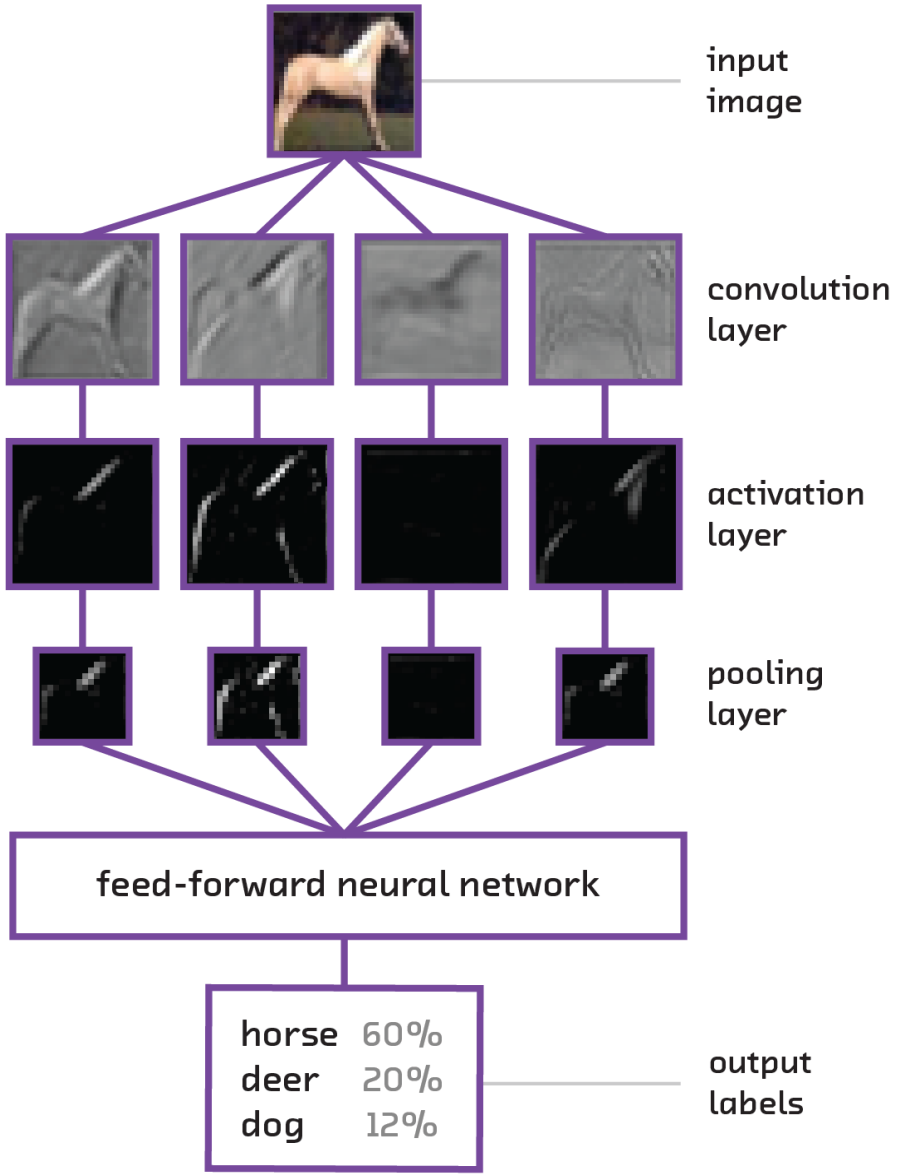
Convolution Neural Network



Convolution Neural Network

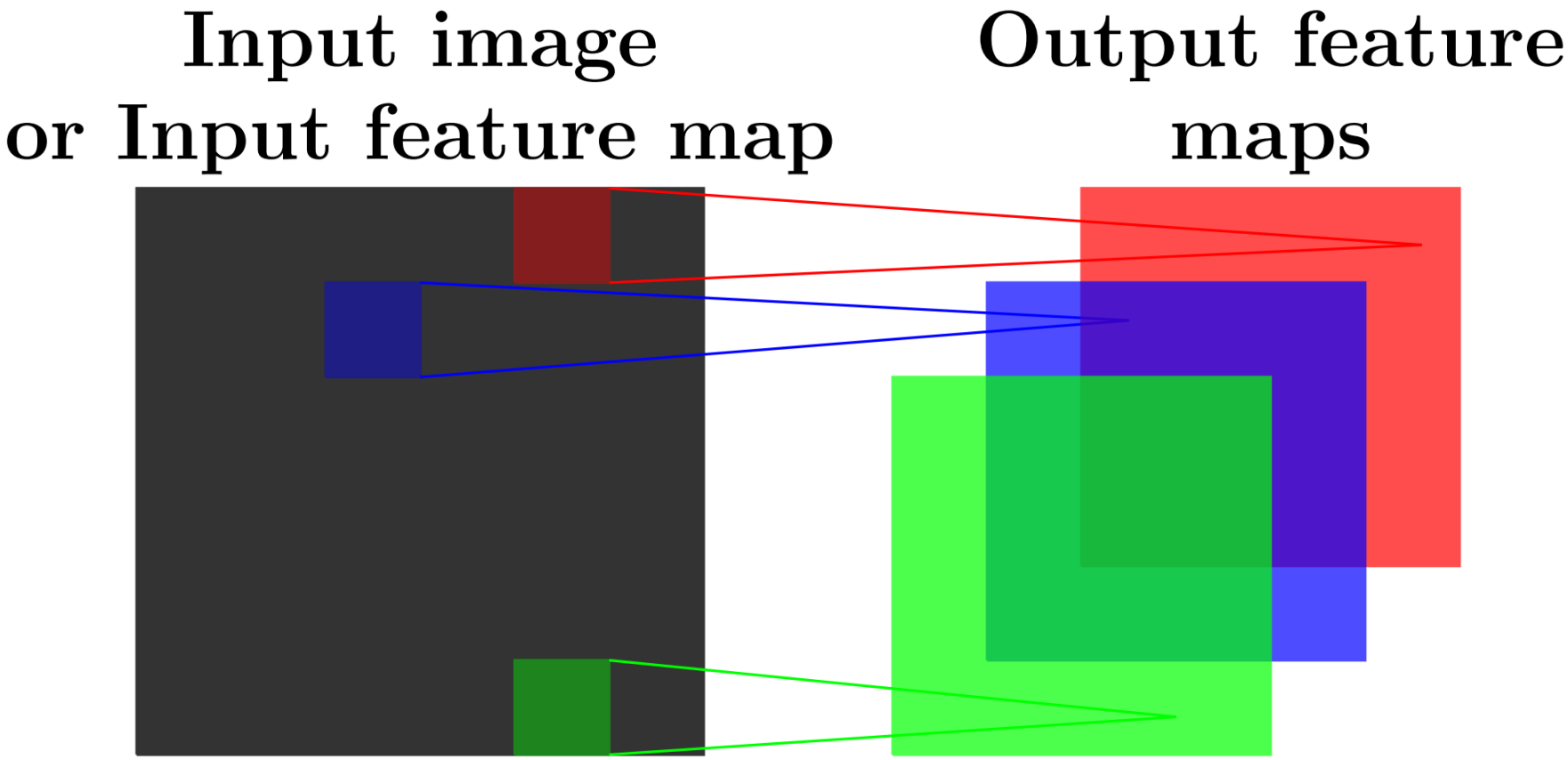


Convolution Neural Network

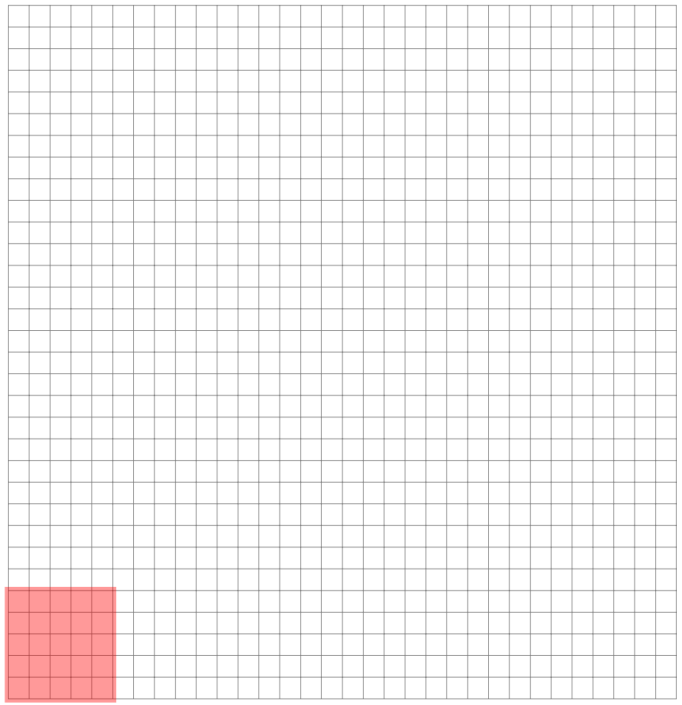


Convolutional layers

- Convolutional layer acts as a feature extractor that extracts features of the inputs such as edges, corners , endpoints.

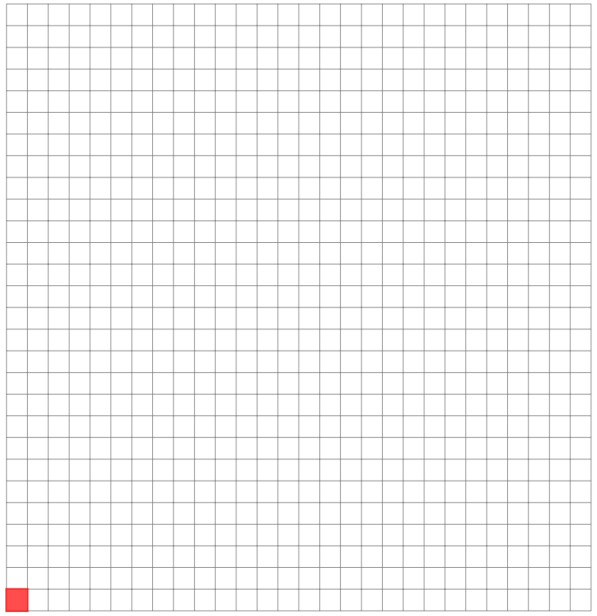


Convolutional layers



$*$

$=$

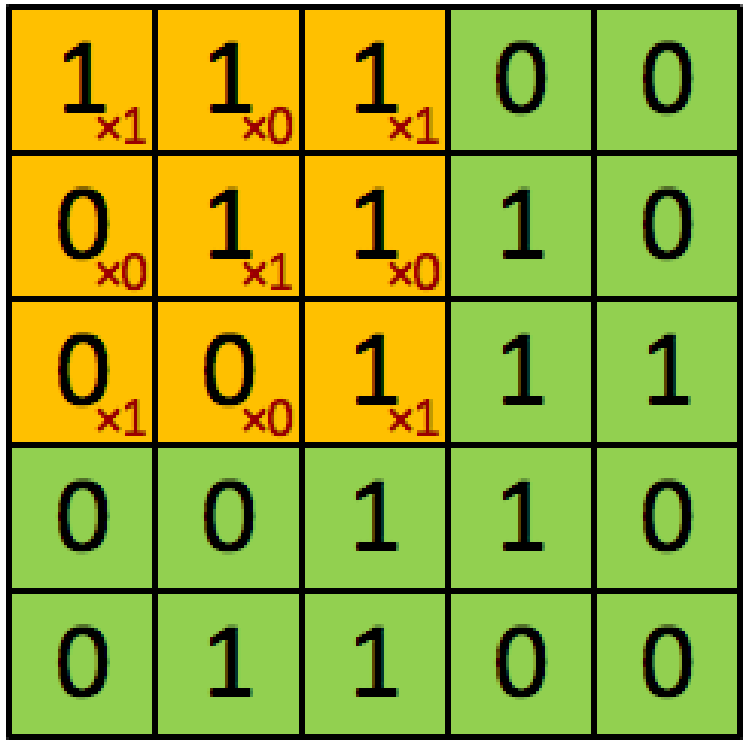


Kernal $N \times N$

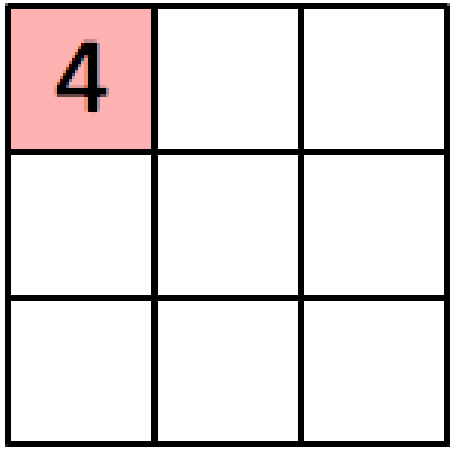
Input $M \times M$

CONV Layer $M - N + 1 \times M - N + 1$

Convolutional layers

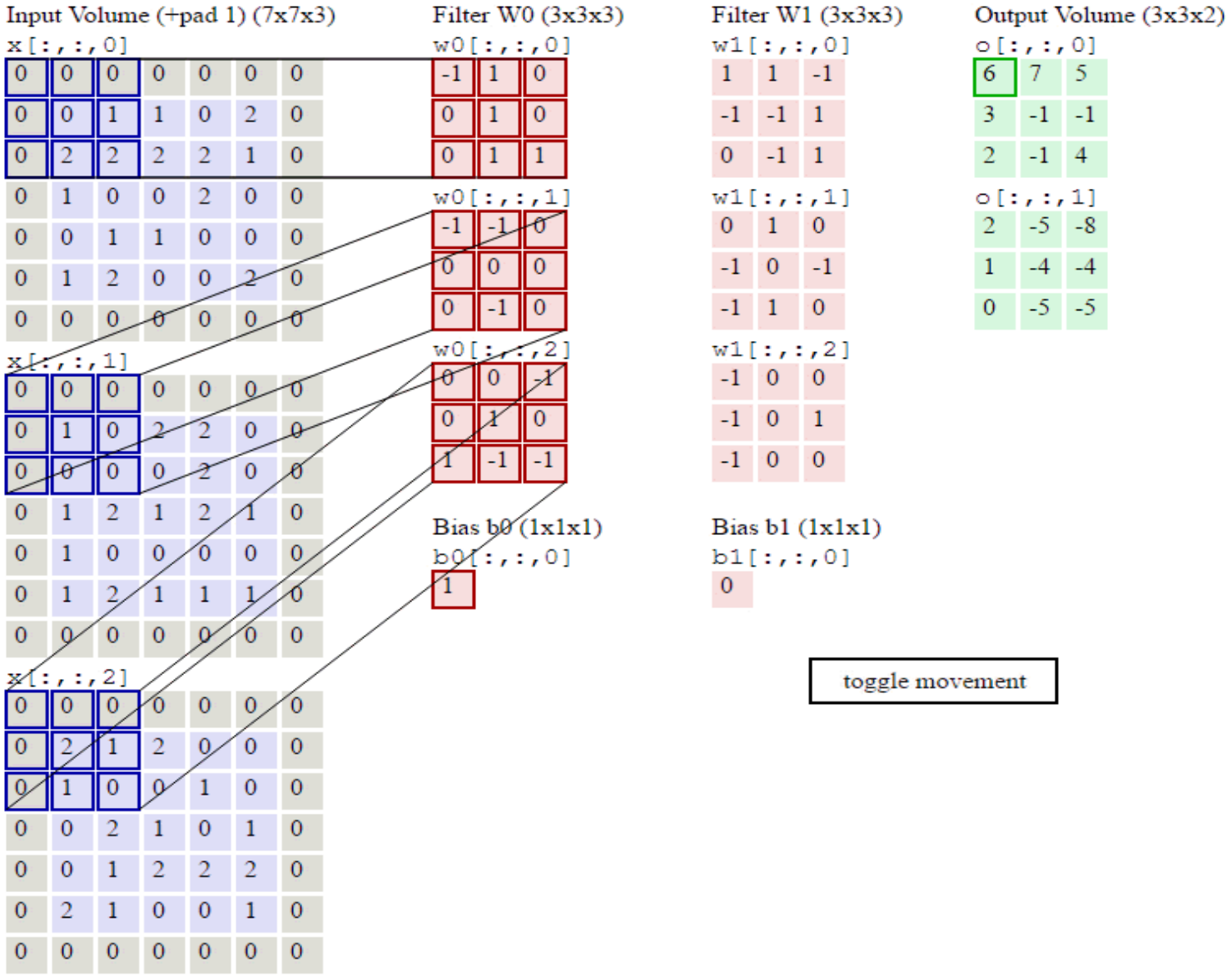


Image



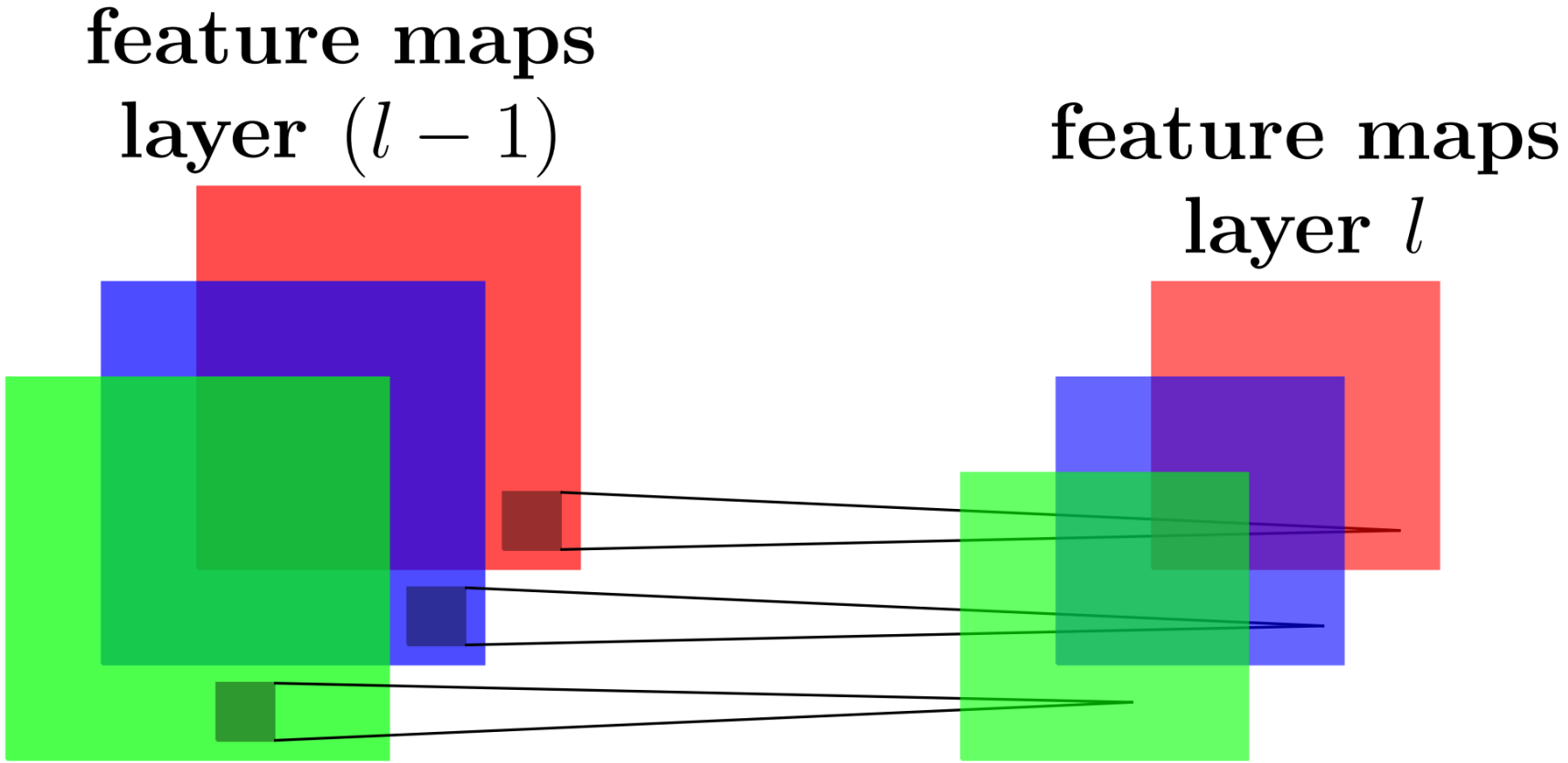
Convolved Feature

Convolutional layers

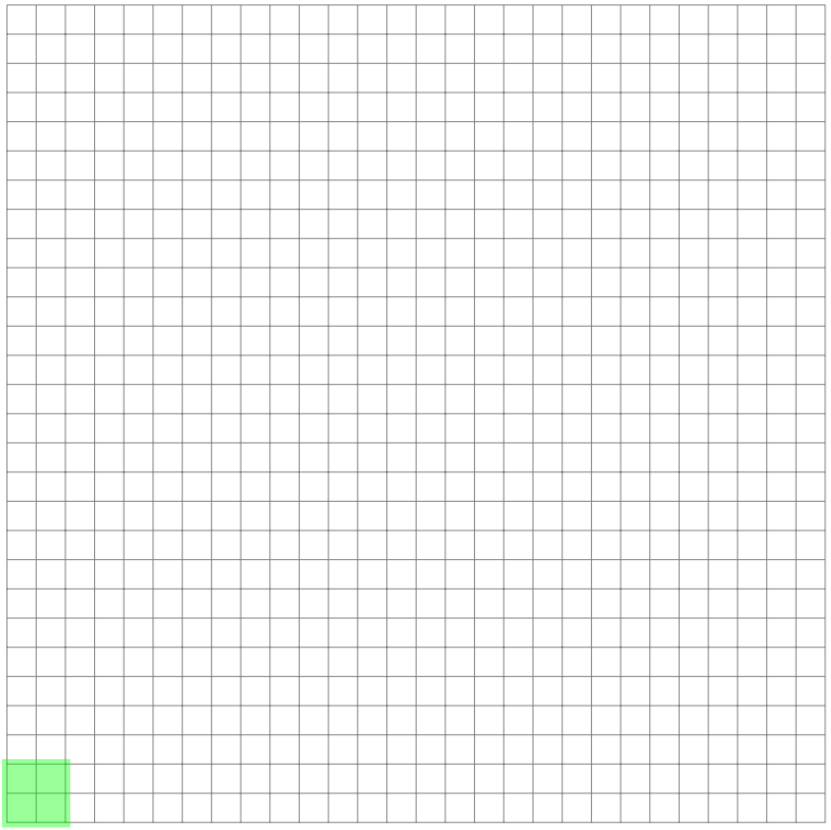


Pooling layers

- The pooling layer reduces the resolution of the image that reduce the precision of the translation (shift and distortion) effect.

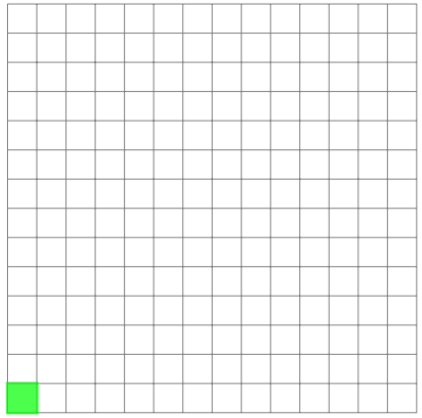


Pooling layers



Feature map

$$M - N + 1 \times M - N + 1$$



POOL Layer

$$\frac{M-N+1}{2} \times \frac{M-N+1}{2}$$

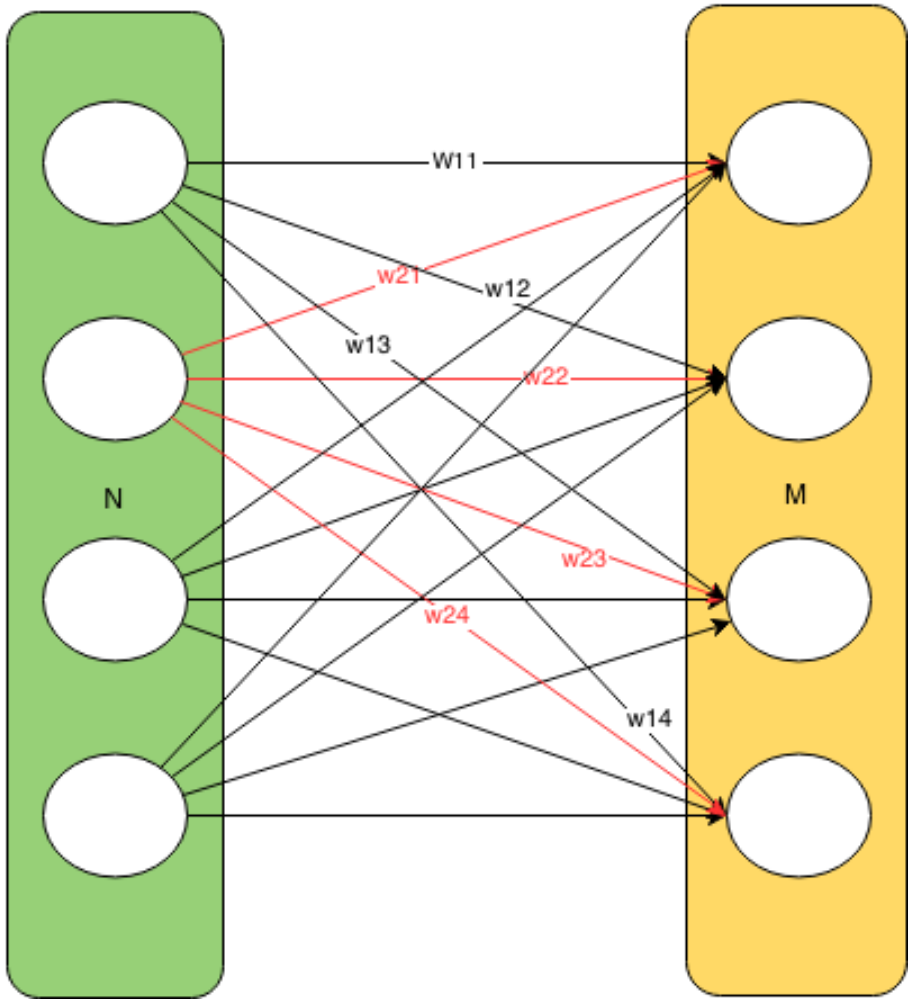
Pooling layers

1	3	2	9
7	4	1	5
8	5	2	3
4	2	1	4

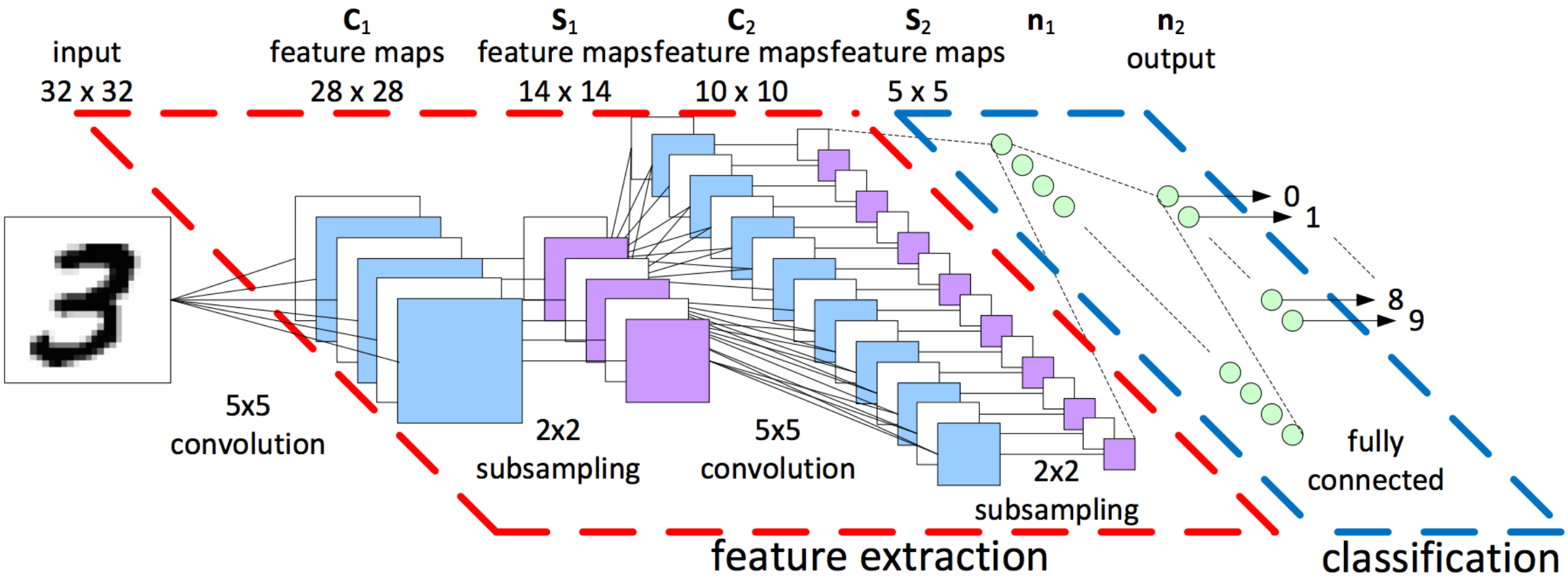
7	9
8	

Fully-Connected Layers

- **fully connected layer** have full connections to all activations in the previous layer.
- Fully connect layer act as classifier.



Convolution Neural Network



Case studies

LeNet : The first successful applications of CNN

AlexNet: The **ILSVRC** 2012 winner

ZFNet: The **ILSVRC** 2013 winner

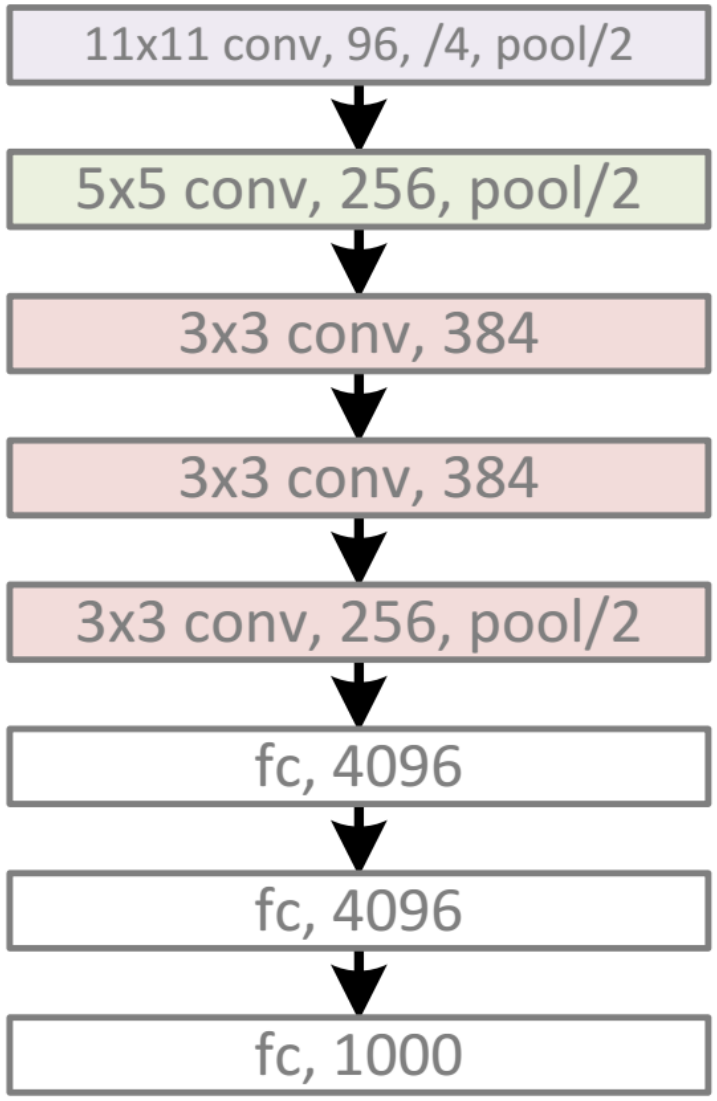
GoogLeNet: The **ILSVRC** 2014 winner

VGGNet: The runner-up in **ILSVRC** 2014

ResNet: The winner of **ILSVRC** 2015

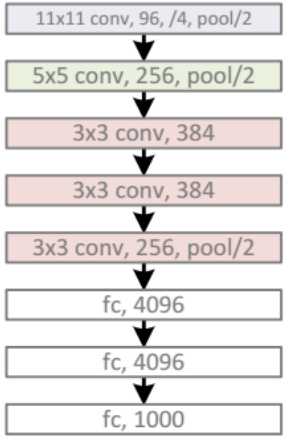
AlexNet

AlexNet, 8 layers
(ILSVRC 2012)

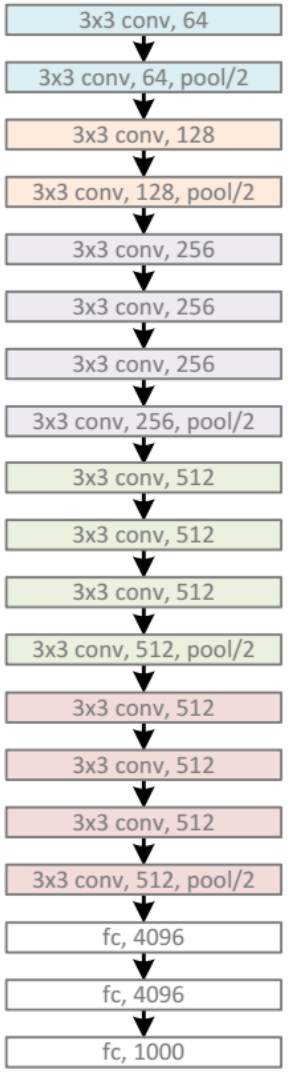


Revolution of Depth

AlexNet, 8 layers
(ILSVRC 2012)



VGG, 19 layers
(ILSVRC 2014)



GoogleNet, 22 layers
(ILSVRC 2014)

Revolution of Depth

AlexNet, 8 layers
(ILSVRC 2012)

VGG, 19 layers
(ILSVRC 2014)

ResNet, 152 layers
(ILSVRC 2015)

Datasets Benchmark

MNIST Handwritten digits – 60000 Training + 10000 Test Data

Google House Numbers from street view - 600,000 digit images

CIFAR-10 60000 32x32 colour images in 10 classes

IMAGENET 1.2 million images, **>**150 GB

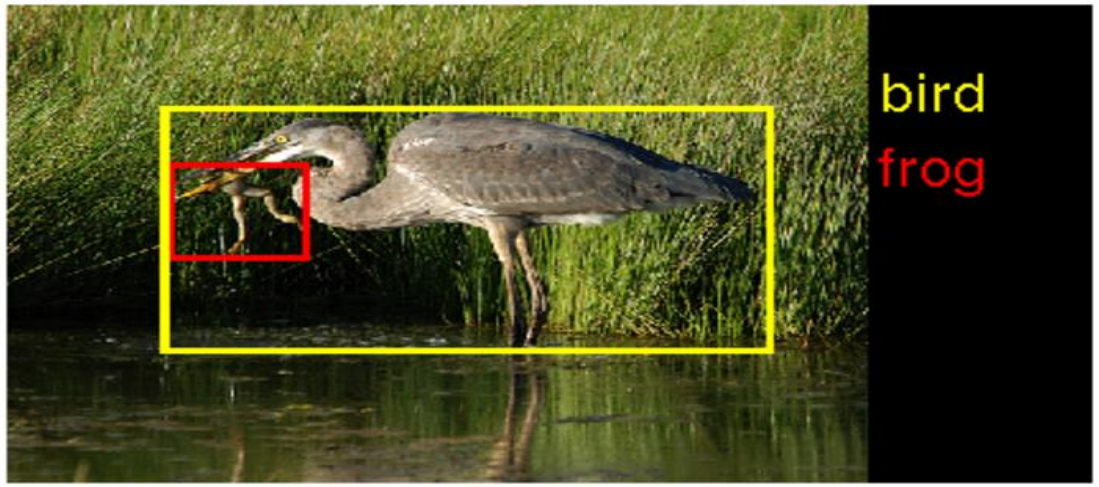
Tiny Images 80 Million tiny images

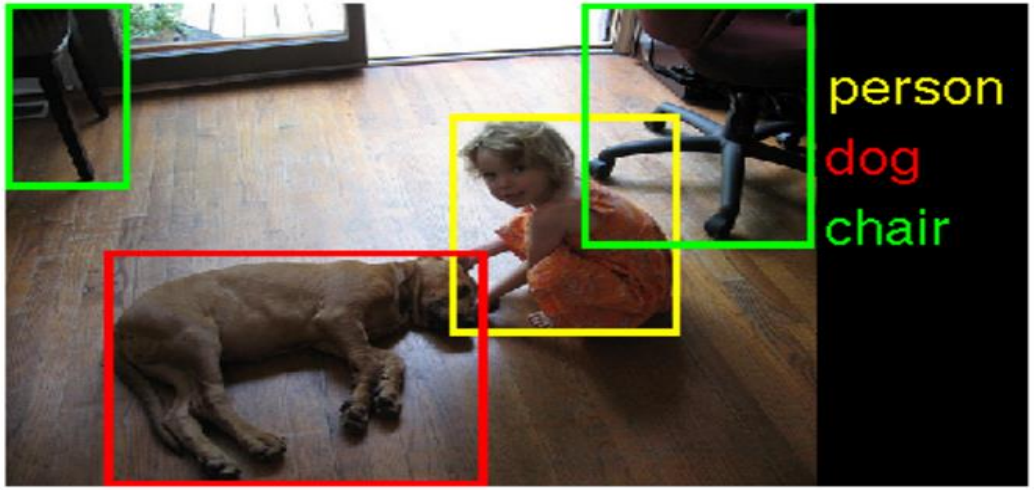
Flickr Data 100 Million Yahoo dataset

ILSVRC

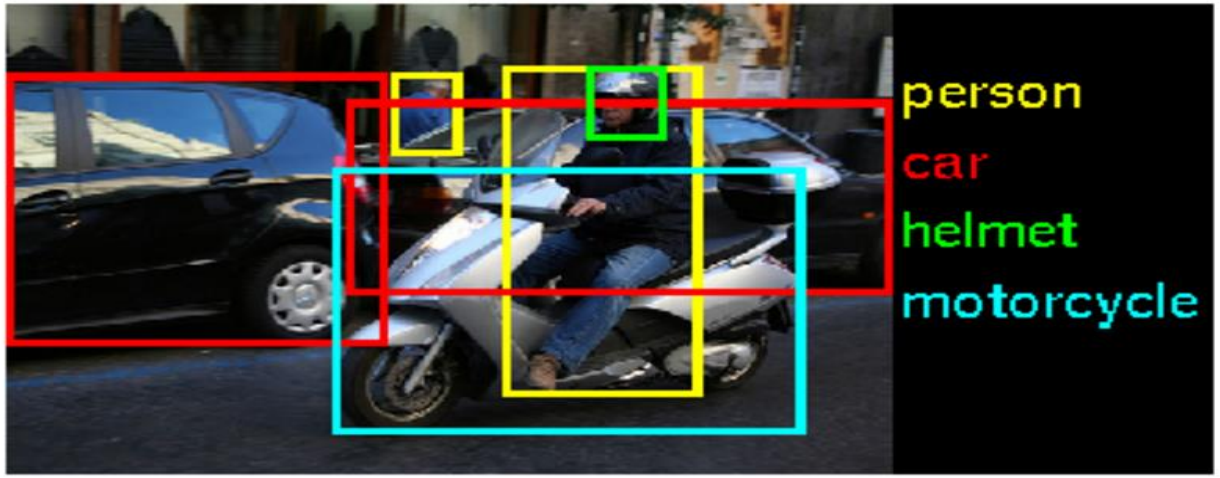
- **The ImageNet Large Scale Visual Recognition Challenge (ILSVRC)** evaluates algorithms for object detection and image classification at large scale.

ILSVRC 2014

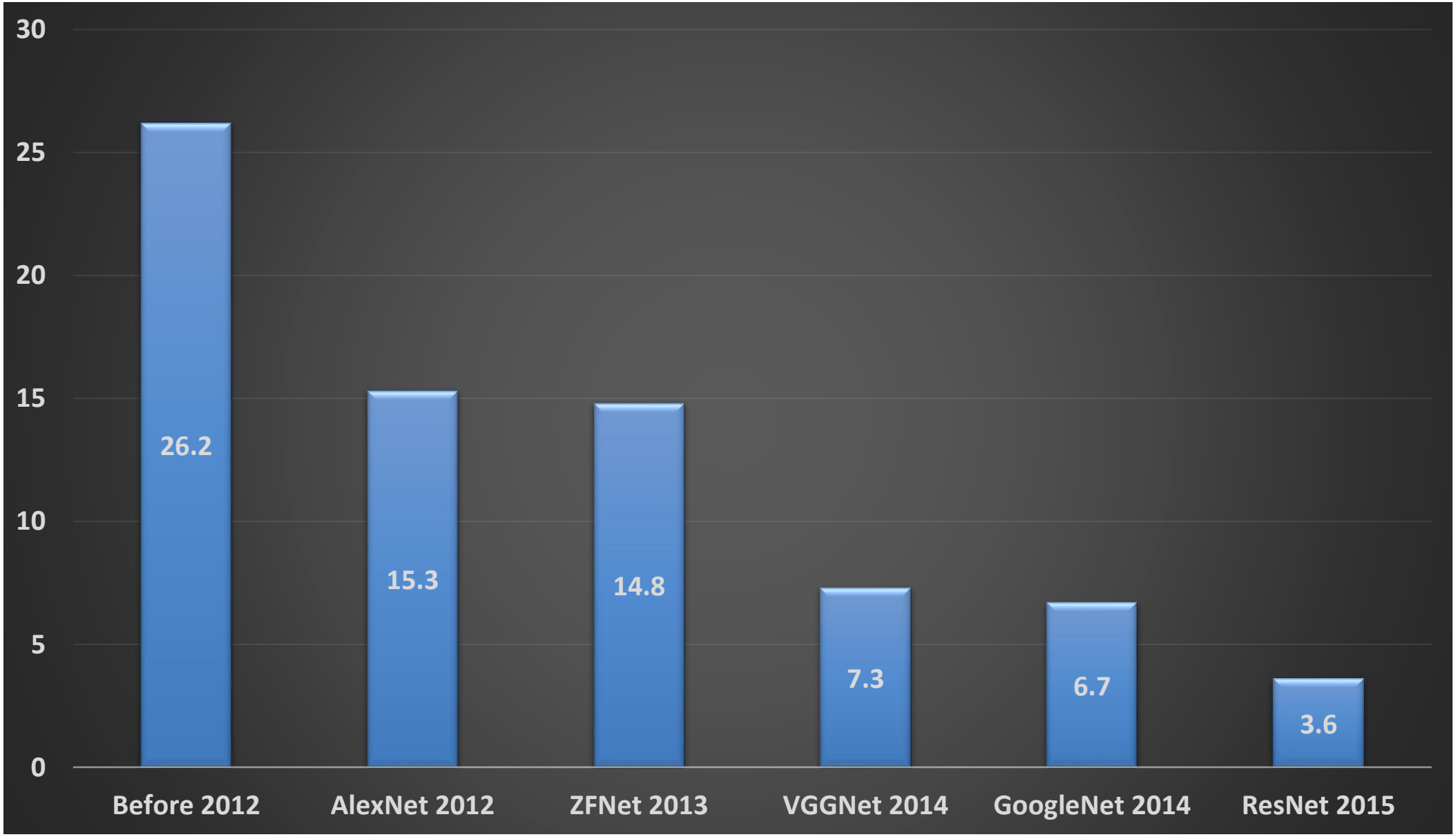

bird
frog


person
dog
chair


person
hammer
flower pot
power drill


person
car
helmet
motorcycle

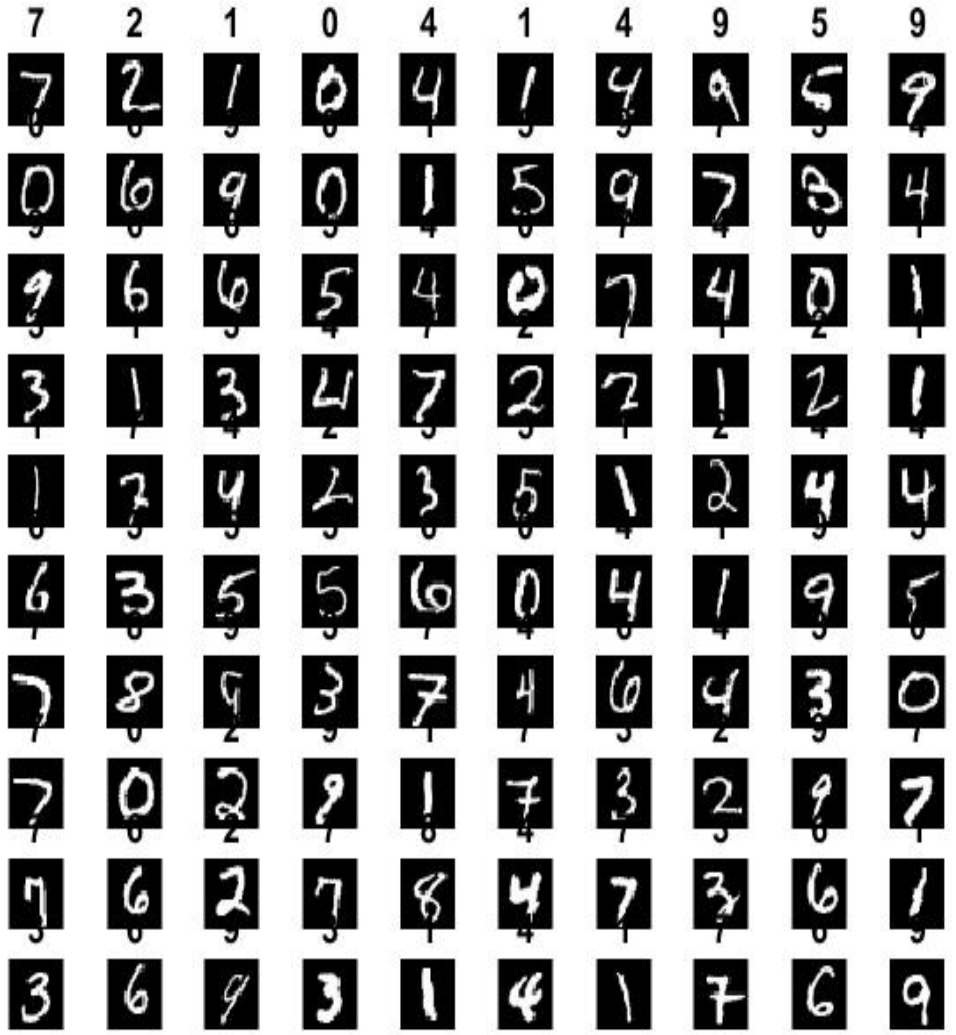
Top-5 error rates on ImageNet



MNIST Dataset

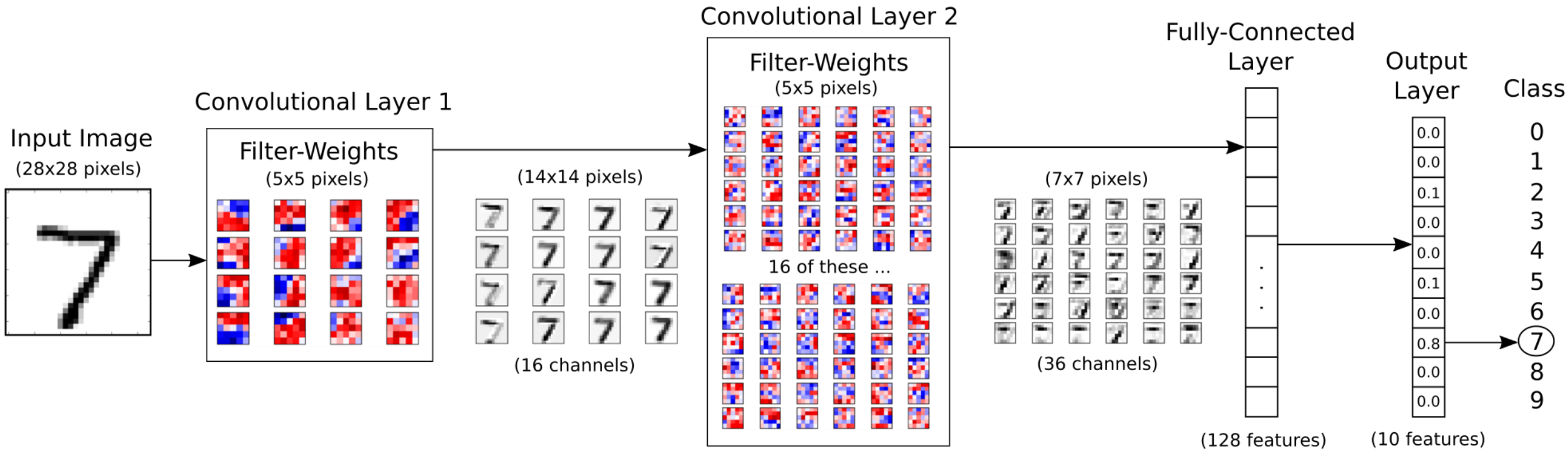
- MNIST is a large database of handwritten digits.

- MNIST contains 60,000 training images and 10,000 testing images



MNIST Dataset

❑ CNN on MNIST Dataset

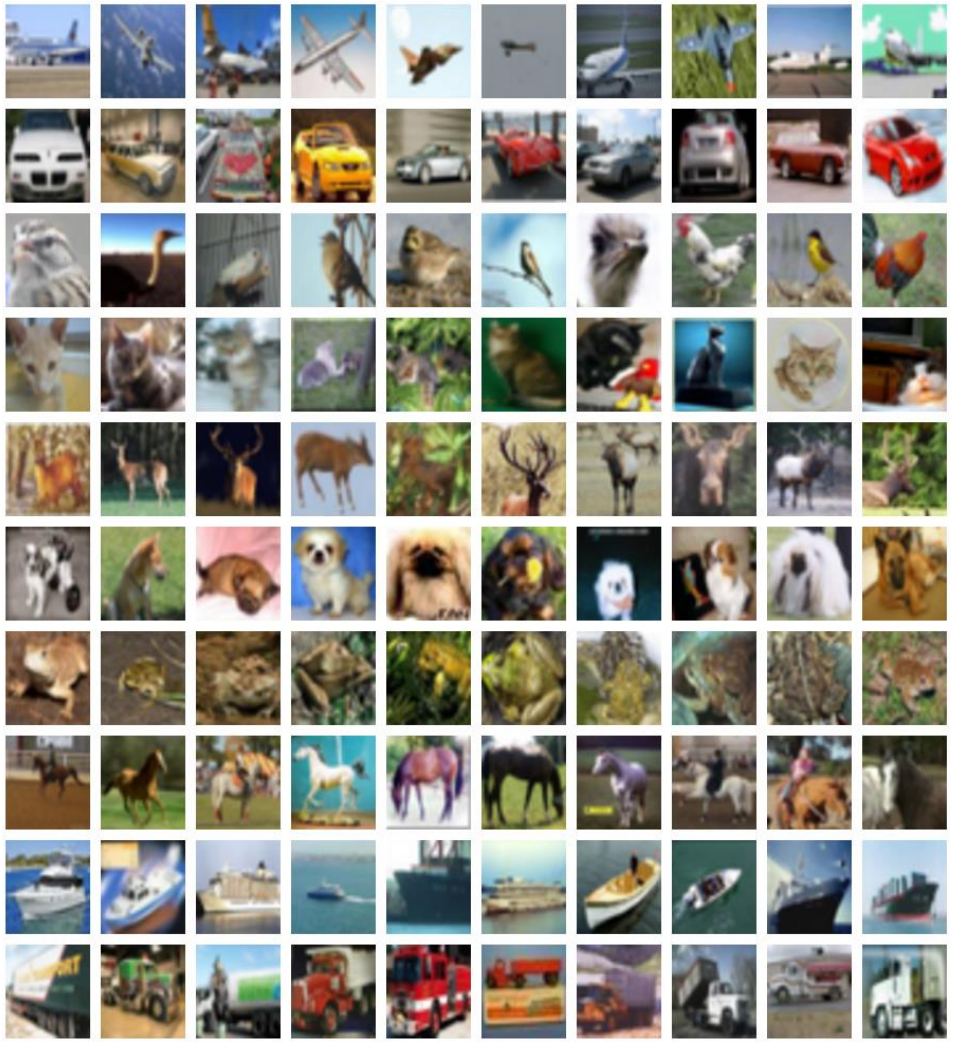


CIFAR-10 Dataset

- CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes

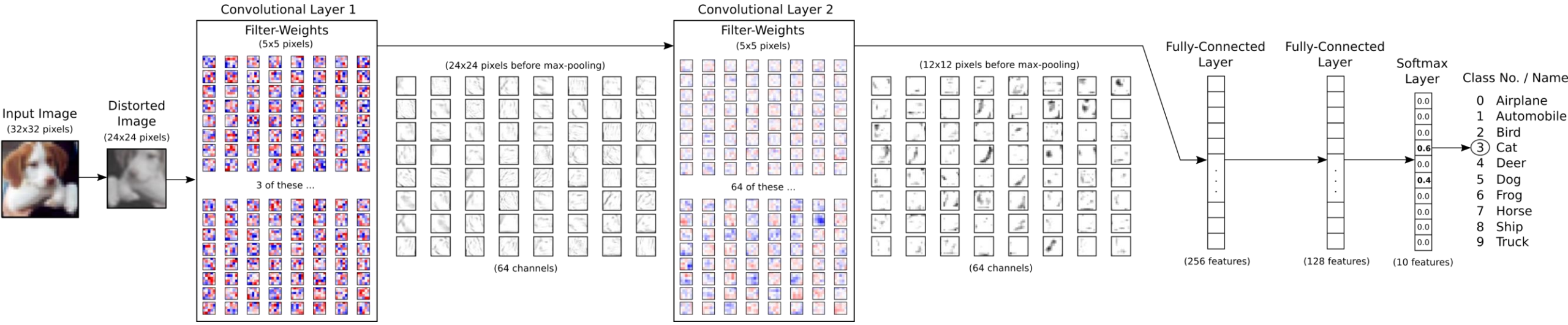
- CIFAR-10 contains 50000 training images and 10000 test images

airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck



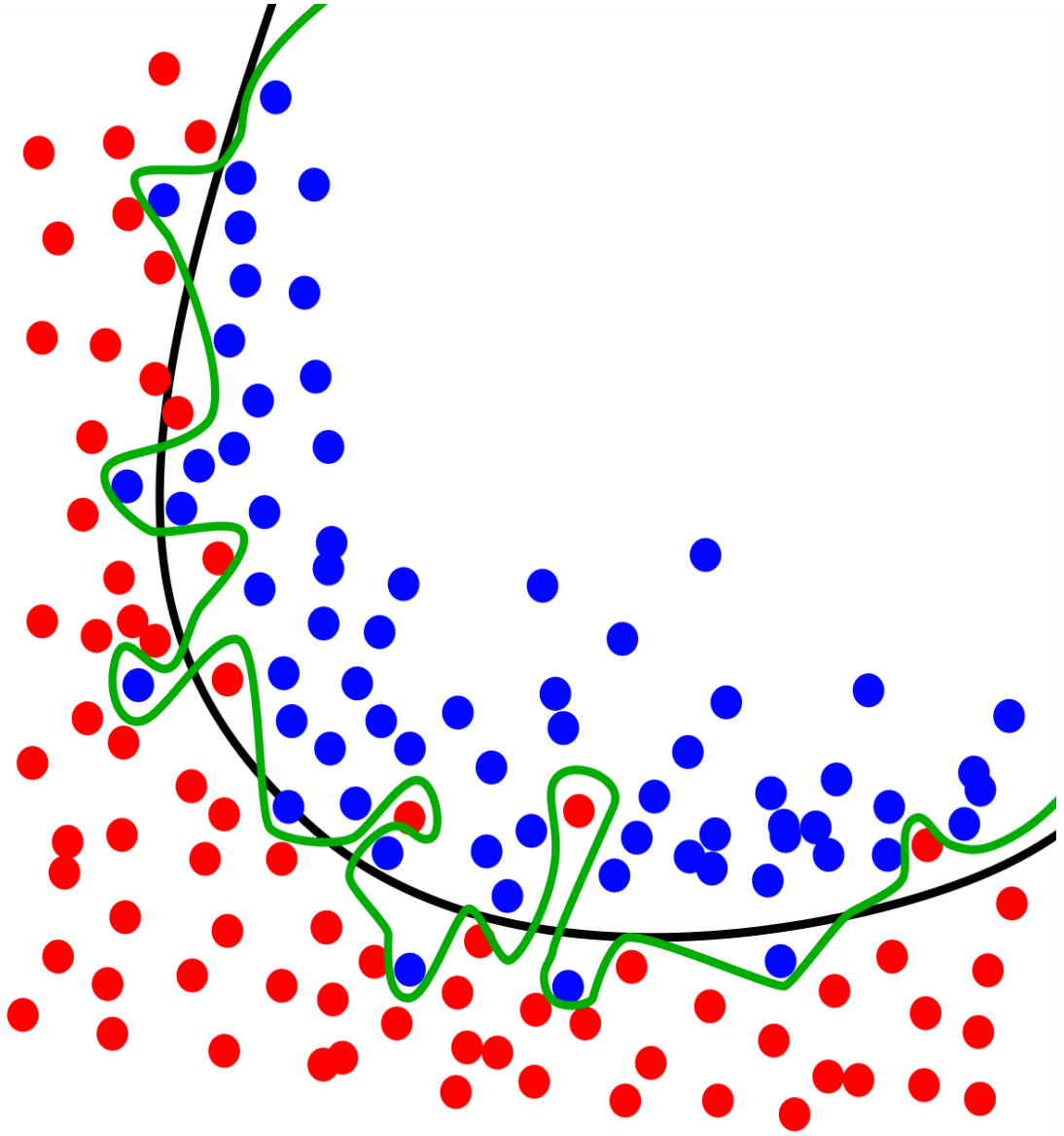
CIFAR-10 Dataset

- CNN on CIFAR-10 Dataset



CNN Optimization

- Overfitting Problem
  - Larger network have a lots of weights this lead to high model complexity
  - Network do excellent on training data but very bad on validation data



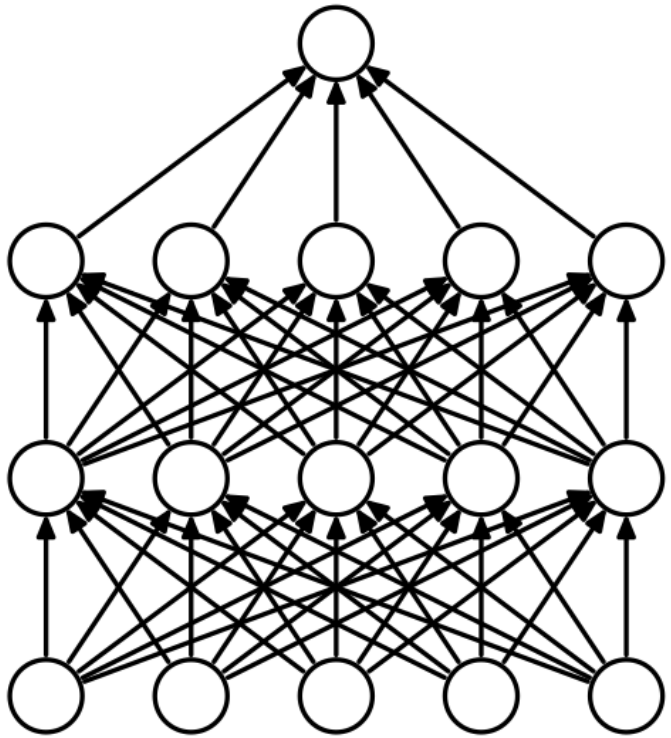
CNN Optimization

- CNN Optimization used to reduce the overfitting problem in CNN by:

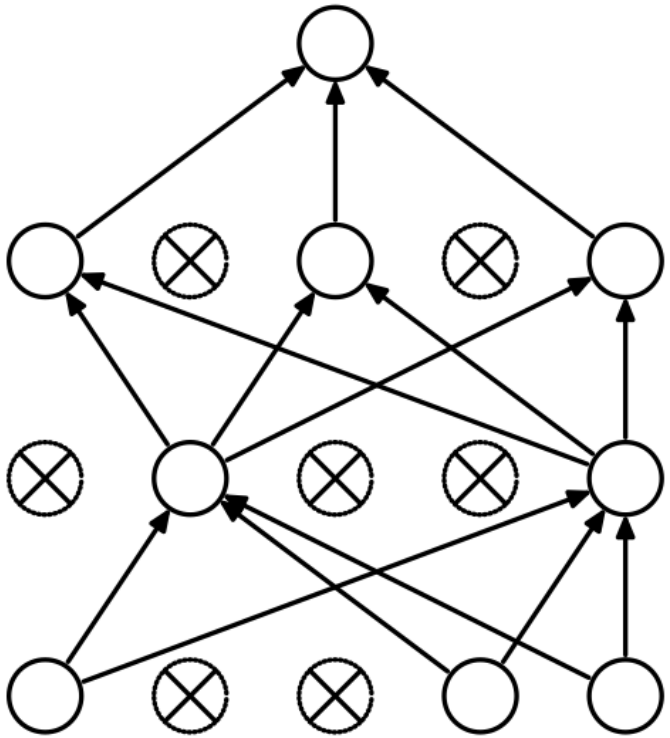
1) Dropout
2) L2 Regularization
3) Mini-batch
4) Gradient descent algorithm
5) Early stopping
6) Data augmentation

CNN Optimization

- **Dropout** is a technique of reducing overfitting in CNN.



(a) Standard Neural Net



(b) After applying dropout.

CNN Optimization

- **L2 Regularization**: Adding a regularization term for the weights to the loss function is a way to reduce overfitting.

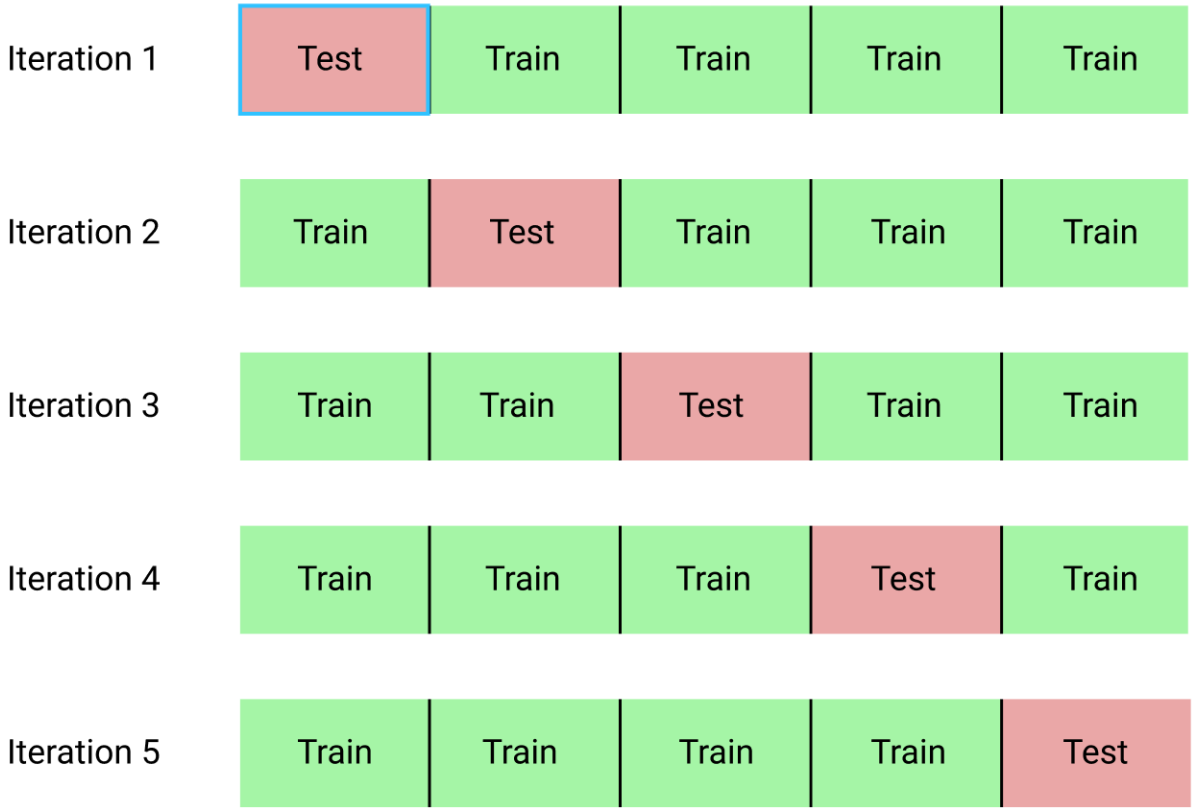
$$E_R(P_i) = E(P_i) + \lambda \Omega(w)$$

- where w is the weight vector, λ is the regularization factor (coefficient), and the regularization function, $\Omega(w)$ is:

$$\Omega(w) = \frac{1}{2} w^t w$$

CNN Optimization

- **Mini-batch** is to divide the dataset into small batches of examples, compute the gradient using a single batch, make an update, then move to the next batch.



CNN Optimization

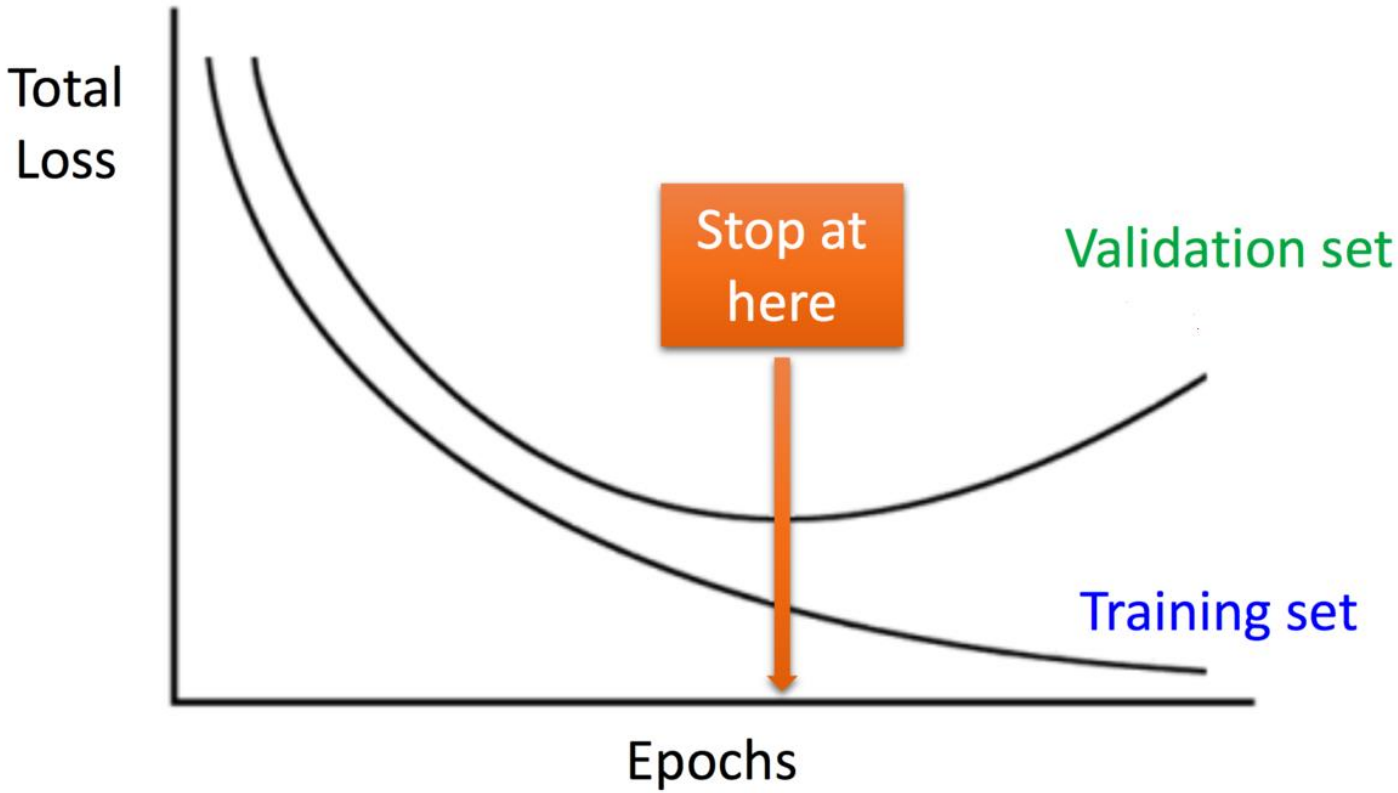
- **The gradient descent algorithm** updates the coefficients (weights and biases) so as to minimize the error function by taking small steps in the direction of the negative gradient of the loss function

$$P_{i+1} = P_i - \alpha \nabla E(P_i)$$

- where i stands for the iteration number, $\alpha > 0$ is the learning rate, P is the parameter vector, and E(Pi) is the loss function.

CNN Optimization

- **Early stopping** monitoring the deep learning process of the network from overfitting.
- If there is no more improvement, or worse, the performance on the test set degrades, then the learning process is aborted



CNN Optimization

- Data augmentation means increasing the number of dataset.

NVIDIA TITAN X

Today NVidia Support my work with

NVIDIA TITAN X

THE MOST ADVANCED GPU EVER BUILT

NVIDIA®

NVIDIA TITAN X

TITAN X Specifications	
GPU Architecture	Pascal
Standard Memory Config	12 GB GDDR5X
Memory Speed	10 Gbps
Boost Clock	1531 MHz
NVIDIA CUDA® Cores	3584
Transistors	12,000 million

NVIDIA TITAN X

TITAN X In Research	
Deep Learning	Augmented Reality
Machine Learning	Image Recognition
Computer Vision	Data Science

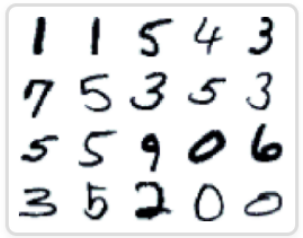
Summary

- **Deep learning** is a class of machine learning algorithms.

- Harder problems such as video understanding, image understanding , natural language processing and Big data will be successfully tackled by **deep learning algorithms.**

Task3

- Each student select one of Benchmark Dataset from https://goo.gl/DNQmtj



MNIST 50 results collected

Units: error %

> Classify handwriten digits. Some additional results are available on the original dataset page.

Result	Method	Venue	Details
0.21%	Regularization of Neural Networks using DropConnect	ICML 2013	

Task3

CIFAR-10

who is the best in CIFAR-10 ?

CIFAR-10 49 results collected

Units: accuracy %

> Classify 32x32 colour images.

89.14%	Deep Convolutional Neural Networks as Generic Feature Extractors	IJCNN 2015	Details
89%	ImageNet Classification with Deep Convolutional Neural Networks	NIPS 2012	Details

- Download the paper that describe the method used by the authors.
- Make one page using word to summarize your selected paper.

How To Send Tasks

- Use MS Word
- Send me e-mail to **mloey@live.com** with email subject “ **Advanced Topics in CS2 – Task3** “
- Put your **Arabic name** on word and email body
- Finally, press **Send**
- **Deadline Next Lecture**

Contact Me

facebook.com/mloey

mohamedloey@gmail.com

twitter.com/mloey

linkedin.com/in/mloey

mloey@fci.bu.edu.eg

mloey.github.io

THANKS FOR YOUR TIME