Deep Learning- Convolutional Neural Network(CNN)

by Natnael Tilahun

February 26, 2020

Contents

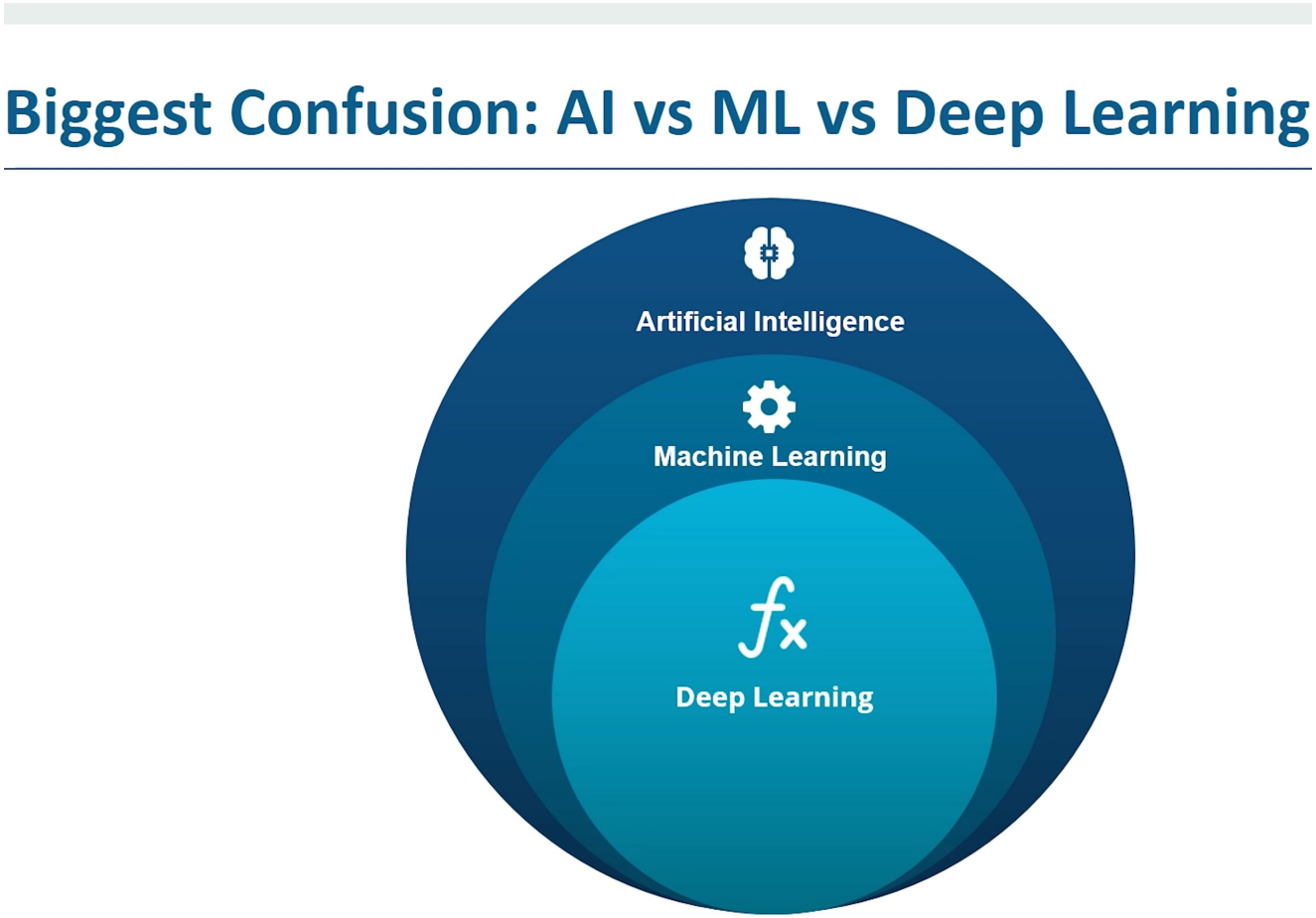
Artificial Intelligence, Machine Learning, and Deep Learning

What exactly is the difference between them?

- As the name suggests, artificial intelligence can be loosely interpreted to mean incorporating human intelligence to machines.
- Whenever a machine completes tasks based on a set of stipulated rules that solve problems (algorithms), such an "intelligent" behavior is what is called artificial intelligence.
- Machine Learning is to enable machines to learn by themselves without explicitly programmed, using the provided data and make accurate predictions.
- Deep Learning is machine learning sub-field of learning representations of data. And inspired by information processing of human brain.
- Deep learning is considered as the next evolution of machine learning.

AI,ML,and DL

What is the relationship between?

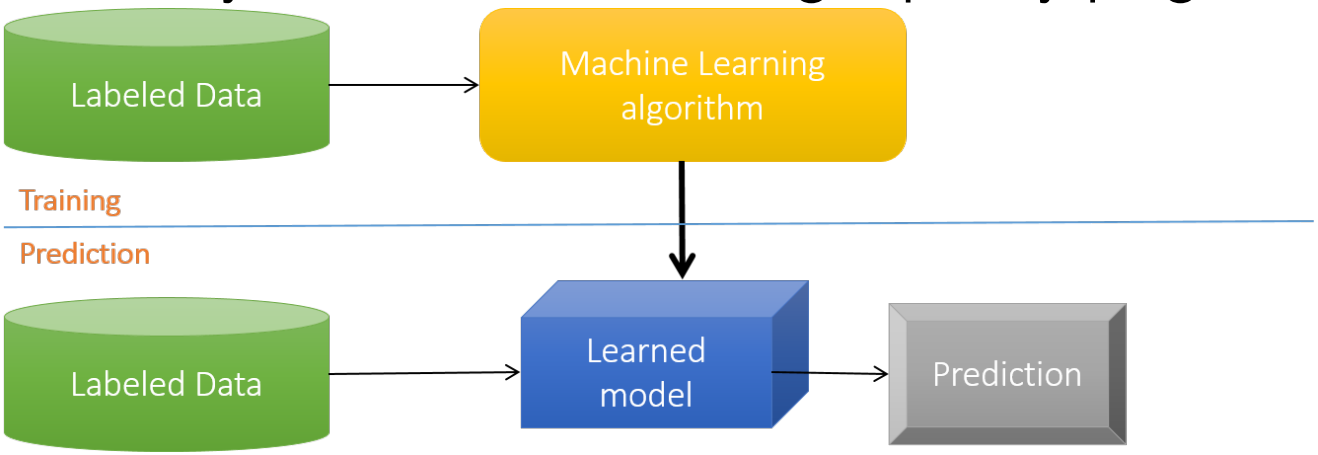


The Relationship Between AI,ML, and DI

Machine Learning Basics

What exactly is ML?

- Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed.



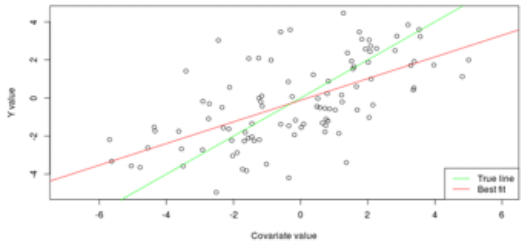
Machine Learning Example

- Methods that can learn from and make predictions on data.

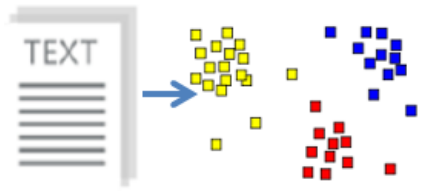
Types of Learning

- **Supervised Learning**: with a labeled training set Example: email classification with already labeled emails
- **Unsupervised Learning** : Discover patterns in unlabeled data Example: cluster similar documents based on text
- **Reinforcement learning**: learn to act based on feedback/reward Example: learn to play Go, reward: win or lose
- Methods that can learn from and make predictions on data.

Classification
Anomaly Detection
Sequence labeling



Regression

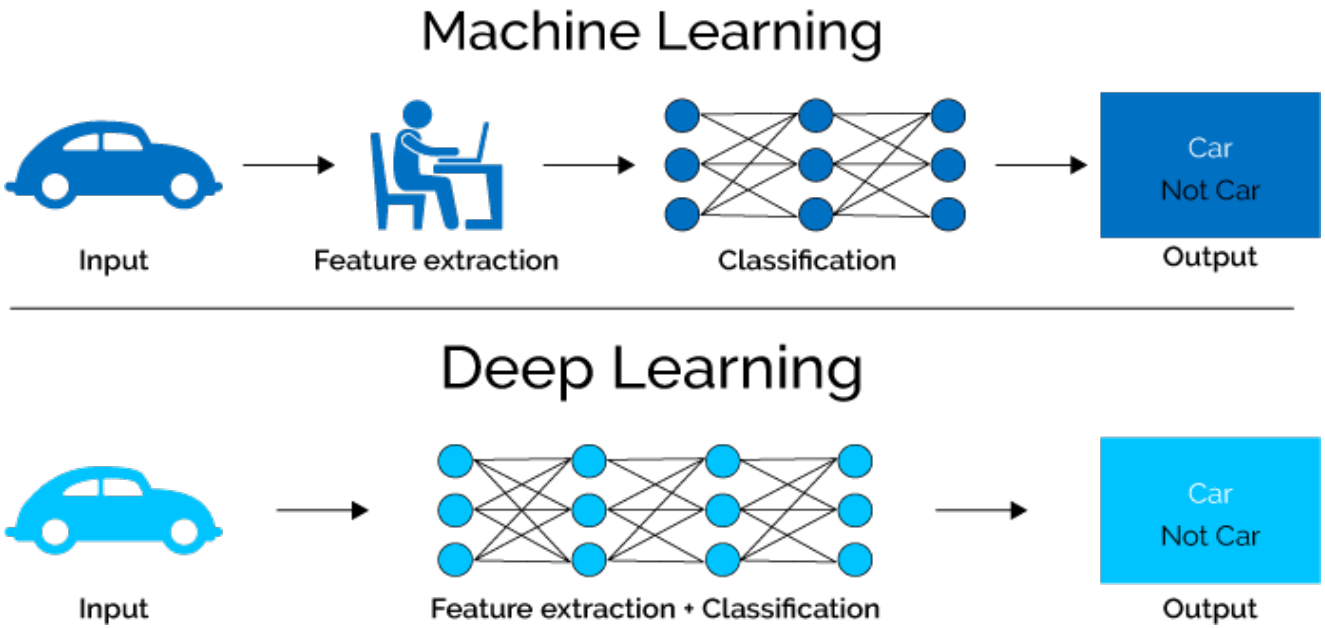


Clustering

Examples of Types of Learning

What is Deep Learning ?

- A **machine learning** sub-field of learning representations of data. Exceptional effective at learning patterns.
- **Deep learning algorithms** attempt to learn (multiple levels of) representation by using a hierarchy of multiple layers.
- If you provide the system tons of information, it begins to understand it and respond in useful ways.



Classification ML Vs DL

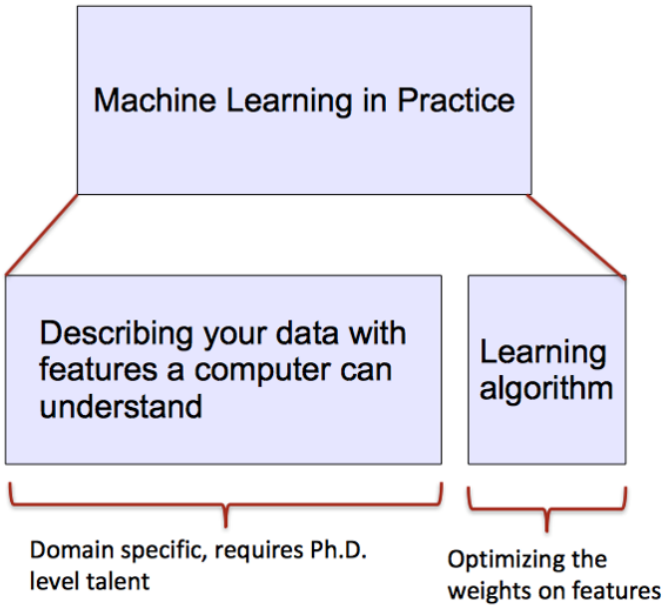
What is Deep Learning ? (Cont.)

The "Deep" in the Deep Learning

- The **"deep"** in deep learning isn't a reference to any kind ofdeeper understanding achieved by the approach.
- Rather, it stands for this idea of **successive layers of representations**
- Other appropriate names for the field could have been **layered representations learning** and **hierarchical representations learning**

Machine Learning Vs Deep Learning

- Most machine learning methods work well because of **human-designed representations and input features**
- ML becomes just **optimizing weights** to best make a final prediction

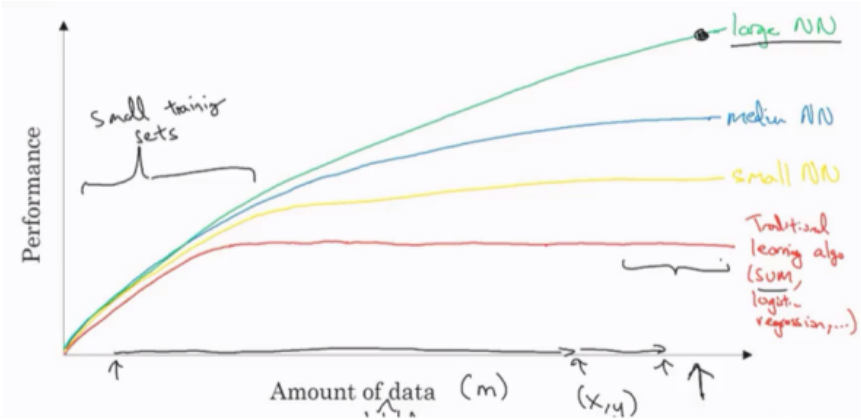


ML in Practice

Feature	NER
Current Word	✓
Previous Word	✓
Next Word	✓
Current Word Character n-gram	all
Current POS Tag	✓
Surrounding POS Tag Sequence	✓
Current Word Shape	✓
Surrounding Word Shape Sequence	✓
Presence of Word in Left Window	size 4
Presence of Word in Right Window	size 4

ML in Practice

- **Data**: Hopefully we have a lot of data because the world is using the computer a little bit more, Mobiles, IOT (Internet of things).
- **Computation**: GPUs, Powerful CPUs, Distributed computing.
- **Algorithm**: Creative algorithms has appeared that changed the way NN works. For example using RELU function is so much better than using SIGMOID function in training a NN because it helps with the vanishing gradient problem.



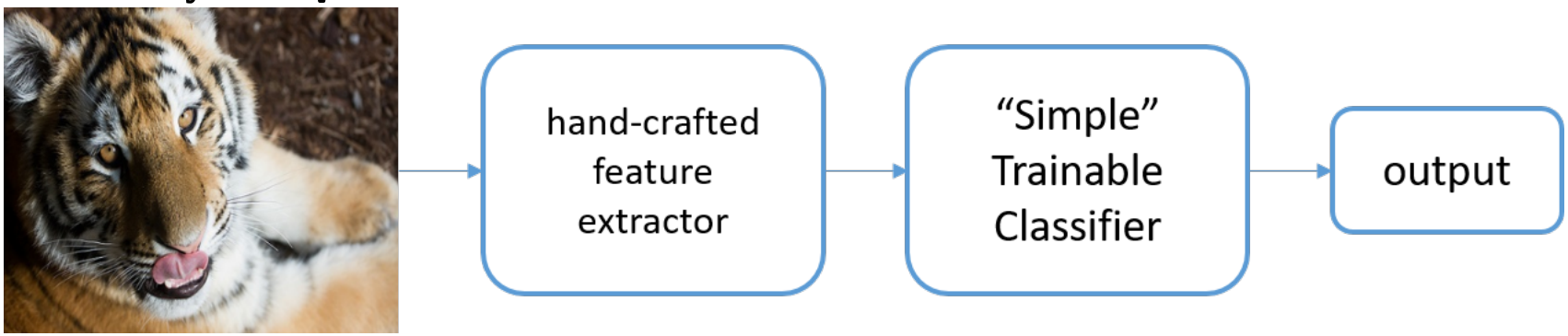
DL Vs Traditional Algorithms

From the above graph we can interpret the following:

- For **small data** NN can perform as Linear regression or SVM (Support vector machine)
- For **big data** a small NN is better that SVM
- For **big data** a **big NN** is better that a medium NN is better that small NN

Computer Vision and Deep Learning(Cont.)

- Traditional pattern recognition models use hand-crafted features and relatively **simple trainable classifier**



Traditional Pattern Recognition

This approach has the following limitations:

- It is **very tedious and costly** to develop hand-crafted features
- The hand-crafted features are usually **highly dependents on one application**, and **cannot** be transferred easily to other applications

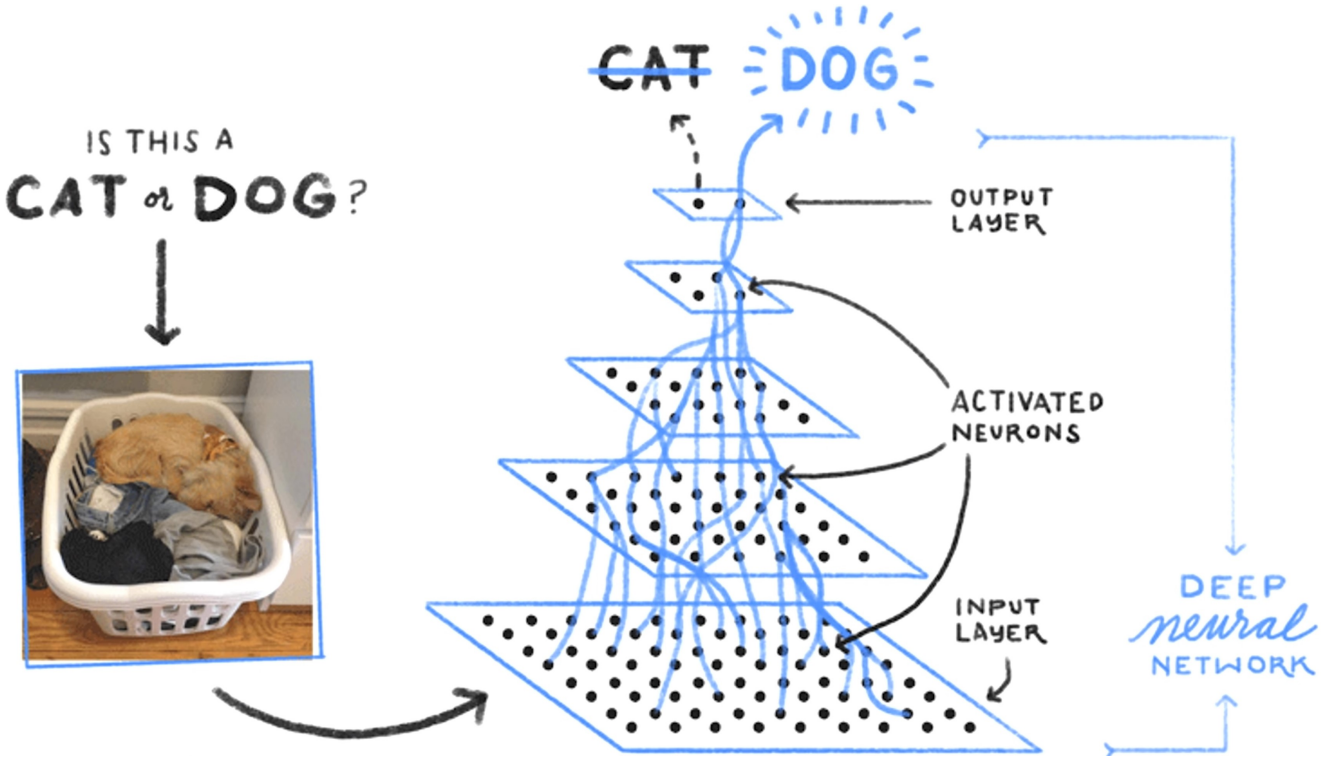
Why is DL useful?

What makes DL a big deal?

- Manually designed features are often **over-specified, incomplete and take a long time to design and validate**
- Learned **Features are easy to adapt**, **fast to learn**
- Deep learning provides a **very flexible**, (almost?) universal, learnable framework for representing world, visual and linguistic information
- Can learn both **unsupervised and supervised**
- Effective end-to-end joint system learning
- Utilize **large amounts of training data**

CNN Classification

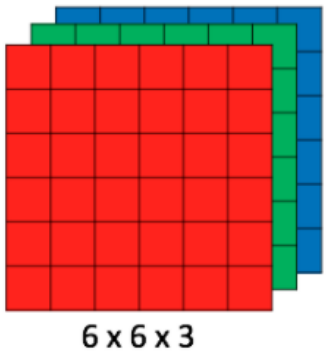
Classifying a given image as a Cat or Dog



CNN Classification

Deep CNN Classification

- CNN image classifications takes an input image, process it and **classify it under certain categories** (Eg., Dog, Cat).
- Computers sees an input **image as array of pixels** and it depends on the **image resolution**

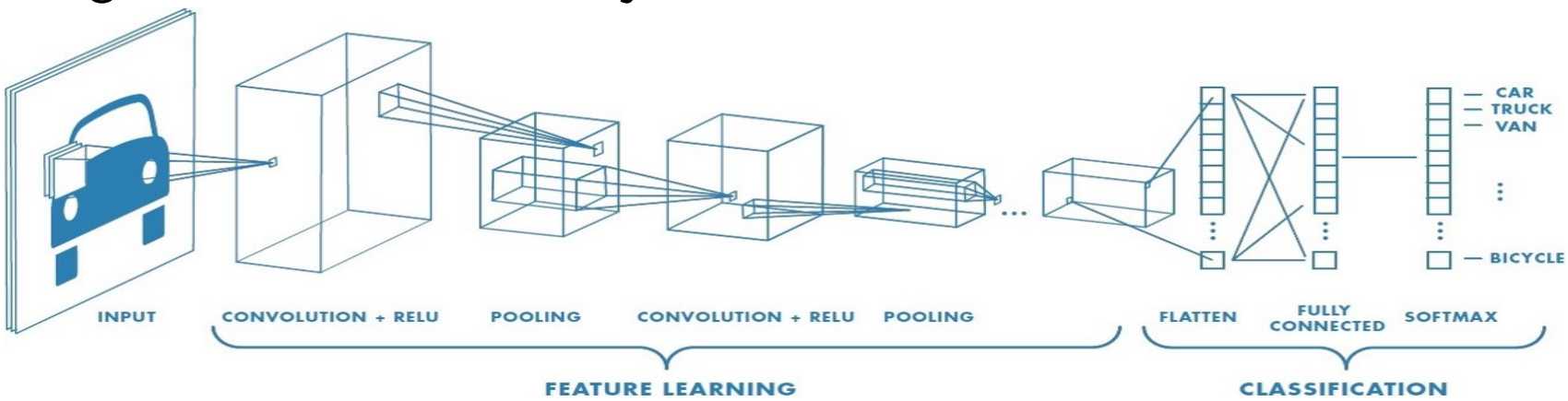


RGB Image

- Based on the image resolution, it will see **h x w x d**(h = Height, w = Width, d = Dimension).
- Eg., An image of 6 x 6 x 3 array of matrix of RGB (3 refers to RGB values) and an image of 4 x 4 x 1 array of matrix of grayscale image.

Deep CNN Classification(Cont.)

- Deep learning CNN models to train and test, each input image will **pass it through many phases or layers**
- Those are: **convolution layers with filters (Kernals), Pooling, fully connected layers (FC) and apply Softmax function**
- Softmax function used to classify an object with probabilistic values between 0 and 1
- The below figure is a complete flow of CNN to process an input image and classifies the objects based on values.



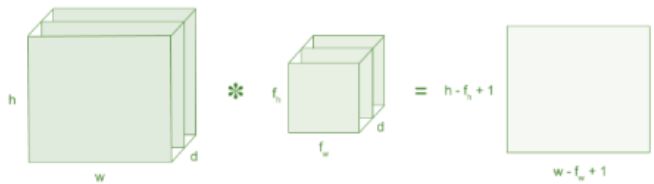
Deep CNN Classification

Deep CNN Layers

1. Convolutional Layer

- Is the first layer to extract features from an input image
- Convolution preserves the relationship between pixels by learning image features using small squares of input data
- It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel

- An image matrix (volume) of dimension **(h x w x d)**
- A filter **(f_h x f_w x d)**
- Outputs a volume dimension **(h - f_h + 1) x (w - f_w + 1) x 1**



Convolution Operation

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

5 x 5 – Image Matrix

*

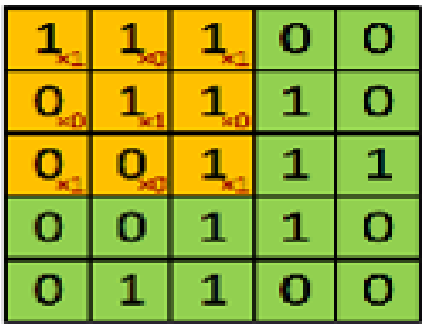
1	0	1
0	1	0
1	0	1

3 x 3 – Filter Matrix

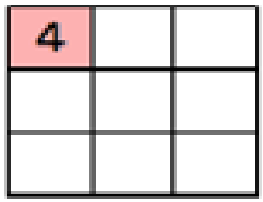
Convolution Operation

Deep CNN Layers(Cont.)

- Then the convolution of 5 x 5 image **matrix multiplies** with 3 x 3 filter matrix which is called "Feature Map" as output shown
- Convolution of an image with different filters can perform operations such as **edge detection, blur and sharpen by applying filters**
- The below example shows various convolution image after applying different types of filters (Kernels)



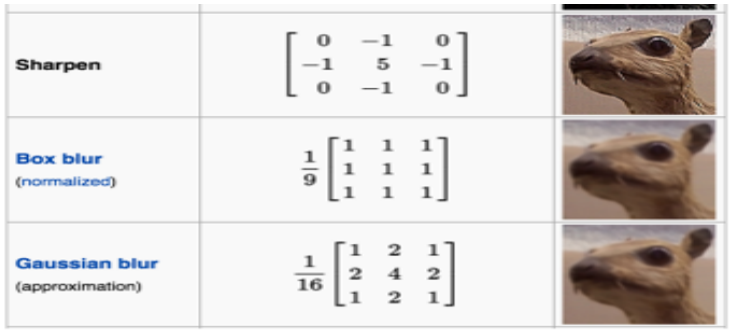
Image



Convolved Feature

An output Matrix

- Types of filters: Example



Stride and Padding

Stride

- Stride is the **number of pixels shifts over the input matrix**
- When the stride is 1 then we move the filters to 1 pixel at a time
- When the stride is 2 then we move the filters to 2 pixels at a time and so on

Padding

- Sometimes **filter does not fit perfectly fit the input image**
- We have two options:
  1. Pad the picture with zeros **(zero-padding)** so that it fits
  2. Drop the part of the image where the filter did not fit. This is called **valid padding** which keeps only valid part of the image

0	0	0	0	0	0	0	0
0	3	1	1	2	8	4	0
0	1	0	7	3	2	6	0
0	2	3	5	1	1	3	0
0	1	4	1	2	6	5	0
0	3	2	1	3	7	2	0
0	9	2	6	2	5	1	0
0	0	0	0	0	0	0	0

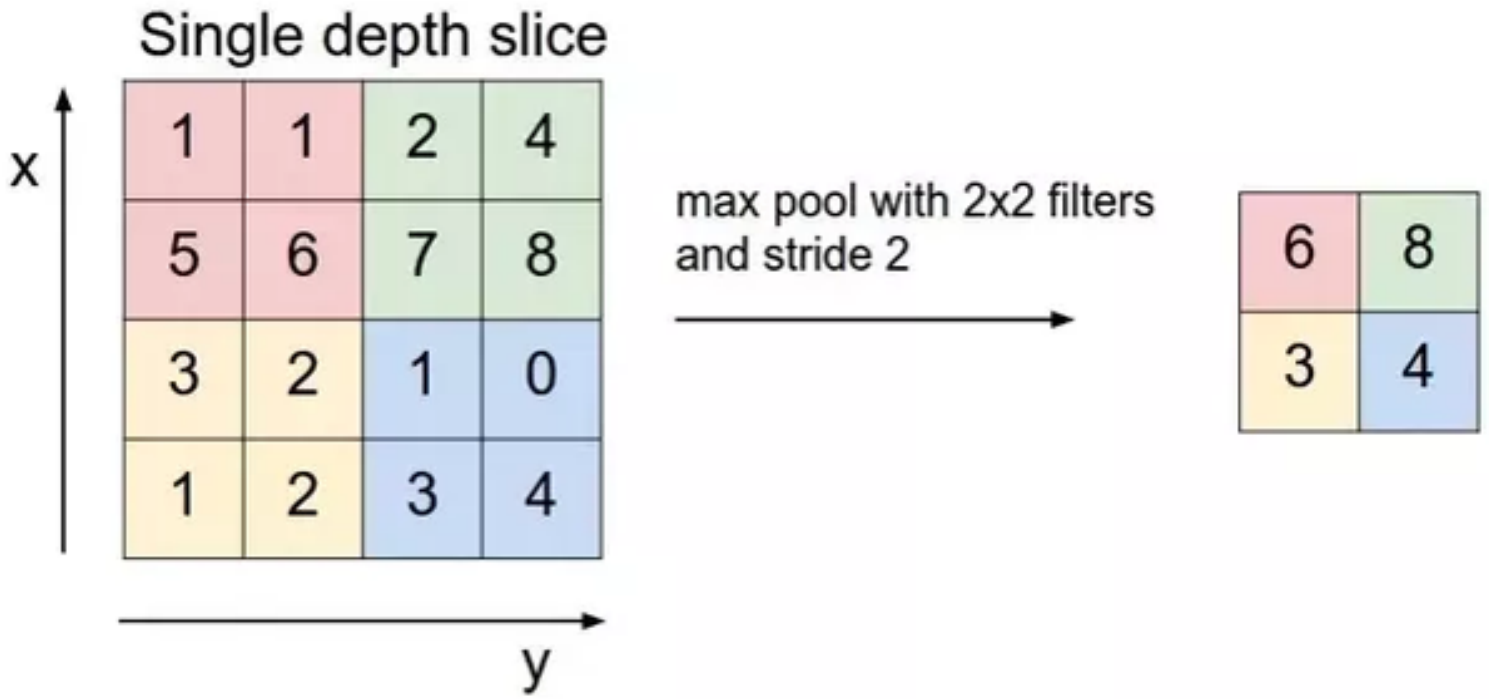
Zero Padding

Pooling Layer

- Pooling layers section would **reduce the number of parameters** when the images are too large
- Spatial pooling also called sub-sampling or down-sampling which reduces the dimensionality of each map but retains important information.
- There are different types of Spatial Pooling
  1. **Max Pooling**
  2. **Average Pooling**
  3. **Sum Pooling**
- **Max pooling** takes the largest element from the rectified feature map
- Taking the largest element could also take the average pooling
- Sum of all elements in the feature map call as **sum pooling**

Pooling Layer(Cont.)

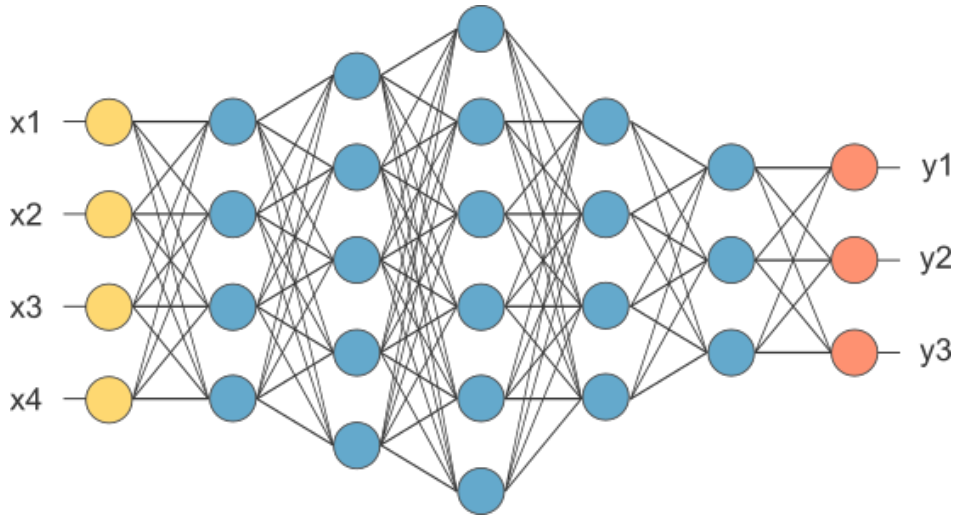
- Example: Max Pooling



Max Pooling

Fully Connected Layer(FC)

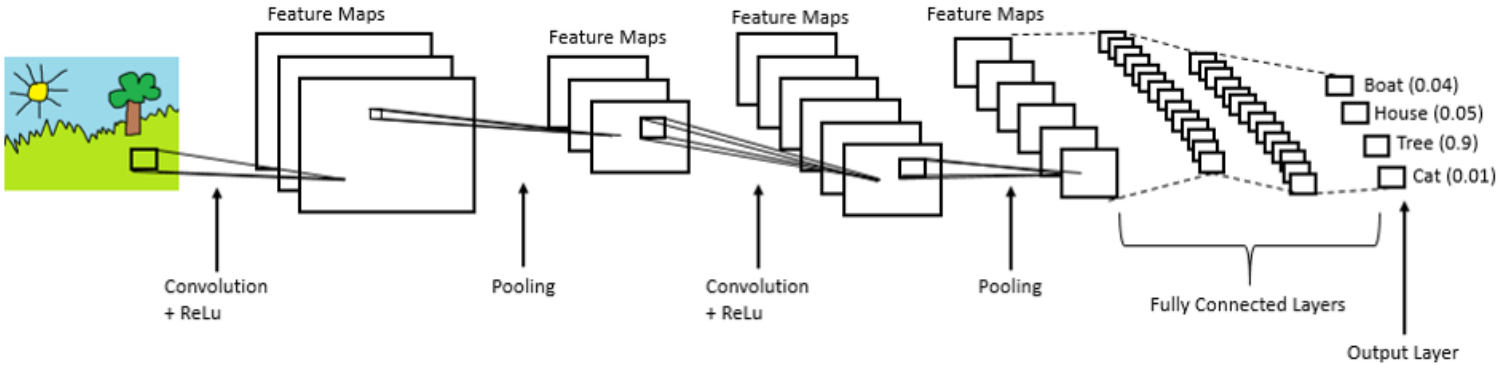
- The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like a neural network



Fully Connected Layer

- In the above diagram, the feature map matrix will be converted as vector (x1, x2, x3, . . .)
- With the fully connected layers, we combined these features together to create a model
- Finally, we have an activation function such as softmax or sigmoid to classify the outputs as cat, dog, car, truck etc.,

The Complete CNN Architecture

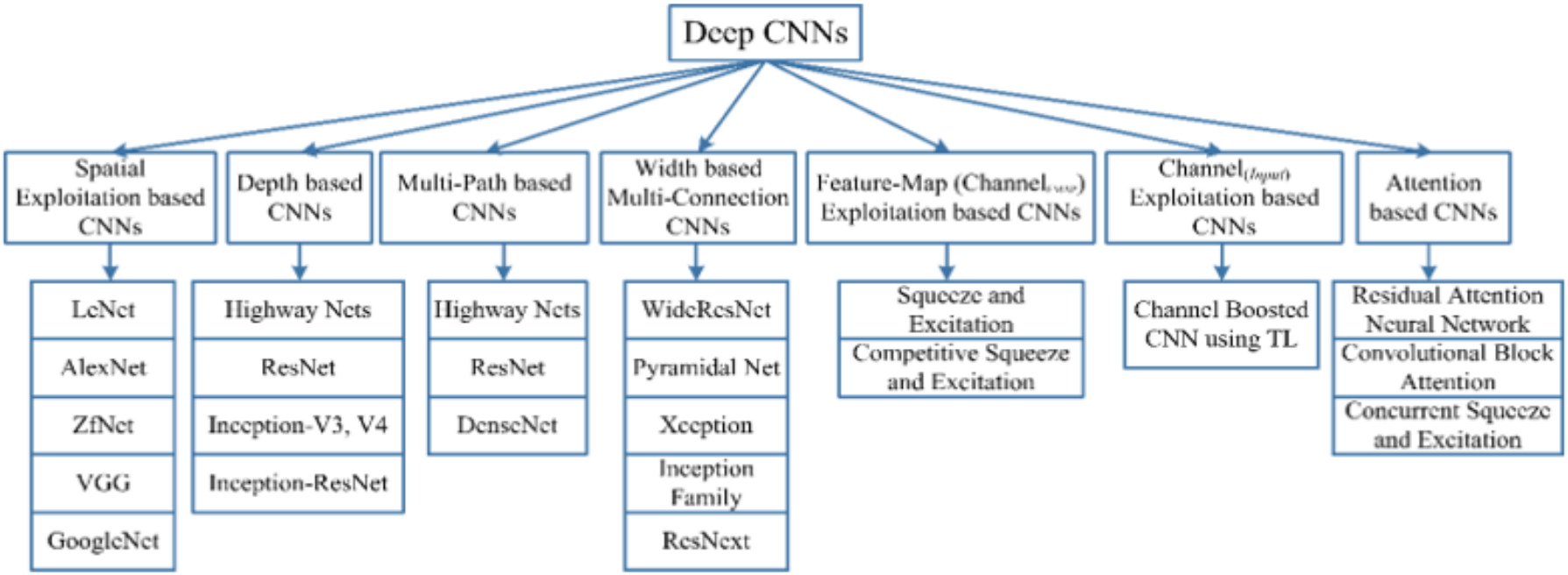


The Final Architecture

Summary

- Provide input image into **convolution layer**
- Choose **parameters, apply filters with strides, padding if requires**. Perform convolution on the image and apply ReLU activation to the matrix
- Add as many **convolutional layers** until satisfied
- Flatten the output and feed into a fully connected layer (FC Layer)
- Output the class using an **ctivation functiona** (Logistic Regression with cost functions) and classifies images

Deep Learning CNN Taxonomy



Deep CNN Taxonomy

Some Deep Learning Frameworks

- A deep learning framework is an interface, library or a tool which allows us to build deep learning models
- It has to be easy and quick, without getting into the details of underlying algorithms
- DL framework provide a clear and concise way for defining models using a collection of pre-built and optimized components

25

TensorFlow

- TensorFlow is an end-to-end open source platform for machine learning
- It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML
- Developers easily build and deploy ML powered applications

Why TenserFlow?

1. Easy model building
2. Robust ML production anywhere
3. Powerful experimentation for research
4. Runs seamlessly on CPU and GPU

Keras

- Keras is a high-level neural networks API, written in Python and capable of running on top of TenserFlow, CNTK, or Theano
- It was developed with a focus on enabling fast experimentation
- Being able to go from idea to result with the least possible delay is key to doing good research

Why Keras?

1. Allows for easy and fast prototyping (through user friendliness, modularity, and extensibility)
2. Supports both convolutional networks and recurrent networks, as well as combinations of the two
3. Runs seamlessly on CPU and GPU

PyTorch

- PyTorch is an open source machine learning library based on the Torch library, used for applications such as computer vison and natural language processing
- It is primarily developed by Facebook's AI Research lab (FAIR)
- PyTorch is an optimized tensor library for deep learning using GPUs and CPUs

Some Advanatge of PyTorch

1. The modeling process is simple and transparent thanks to the framework's architectural style
2. It supports distributed training

Applications of Deep Learning

Automated Driving

Automotive researchers are using deep learning to automatically detect objects such as stop signs and traffic lights.

Aerospace and Defense

Deep learning is used to identify objects from satellites that locate areas of interest, and identify safe or unsafe zones for troops.

Medical Research

Cancer researchers are using deep learning to automatically detect cancer cells.

Applications of Deep Learning(Cont.)

Industrial Automation

Deep learning is helping to improve worker safety around heavy machinery by automatically detecting when people or objects are within an unsafe distance of machines.

Electronics

Deep learning is being used in automated hearing and speech translation.

The End!!!
Thank You!
Question?