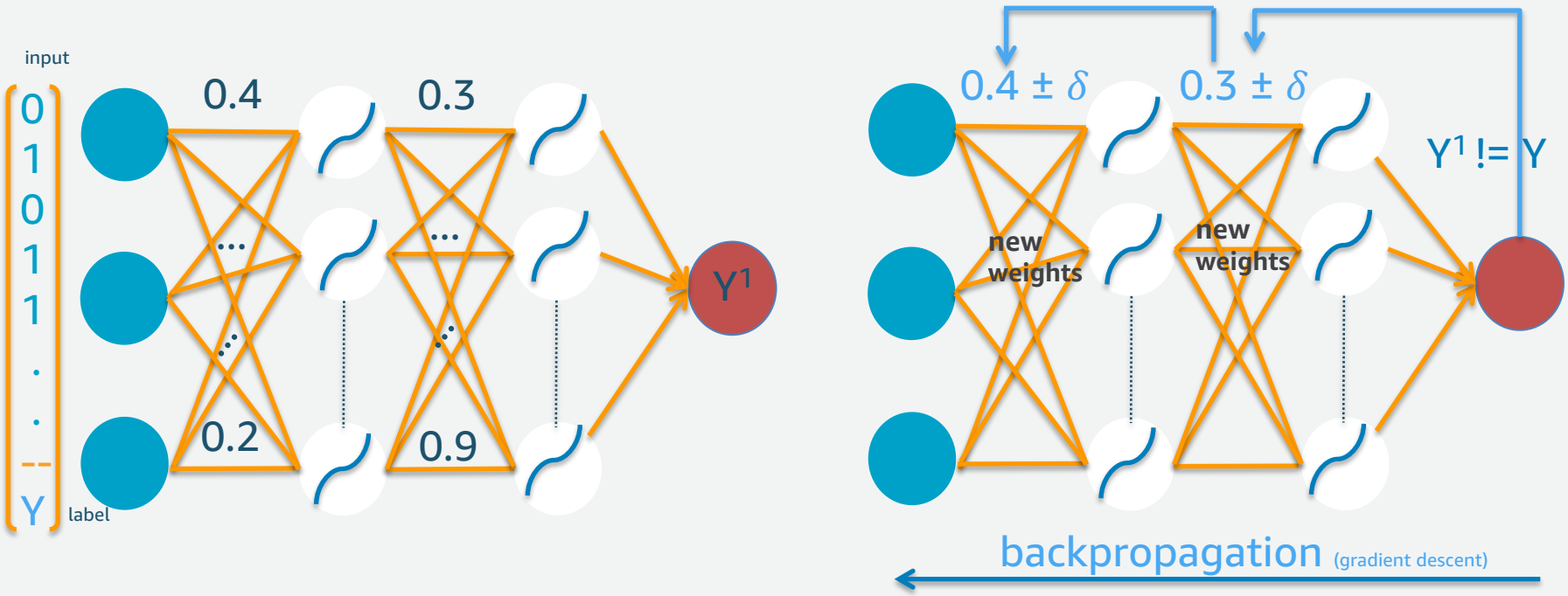
Recap: Neural Network and CNN

Muhyun Kim, AI Labs

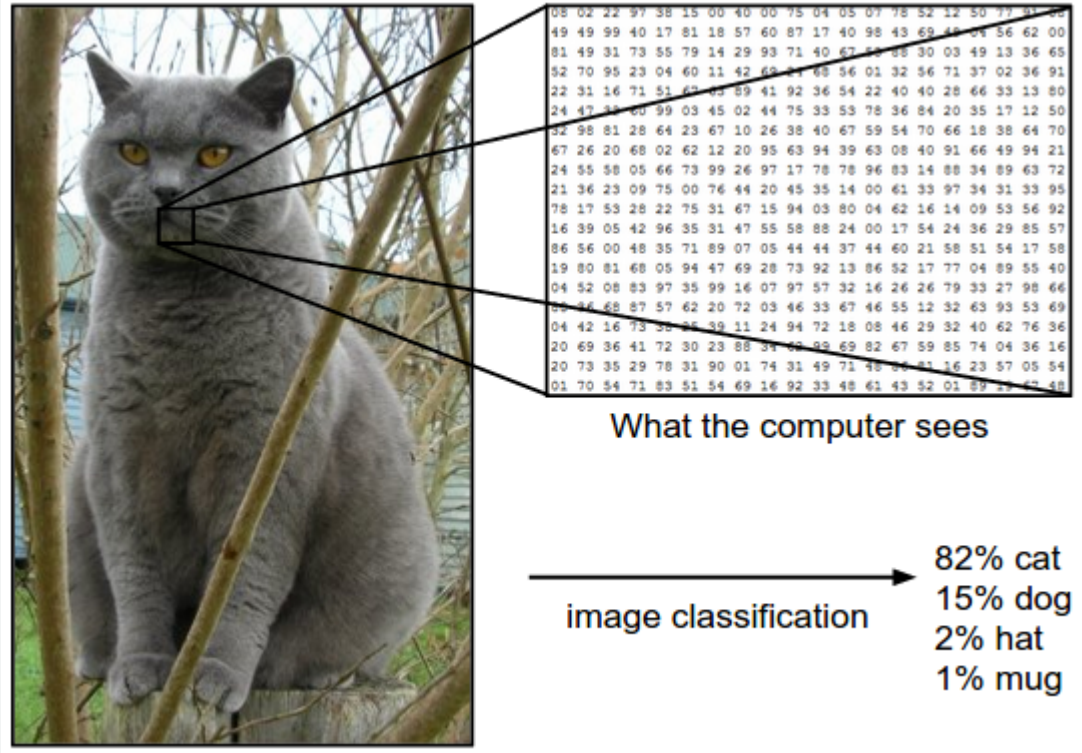
August 13, 2018

Learning from Data



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Images seen by Computer

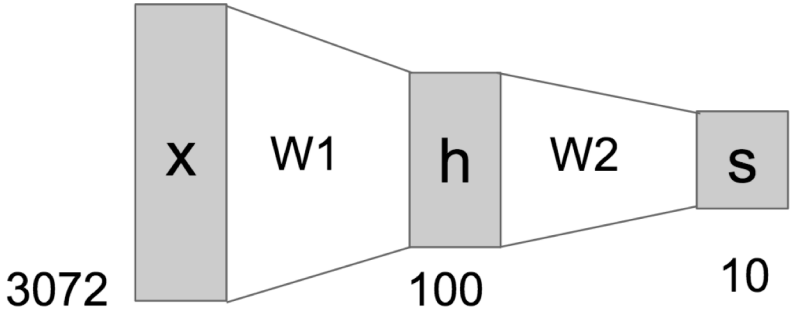


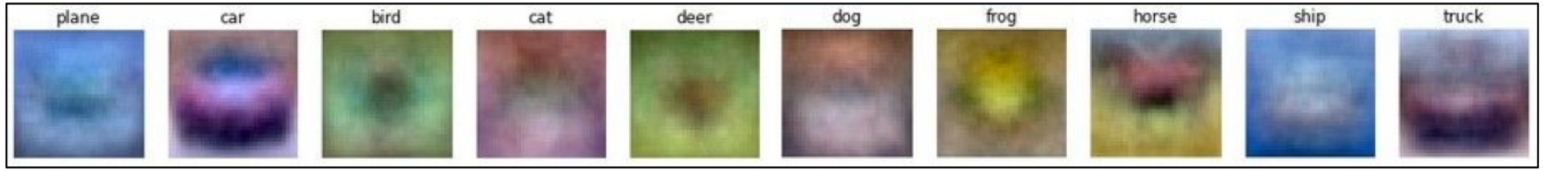
© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Multilayer Neuron Network Approach

Linear score function: $f = Wx$

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

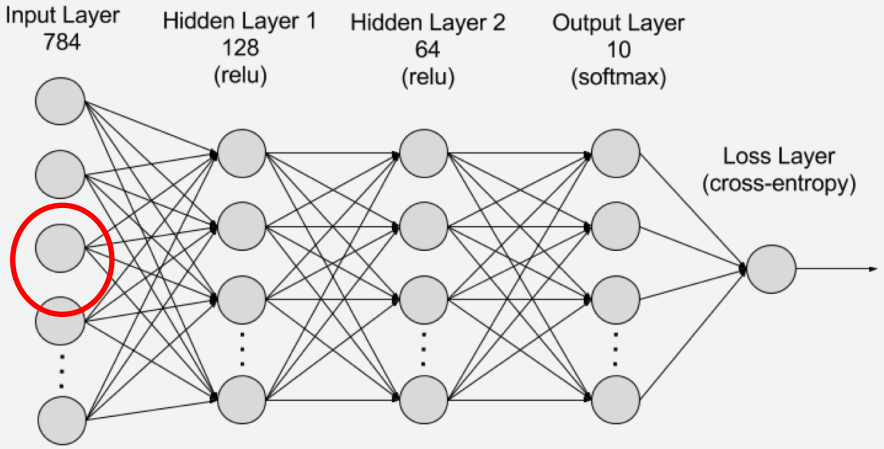




© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Multilayer Perception is NOT enough in CV

28 x 28
Array

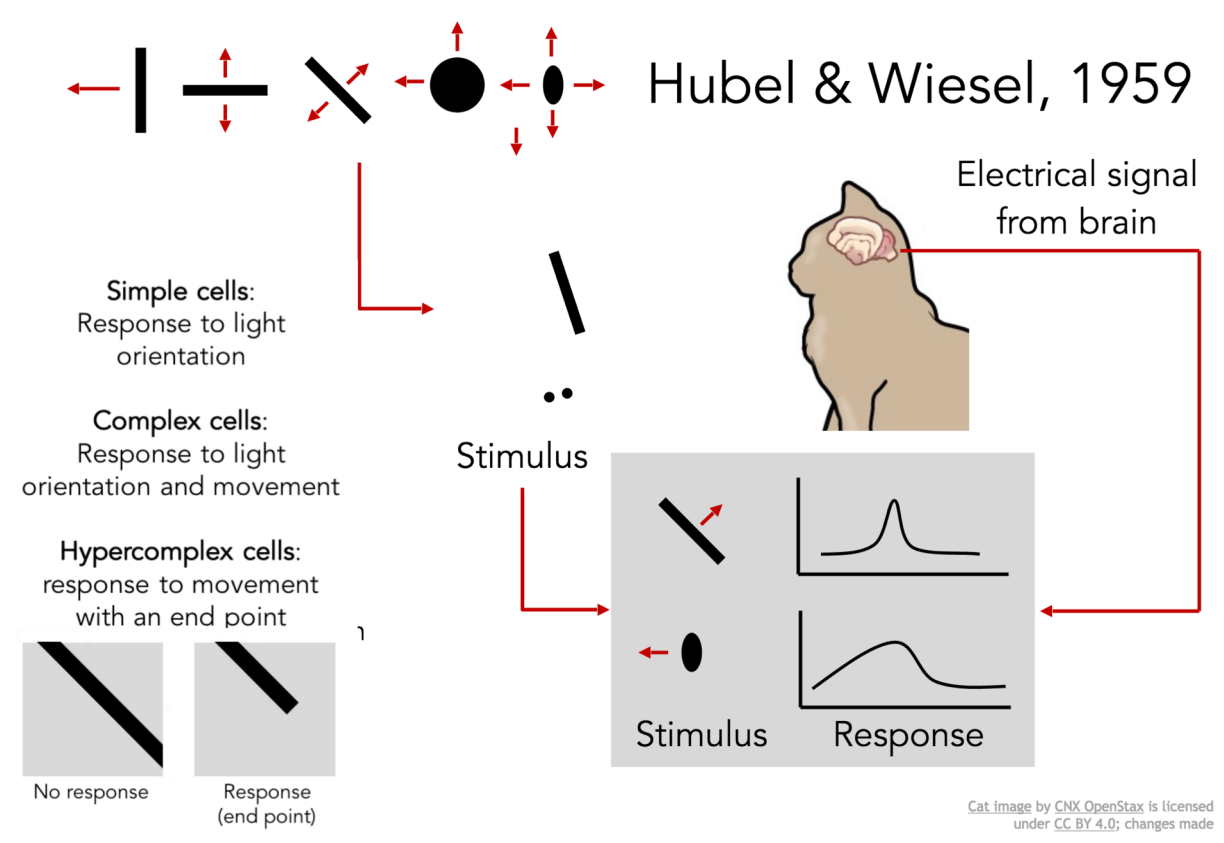


784 1
Dimensional
Array

Context in Image is NOT preserved

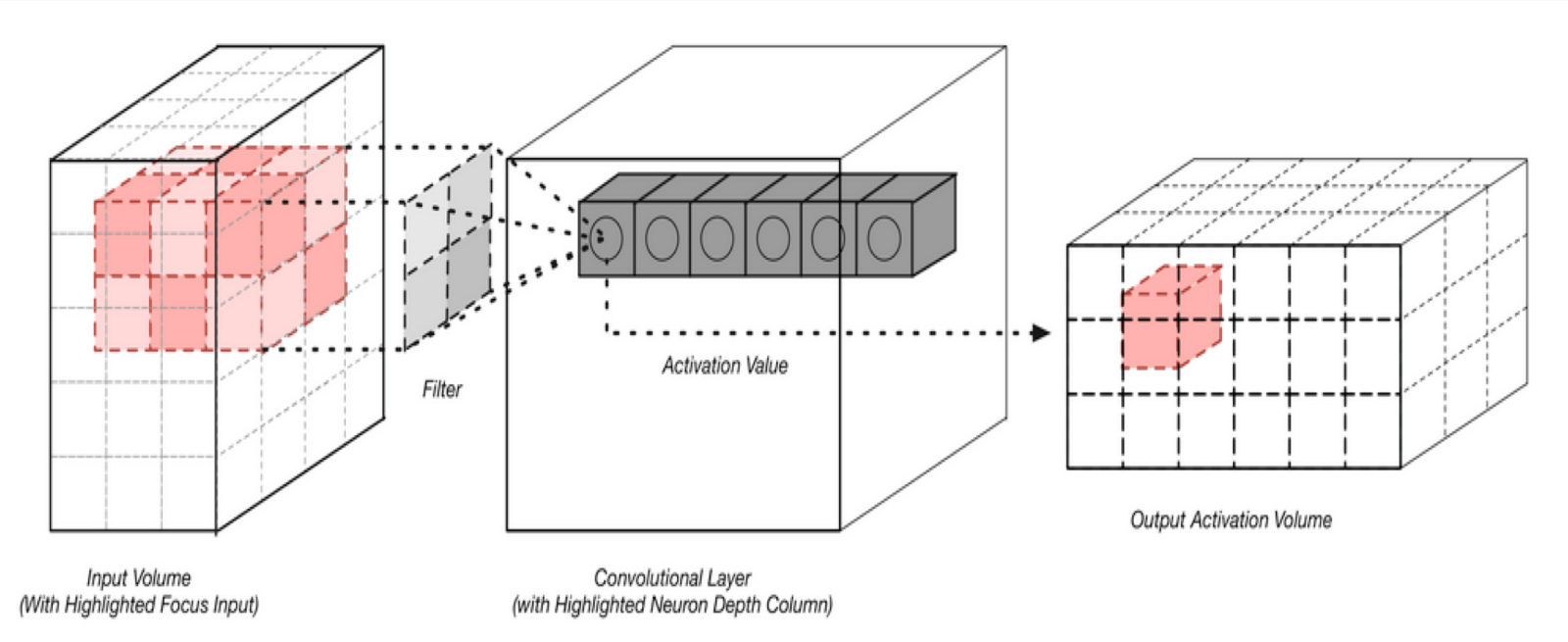
© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Neural Basis of Visual Perception



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Convolutional Neural Network

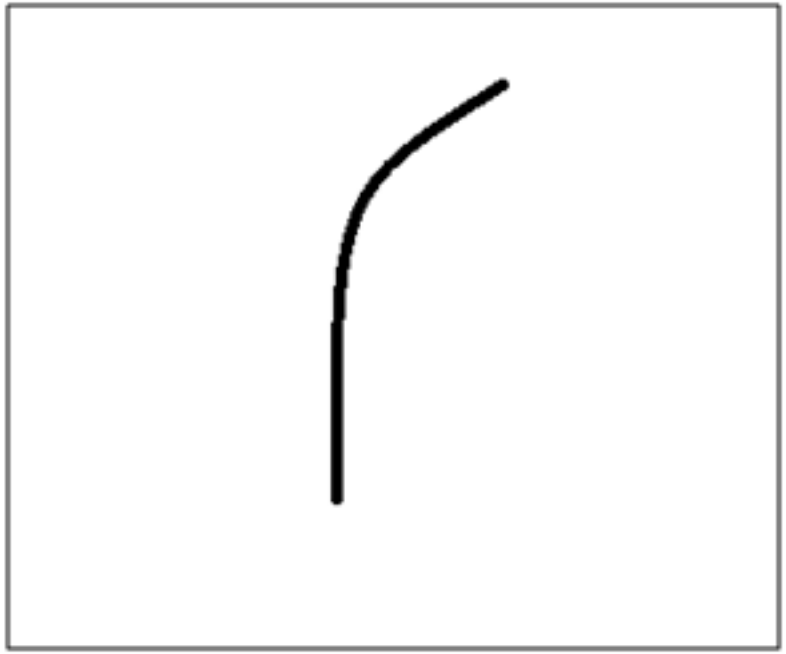


© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

How to find patterns in Image - Filter

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter



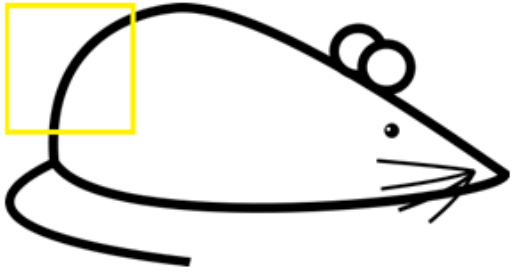
Visualization of a curve detector filter

Source: https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

A Filter to detect the curved line

Original image



Visualization of the filter on the image

Visualization of the receptive field

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Pixel representation of the receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = (50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600 (A large number!)

Source: https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

A Filter to detect the curved line

Visualization of the filter on the image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

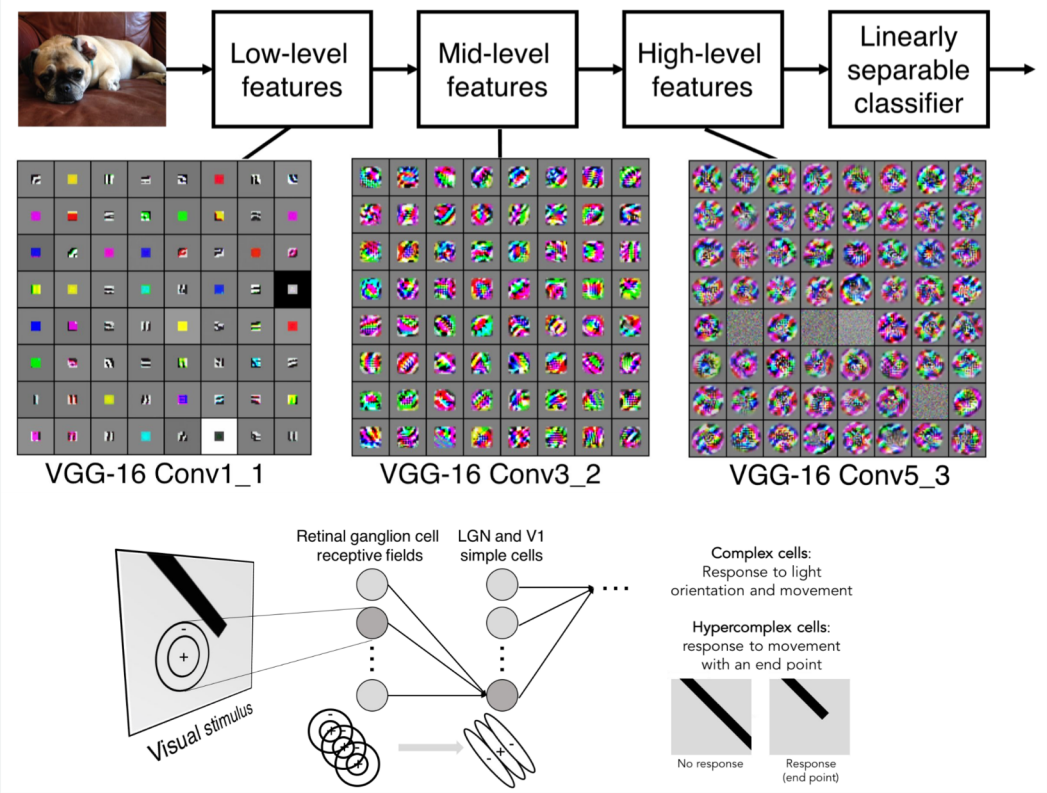
Pixel representation of filter

Multiplication and Summation = 0

Source: https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

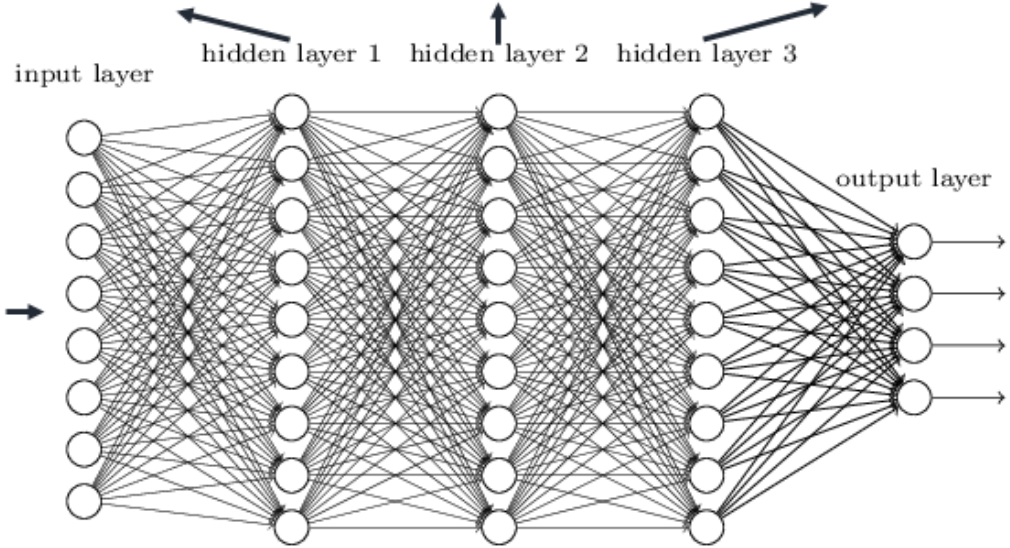
Filters in Layers



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Filters needed in Neural Network

Deep neural networks learn hierarchical feature representations



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

What is the problem to solve with ML?

Define a Neural Network composed of

multiple filters

and

find the filters

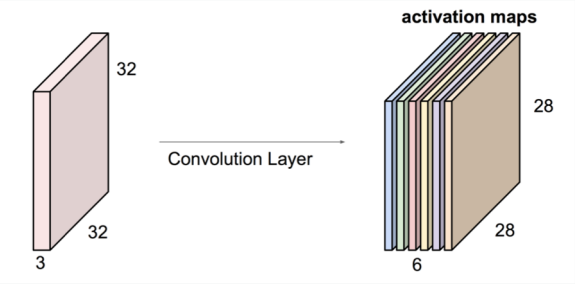
by learning from Dataset

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Types of Layers used in CNN

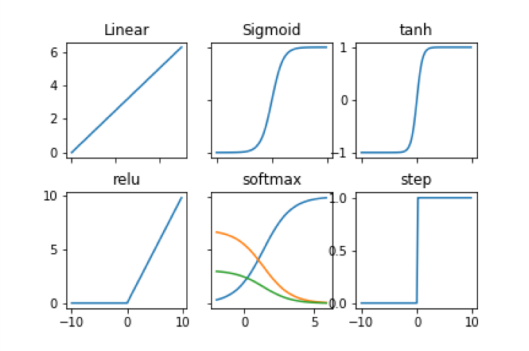
Convolutional Layer

Filter to find patterns



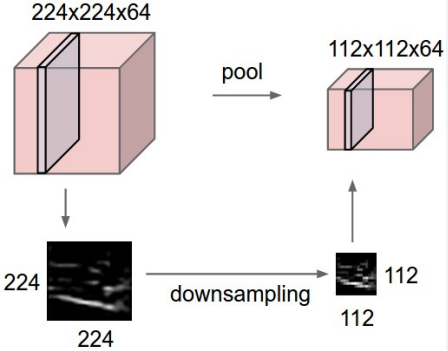
Activation Layer

Non-Linearity



Pooling Layer

Size Reduction



Fully Connected Layer

Convert tensors into vectors

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Convolutions

- **Kernel: 3 Dimension Array to detect a pattern**
- **Receptive Field**: Area to apply the filter
- **Stride**: Steps to move Receptive Field
- **Padding**: Padding around the original tensor

Input Volume (1x5x5 +pad 1)

0	0	0	0	0	0	0
0	12	31	12	167	255	0
0	**45**	**54**	67	254	254	0
0	**0**	**25**	238	254	254	0
0	**10**	**196**	254	254	251	0
0	127	254	254	238	98	0
0	0	0	0	0	0	0

Receptive Field I_{21}

0	45	54
0	0	25
0	10	196

Kernel K (1x3x3)

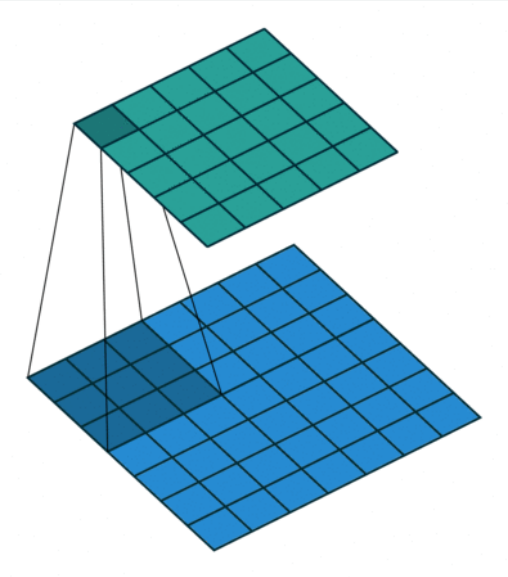
1	0	-1
1	1	0
-1	0	1

Output

0	243	
142		

Inner Product $< I_{21}, K >$

$$= 0 \cdot 1 + 45 \cdot 0 + 54 \cdot -1 + 0 \cdot 1 + 0 \cdot 1 + 25 \cdot 0 + 0 \cdot -1 + 10 \cdot 0 + 196 \cdot 1 = \mathbf{142}$$

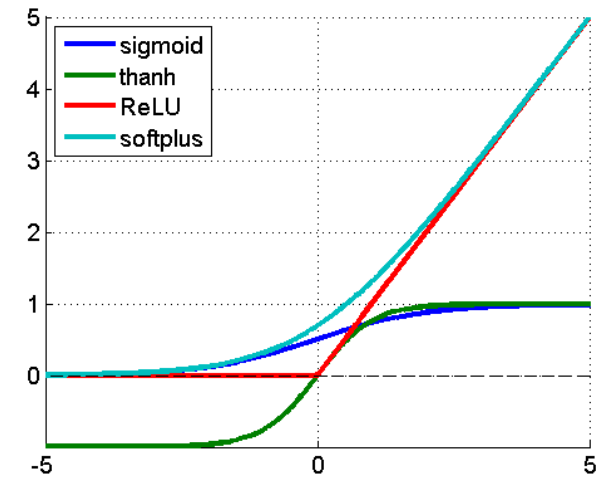


See More Conv.

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.
Image Credit : https://github.com/vdumoulin/conv_arithmetic

Activation Function

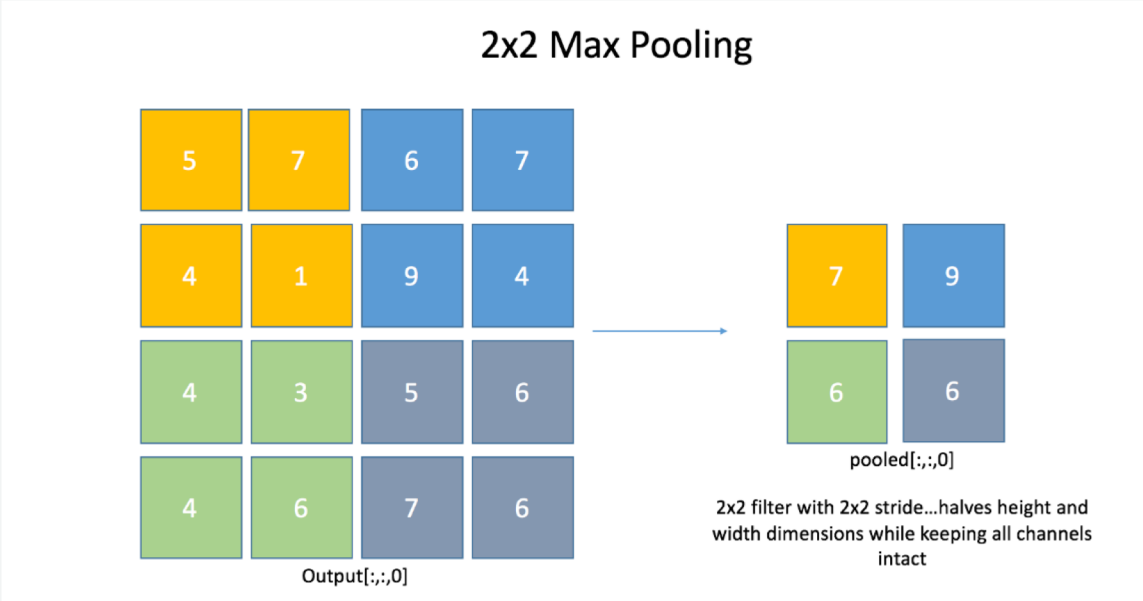
- Adding Non-linearity
- Applied to the output of Convolution Layer
- Activations
  - ReLu (**Re**ctified **L**inear **U**nit)
  - Sigmoid
  - Hiberbolic Tangent (tanh)
  - Softmax
  - Step function



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

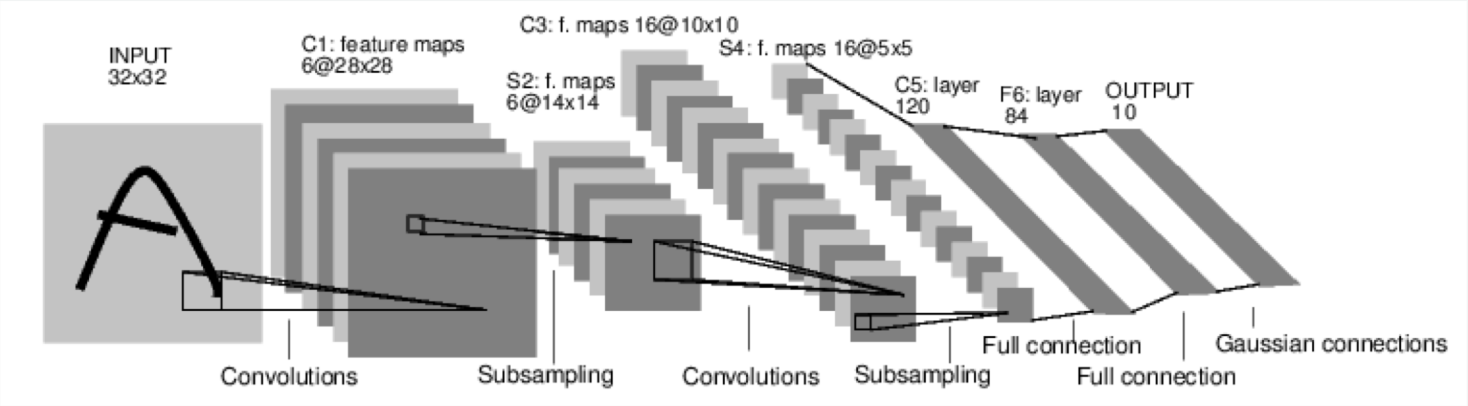
Pooling Layer

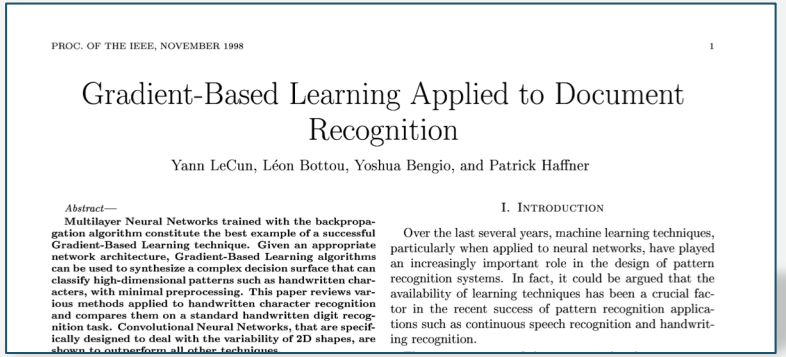
- **Kernel**: Area to extract the value from
- **Stride**: Steps to move Kernel



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

LeNet-5 (1998)





PROC. OF THE IEEE, NOVEMBER 1998 1

Gradient-Based Learning Applied to Document Recognition

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner

Abstract—
Multilayer Neural Networks trained with the backpropagation algorithm constitute the best example of a successful Gradient-Based Learning technique. Given an appropriate network architecture, Gradient-Based Learning algorithms can be used to synthesize a complex decision surface that can classify high-dimensional patterns such as handwritten characters, with minimal preprocessing. This paper reviews various methods applied to handwritten character recognition and compares them on a standard handwritten digit recognition task. Convolutional Neural Networks, that are specifically designed to deal with the variability of 2D shapes, are shown to outperform all other techniques.

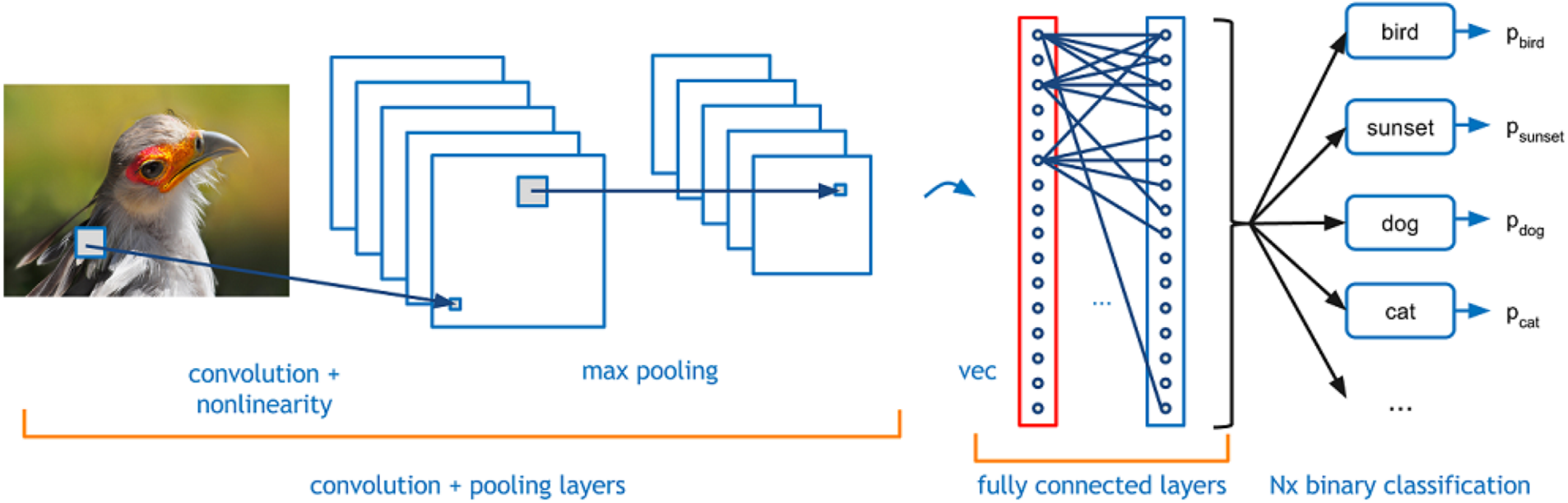
I. Introduction

Over the last several years, machine learning techniques, particularly when applied to neural networks, have played an increasingly important role in the design of pattern recognition systems. In fact, it could be argued that the availability of learning techniques has been a crucial factor in the recent success of pattern recognition applications such as continuous speech recognition and handwriting recognition.

image credit - Le Cun et al, 1998

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Convolutional Neural Network for Image Classification



Source: https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

END

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved.