

# Artificial Neural Networks

---

CONVOLUTIONAL NEURAL NETWORKS

# Topics

---

## 1. Introduction

- a) Motivation
- b) Drawbacks of multilayer perceptron networks

## 2. Convolution Neural Networks

- a) Feature Extraction by Convolution
- b) Pooling
- c) Classification

# Introduction

---

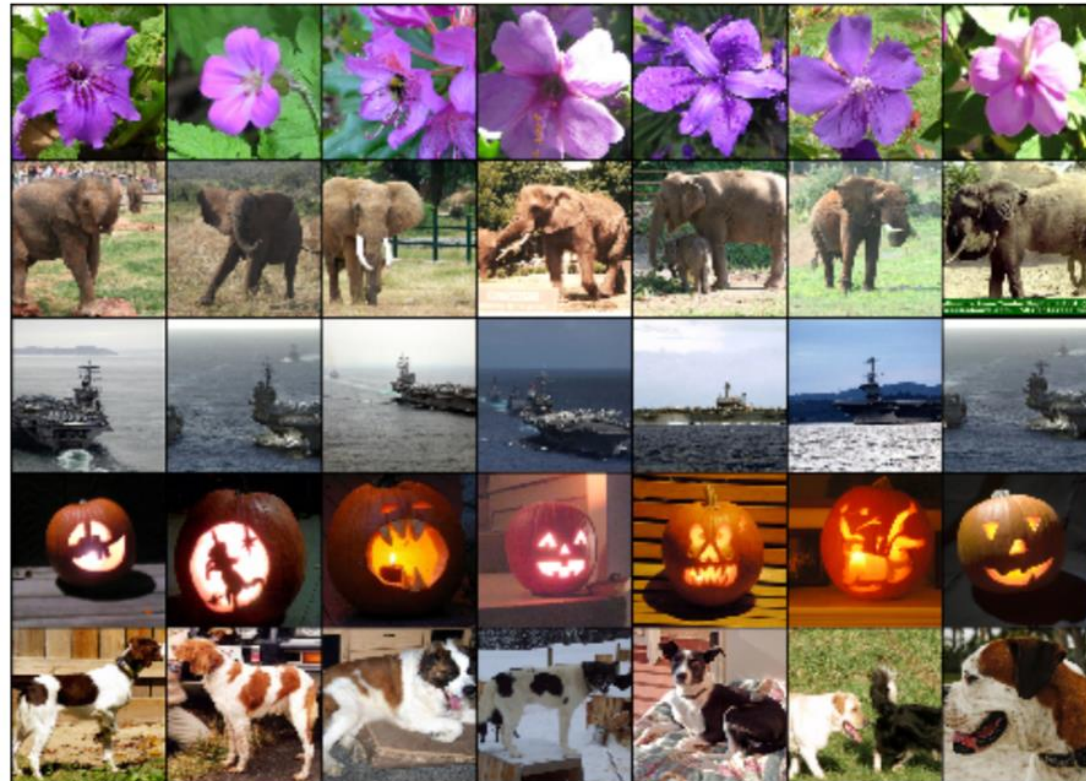
The concept of convolutional neural networks was introduced in 1995 by Yann LeCun and Yoshua Bengio at the New York University.

Convolutional neural networks are mostly used for signal and image processing applications.

# Introduction

---

Sample images classified by convolutional neural networks





**mite**



**container ship**



**motor scooter**



**leopard**

	mite
	black widow
	cockroach
	tick
	starfish

	container ship
	lifeboat
	amphibian
	fireboat
	drilling platform

	motor scooter
	go-kart
	moped
	bumper car
	golfcart

	leopard
	jaguar
	cheetah
	snow leopard
	Egyptian cat



**grille**



**mushroom**



**cherry**



**Madagascar cat**

	convertible
	grille
	pickup
	beach wagon
	fire engine

	agaric
	mushroom
	jelly fungus
	gill fungus
	dead-man's-fingers

	dalmatian
	grape
	elderberry
	ffordshire bullterrier
	currant

	squirrel monkey
	spider monkey
	titi
	indri
	howler monkey

# Motivation

---

Convolutional Neural Networks (CNN) were biologically motivated by the findings of locally sensitive and orientation-selective nerve cells in the visual cortex.

The visual cortex contains a complex arrangement of cells which are sensitive to small sub-regions of the visual field, called a receptive field.



# Motivation

---

The sub-regions are tiled to cover the entire visual field.

These cells act as local filters over the input space and are well-suited to exploit the strong spatially local correlation present in natural images.

Additionally, two basic cell types have been identified:

- Simple cells respond maximally to specific edge-like patterns within their receptive field.
- Complex cells have larger receptive fields and are locally invariant to the exact position of the pattern.

# Comparison with Multilayer Perceptron (MLP)

---

CNNs are a special form of multilayer perceptron networks.

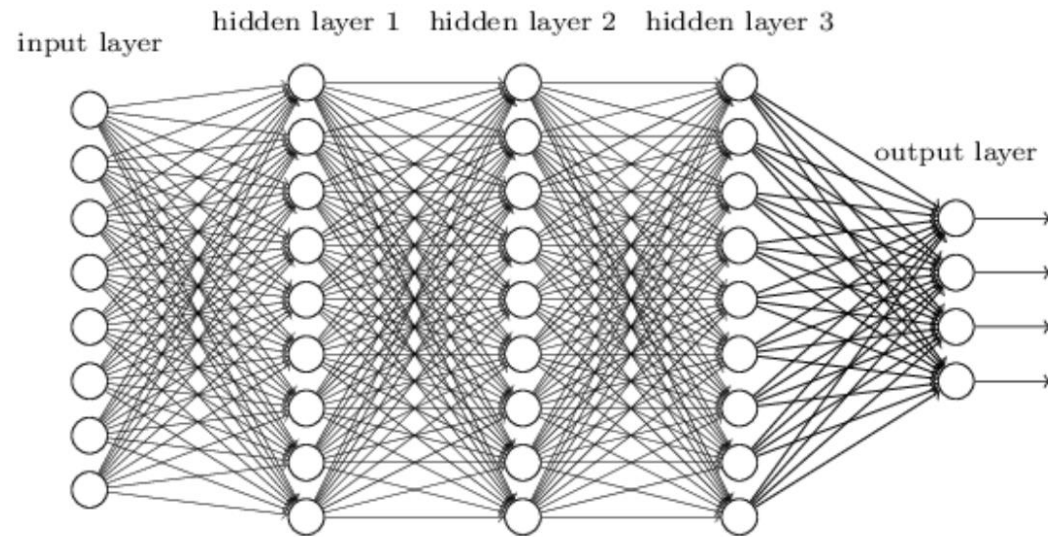
However, CNNs have some advantages over multilayer perceptron neural networks.



# Drawbacks of MLP

---

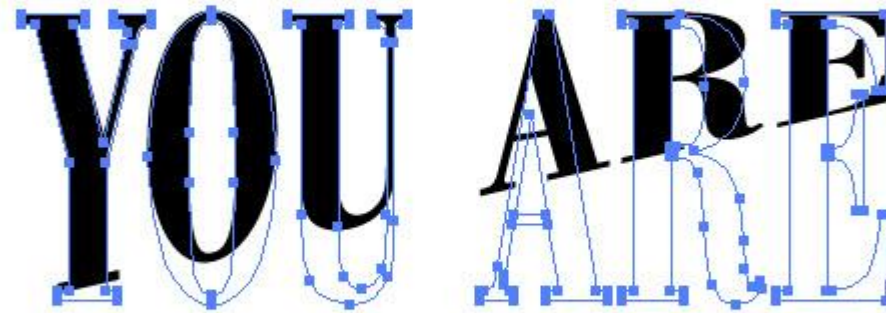
The number of trainable parameters grows exponentially to extremely large values



# Drawbacks of MLP

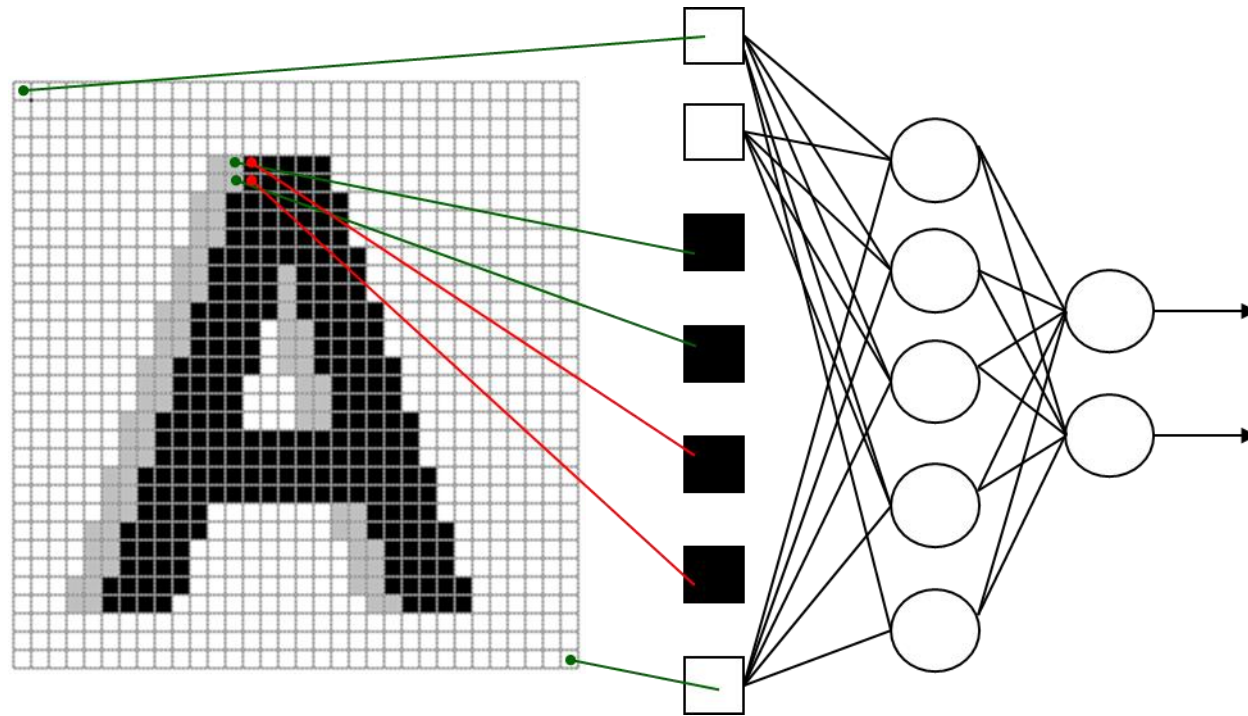
---

Sensitivity to scaling, shifting, and rotation in input data



# Drawbacks of MLP

---



# Features of CNNs

---

CNNs are feed-forward network and can extract topological properties from input data.

CNNs are trained with a modified version of the back-propagation algorithm.

CNNs can recognize visual patterns with extreme variability directly from pixel images with minimal preprocessing.

# CNN Structure

---

A CNN consists of

- A preprocessor which extracts the features
- A classifier which uses the extracted features to identify the input

# Feature Extraction in CNNs

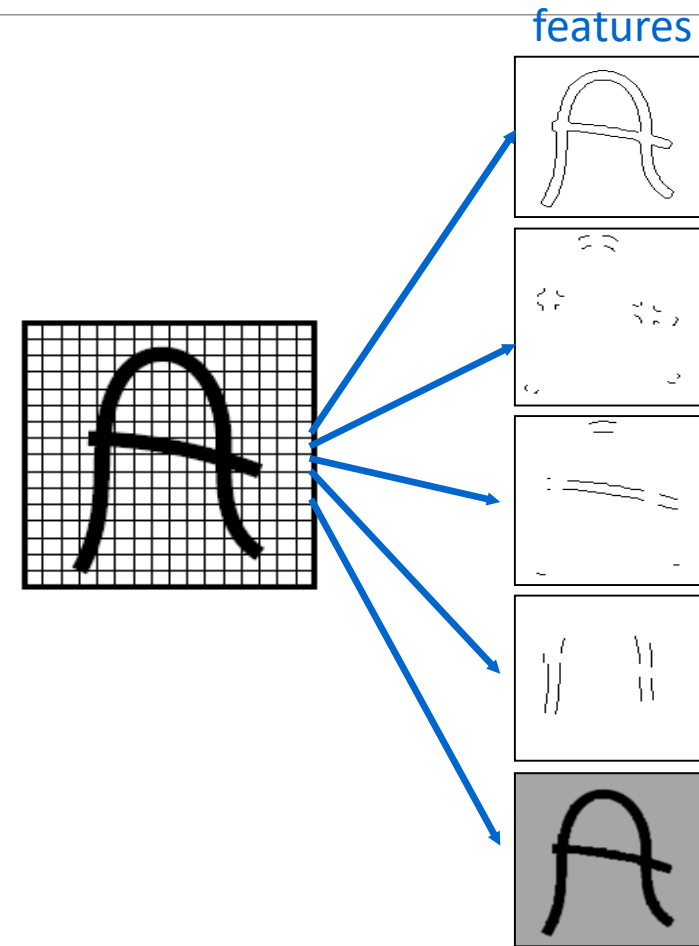
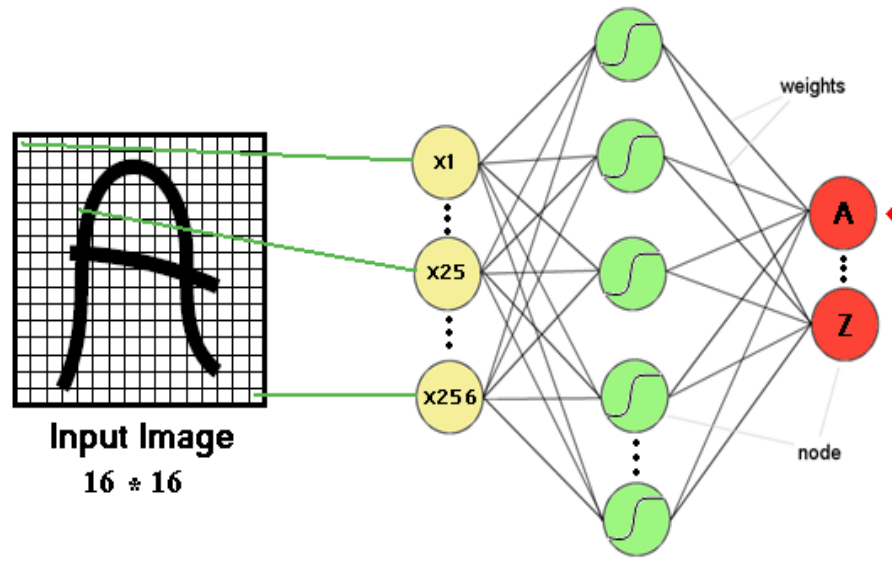
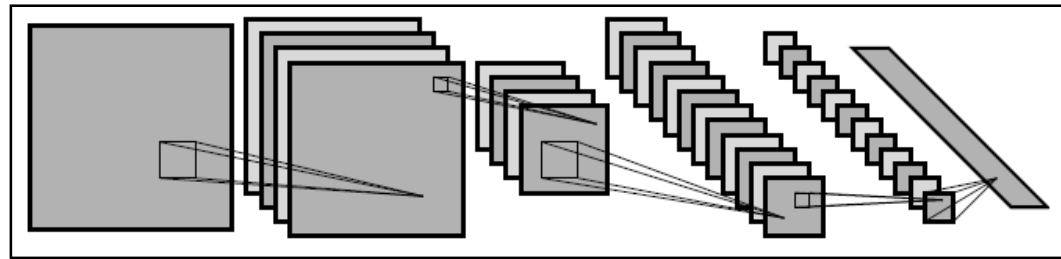
---

The input data (image for instance) is divided into sub-regions (not necessarily disjoint).

Each sub-region is convolved with a pattern

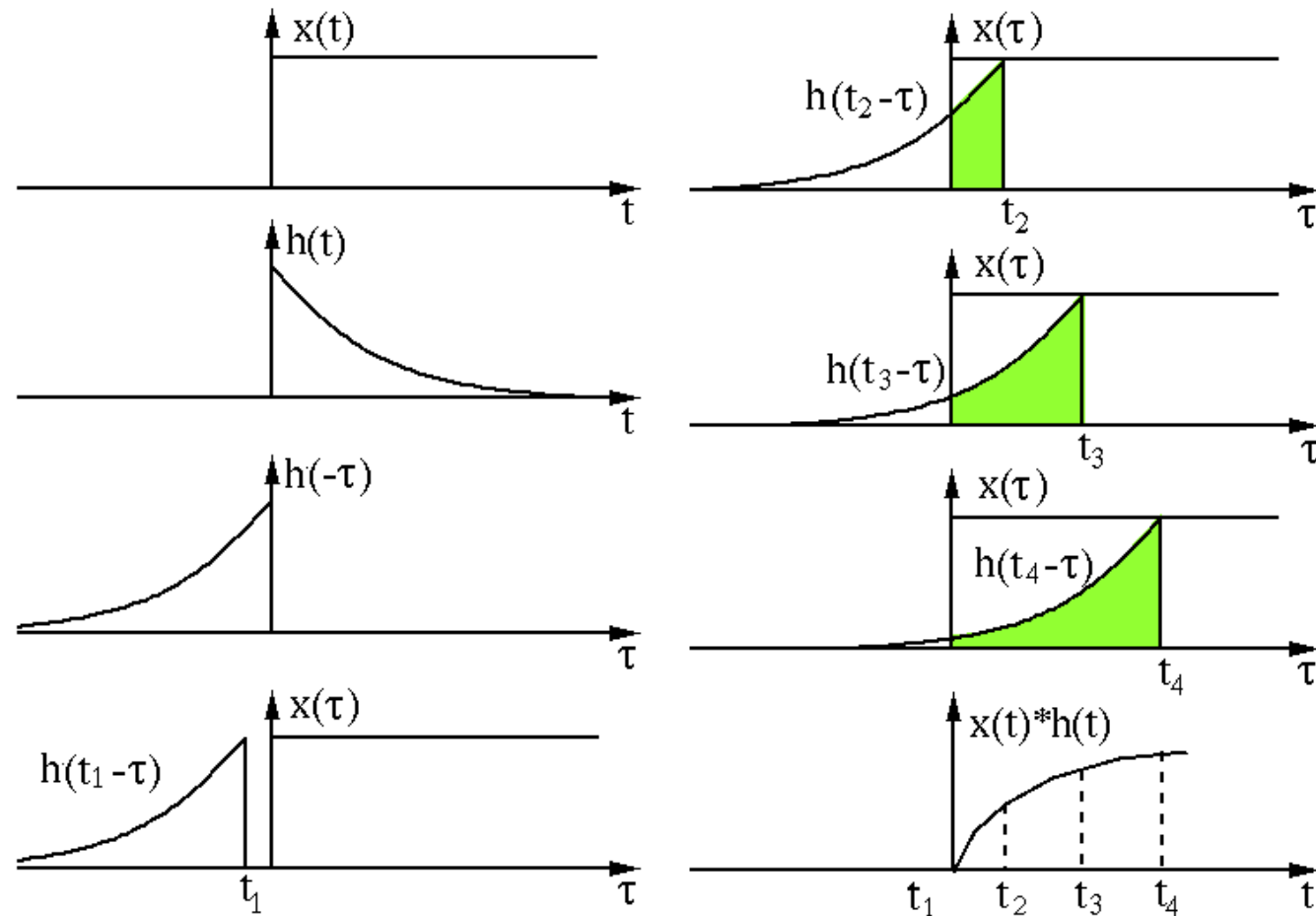
The results are grouped and provided as input to the classifier layers

# Feature extraction layer or Convolution layer

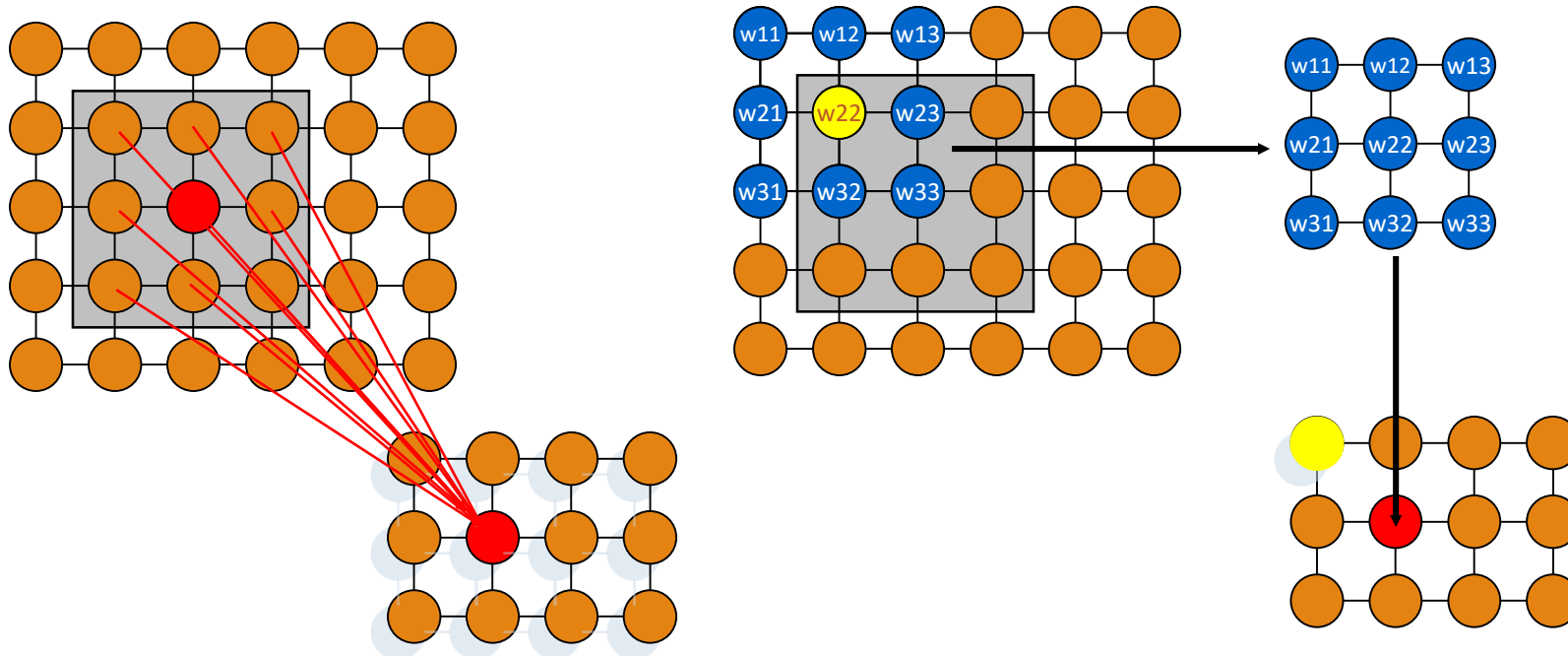
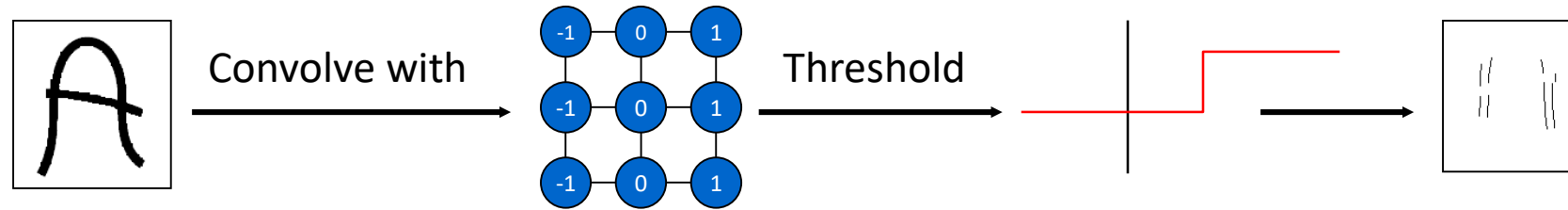




# Convolution Operator



# Convolution Operation in 2-Dimensional Images



# Sparse Connectivity

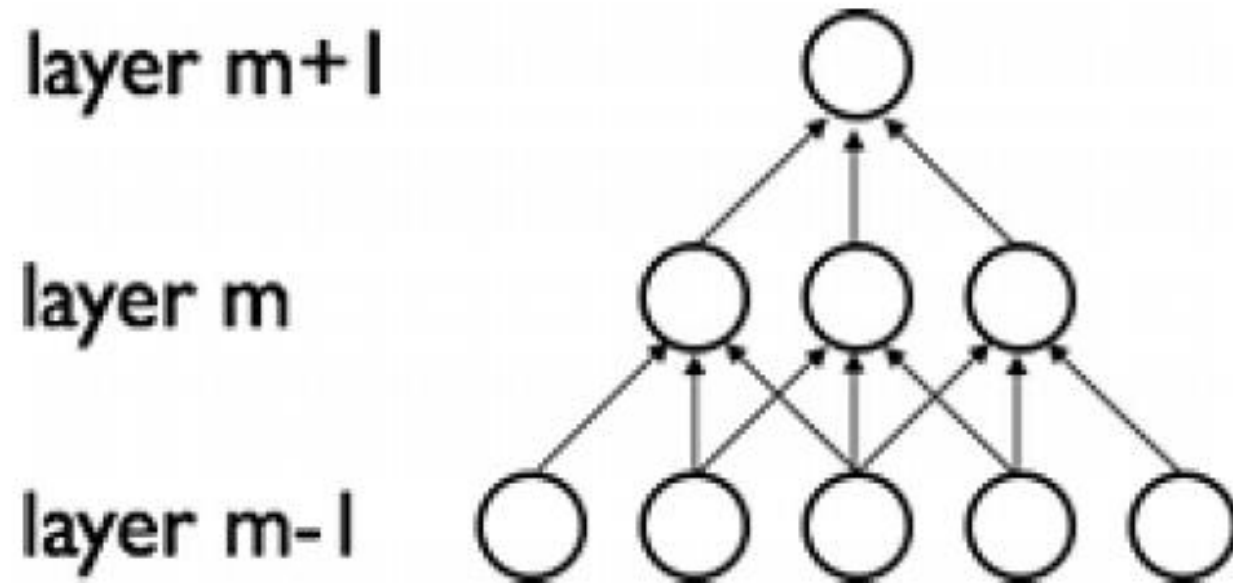
---

CNNs exploit spatially-local correlation by enforcing a local connectivity pattern between neurons of adjacent layers.

In other words, the inputs of hidden units in layer ***m*** are from a subset of units in layer ***m-1***, which have spatially contiguous receptive fields.

# Sparse Connectivity

---



# Feature Maps

---

A feature map is obtained by repeated application of a function across sub-regions of the entire image.

In other words, by convolution of the input image with a linear filter, adding a bias term and then applying a non-linear function.

$$h = f(W.X_{sub-region} + b)$$

# Max-Pooling

---

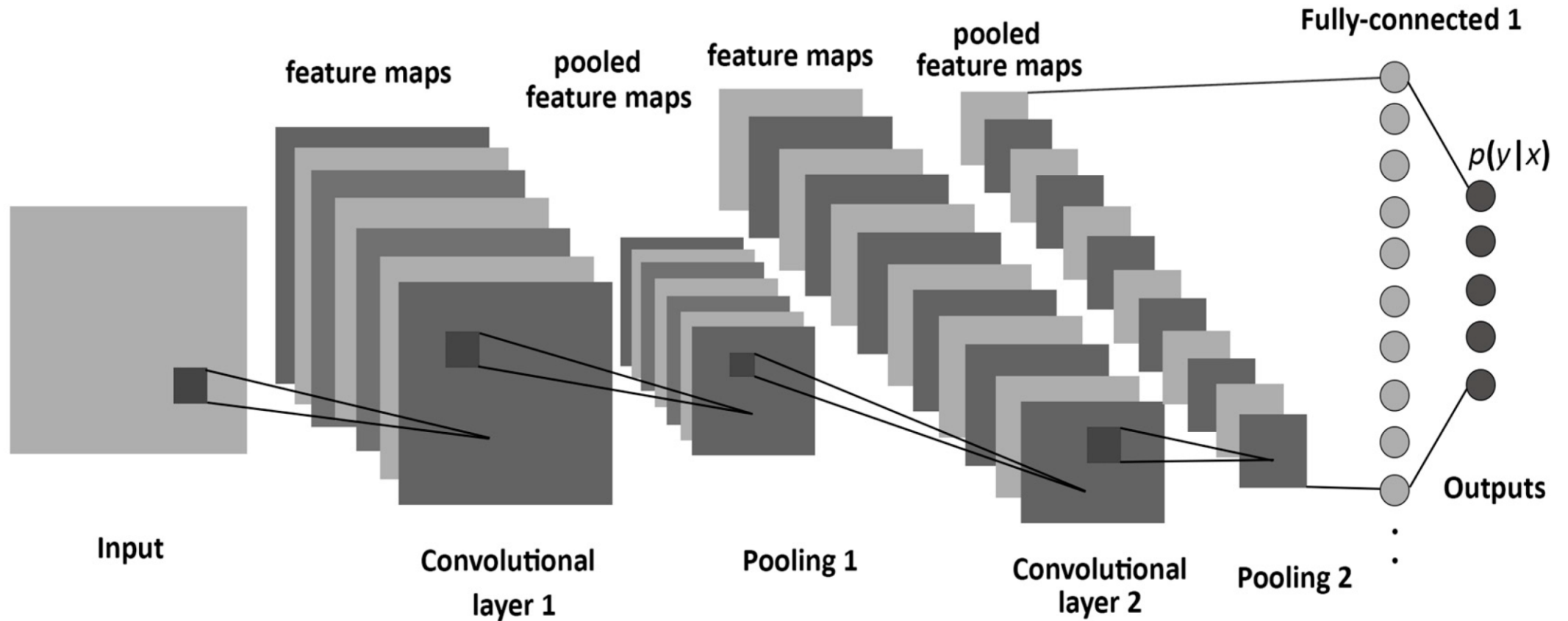
Another important concept of CNNs is max-pooling, which is a form of non-linear down-sampling.

Max-pooling partitions the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum value.

Max-pooling is useful in vision for two reasons:

1. By eliminating non-maximal values, it reduces computation for upper layers.
2. It provides a form of translation invariance. Imagine cascading a max-pooling layer with a convolutional layer.

# Convolutional Neural Network





# Assignment

---

Consider the matrix shown below.

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

Apply this matrix as a filter to the sample image given in the webpage of the course. (convolution). Consider both overlapping and non-overlapping blocks

Repeat the convolution by rotating the matrix one position clockwise.

Convolve the matrix after each rotation.

Discuss the results.

Deadline: December 20<sup>th</sup>, 2017