8. Image Classification using Convolutions (CNN)

Prof. Marcelo José Rovai
rovai@unifei.edu.br

UNIFEI - Universidade Federal de Itajubá, Brazil
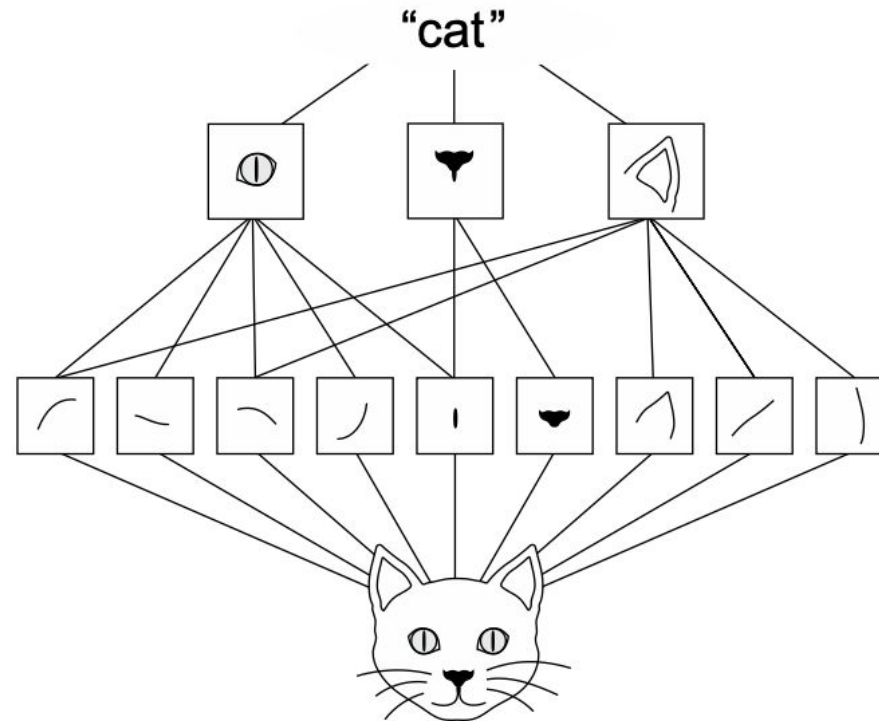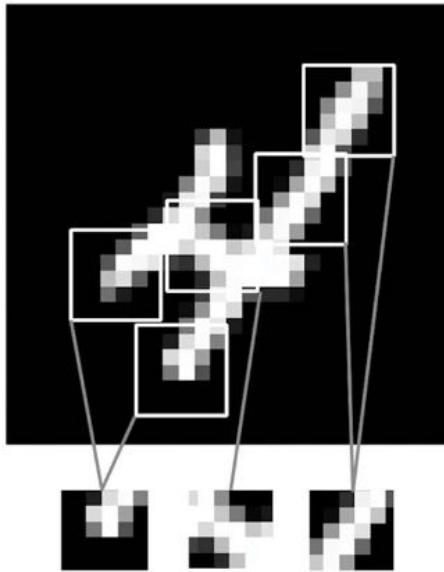
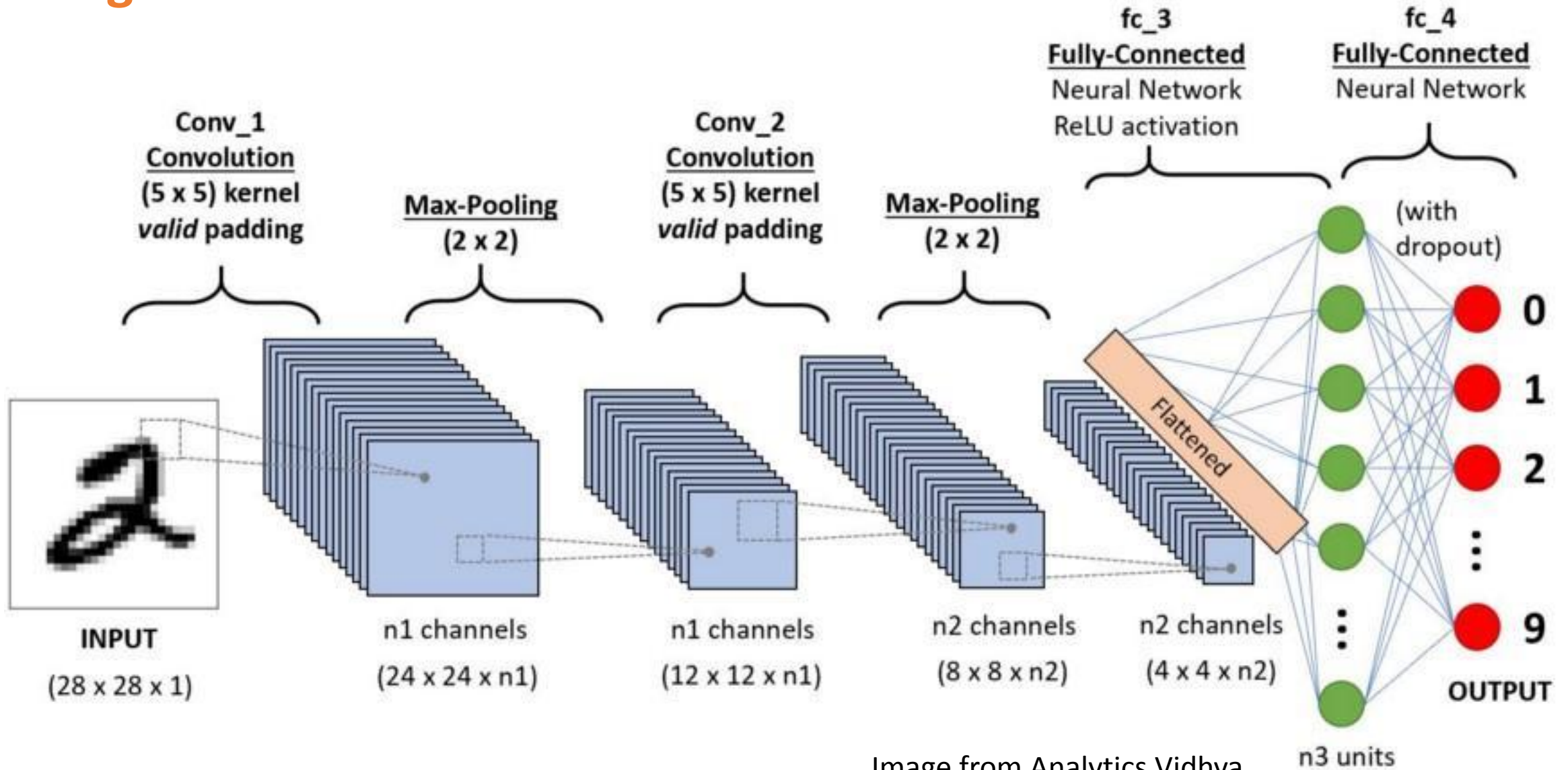# Convolutions

https://poloclub.github.io/cnn-explainer/

The convolution operation

The fundamental difference between a densely connected layer and a convolution layer is this: **Dense layers learn global patterns** in their input feature space (for example, for an MNIST digit, patterns involving all pixels), whereas **convolution layers learn local patterns**—in the case of images, patterns found in small 2D windows of the inputs In the previous example, these windows were all 3 × 3.
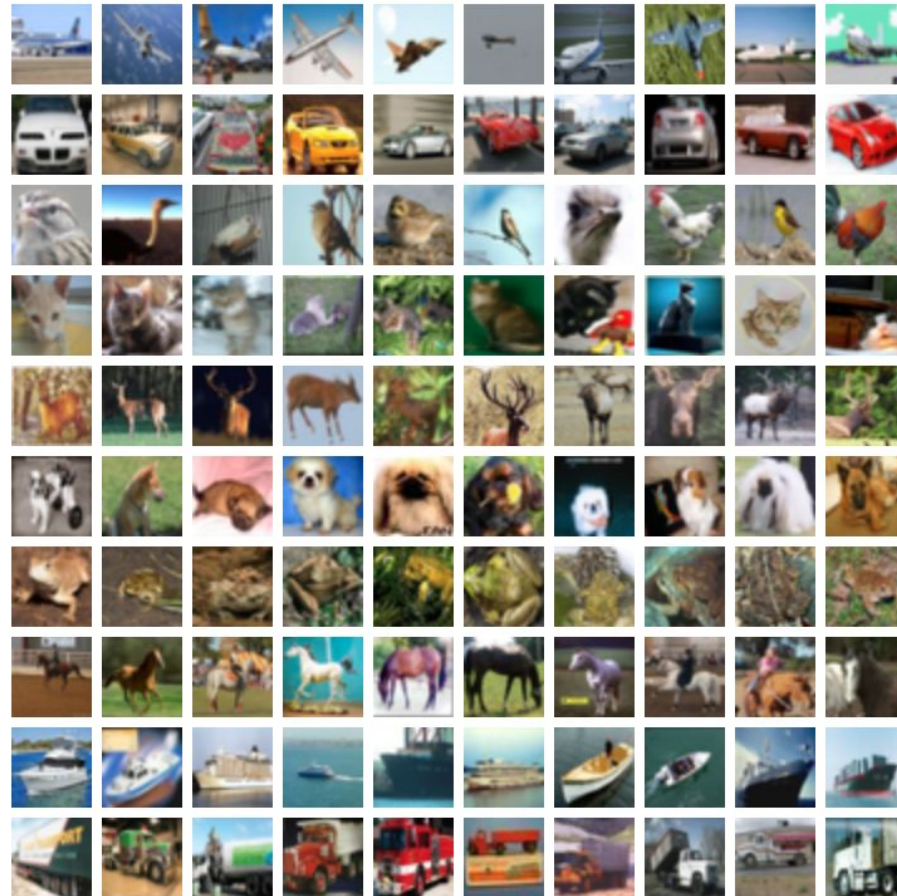

They can learn spatial hierarchies of patterns. A first convolution layer will learn small local patterns such as edges, a second convolution layer will learn larger patterns made of the features of the first layers, and so on.

"Deep Learning with Python" by François Chollet

Image Classification with CNN

Conv_1
Convolution
(5 x 5) kernel
valid padding

Max-Pooling
(2 x 2)

Conv_2
Convolution
(5 x 5) kernel
valid padding

Max-Pooling
(2 x 2)

fc_3
Fully-Connected
Neural Network
ReLU activation

fc_4
Fully-Connected
Neural Network

(with dropout)

Flattened

INPUT
(28 x 28 x 1)

n1 channels
(24 x 24 x n1)

n1 channels
(12 x 12 x n1)

n2 channels
(8 x 8 x n2)

n2 channels
(4 x 4 x n2)

n3 units

OUTPUT

0

1

2

9

Image from Analytics Vidhya

Cifar-10

0 airplane
1 automobile
2 bird
3 cat
4 deer
5 dog
6 frog
7 horse
8 ship
9 truck

https://www.tensorflow.org/datasets/catalog/cifar10

So far …

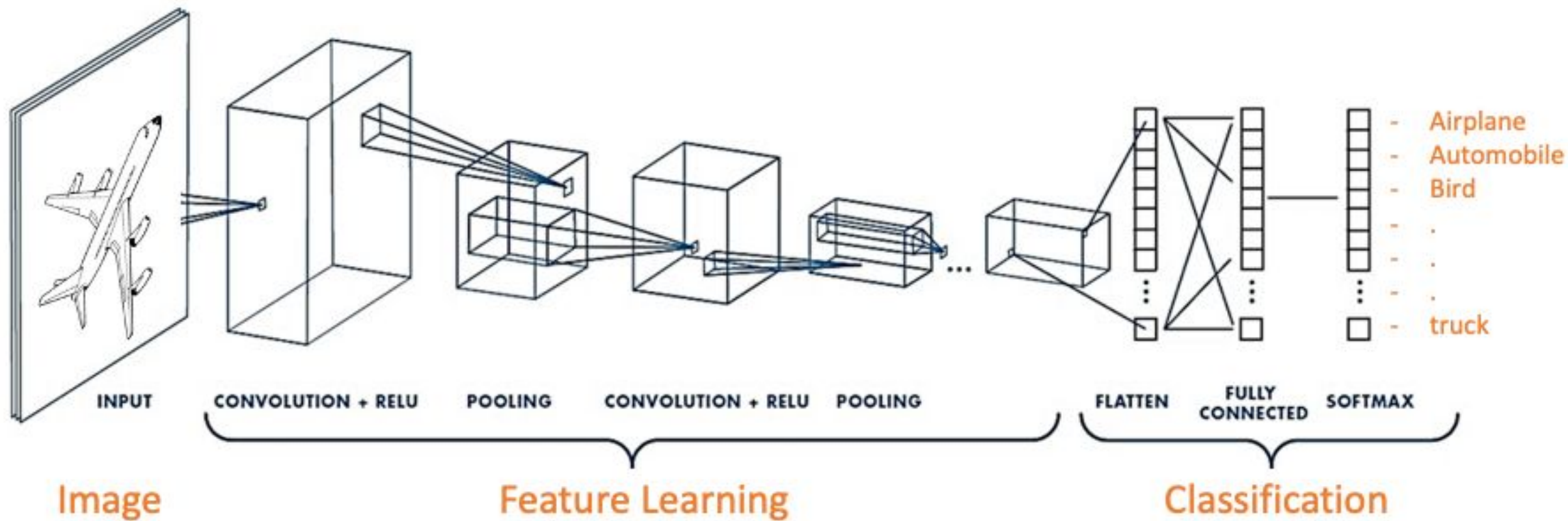We saw how to build Neural Networks (DNN) that classify images of digits (**MNIST**).

Now,

We will instead, recognize the 10 classes of **CIFAR** ('airplane', 'automobile', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship' and 'truck').

There are some key differences between these image datasets that we need to take into account:

- While MNIST has 28x28 monochrome images (1 color channel), CIFAR has 32x32 color images (3 color channels).
- Besides, MNIST images are simple, containing just the object centered in the image, with no background. Conversely, CIFAR ones are not centered and can have the object with a background, such as airplanes that might have a cloudy sky behind them!

Those differences are the main reason to use a CNN instead of a DNN.

INPUT   CONVOLUTION + RELU   POOLING   CONVOLUTION + RELU   POOLING   FLATTEN   FULLY CONNECTED   SOFTMAX

- Airplane
- Automobile
- Bird
- .
- .
- .
- truck

Image   Feature Learning   Classification

Image Classification using CNN
Code Time!

CNN_Cifar-10.ipynb

Thanks

UNIFEI

WALC 2022
Panamá
Presencial y Online
7 al 11 de noviembre

25°
Taller sobre
Tecnologías de Redes Internet
para América Latina y el Caribe

Organización General

30° Fundación
EsLaRed