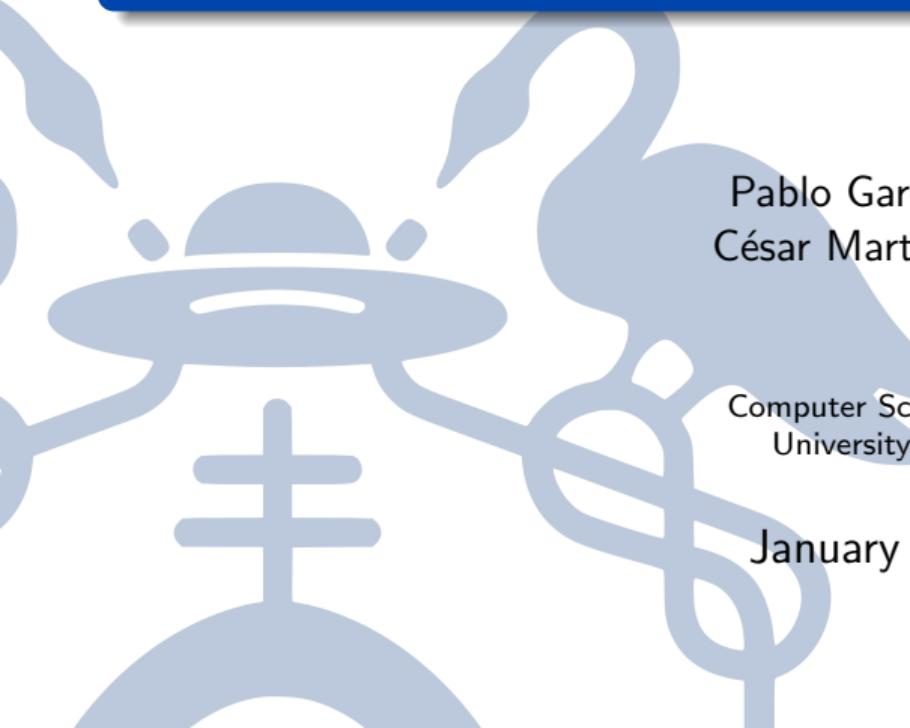


Computer Vision

Object detection using R-CNN Deep Learning



Pablo García García
César Martín Guijarro

Computer Science Degree
University of Alcalá

January 15, 2024

Table of contents

1 Introduction

2 Research

- How a CNN works
- MATLAB example
- How to make new datasets
- CNN over our own datasets
- Improvements

3 Conclusions

Introduction

- Research and understand CNNs.
- Detect objects inside of an image using *Matlab*.
- Expand and improve beyond *Matlab's* example.

Convolutional Neural Networks

CNN are just a type of **Neural Networks** that work better than usual Feedforward Networks in **CV problems**. Their main operation is called convolution.

Convolution

$$(f * g)(x, y) = \iint_S f(x - u, y - v)g(u, v)dudv$$

Feedforward networks input
Vector

Convolutional networks input
Tensor

Convolutional Neural Networks

Convolutional layer

The convolutional layer gives the net its name. Convolution with its parameters is applied to each element in the tensor, usually every RGB component.

Convolution parameters

- **Kernel:** $\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$, e.g. Represents the knowledge of the net.
- **Padding:** add n zeros rows and columns.
- **Pooling:** *maxpooling*, e.g.
- **Stride:** kernel moves by n pixels.

Convolutional Neural Networks

Convolutional layers in MATLAB

① convolution2dLayer

- ① Sets padding size
- ② Sets stride size
- ③ Performs convolution with each filter
- ④ Add bias

② reluLayer: performs ReLU operation

③ maxPooling2dLayer

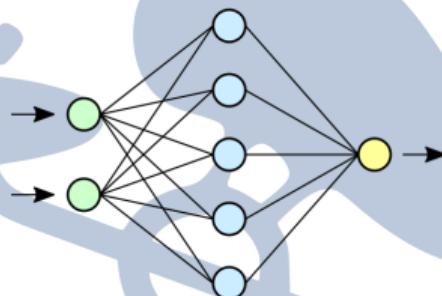
- ① Sets padding size
- ② Sets stride size
- ③ Performs pooling

$$d = \frac{n - k + 2p}{s + 1}$$

Convolutional Neural Networks

Neural network layer

After several convolutional layers, a usual neural network with **Backpropagation** appears to classify, take decisions, etc. Then a **softmax** normalization is performed.



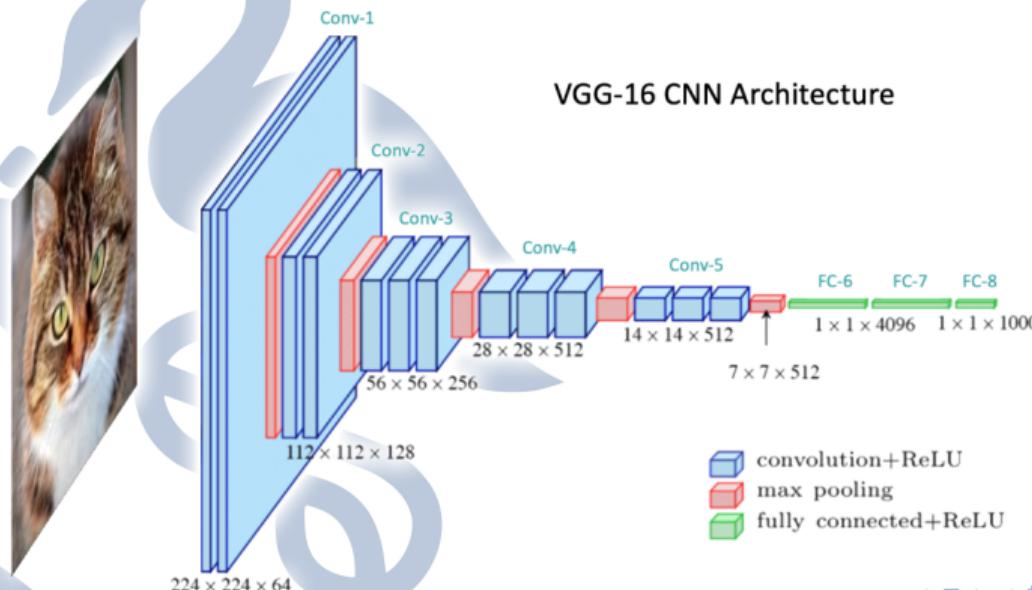
$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

In MATLAB, several `fullyConnectedLayer`, `reluLayer`, and then `softmaxLayer`, `classificationLayer`.

Convolutional Neural Networks

Final architecture

To sum up...



STOP signs example



Make new MATLAB datasets

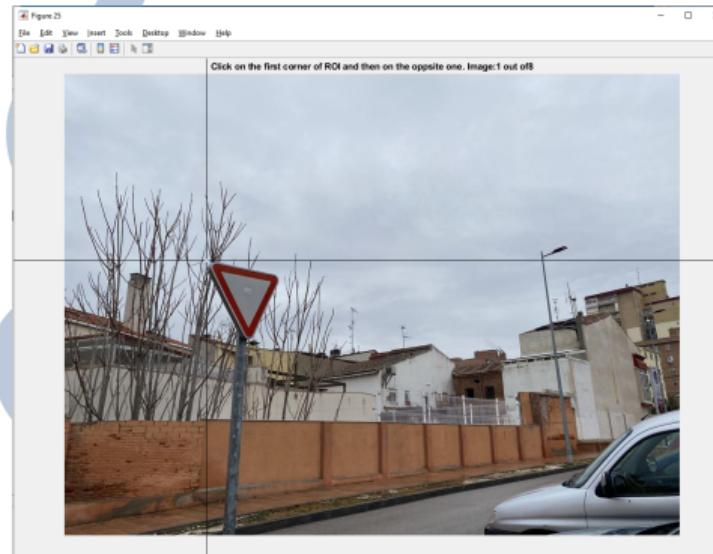
Introduction to the problem

- Bounding boxes defined by a vector `[x, y, width, height]`.
- Inserting the values manually is extremely tedious.
- Create a custom script to make it easier for the user.

Make new MATLAB datasets

Explanation

Simple and intuitive way to generate datasets by clicking.



Pedestrian signs



Wrong way signs



Yield signs



Heatmap representation



Bad example

Traffic lights



Detection for Several Instances of the same Class

Introduction

- Until now, only one object could be recognized at once.
- We decided to challenge ourselves to detect more than one object at a time.
- We eventually came up with two approaches.

Detection for Several Instances of the same Class

First Approach

Instance amount is specified by the user: **5 instances.**



Detection for Several Instances of the same Class

Second Approach - I

Confidence threshold is specified by the user: threshold = 0.6.



Detection for Several Instances of the same Class

Second Approach - II

Confidence threshold is specified by the user: `threshold = 0.7.`



Multiple class identification

Introduction

- We were not able to use a single network to detect different classes with `trainRCNNObjectDetector`.
- We decided to **implement custom code** to detect **more than one class** by mixing together two or more different networks.

Multiple class identification

Example of our implementation

As it is shown in this picture, different types of signs and many signs can be detected in one single picture:



Conclusions & future work

