Computer Vision
hugomcp@di.ubi.pt, 2023/24

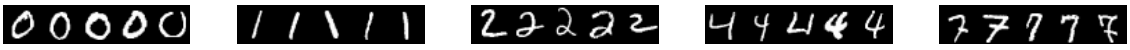
Ficha Prática 7

Self Supervised Learning

In this practical sheet, we will compare the effectiveness attained by a very simple CNN working in full supervised mode (varying the amount of labelled data), with a self-supervised learning (SSL) strategy followed by a supervised module.

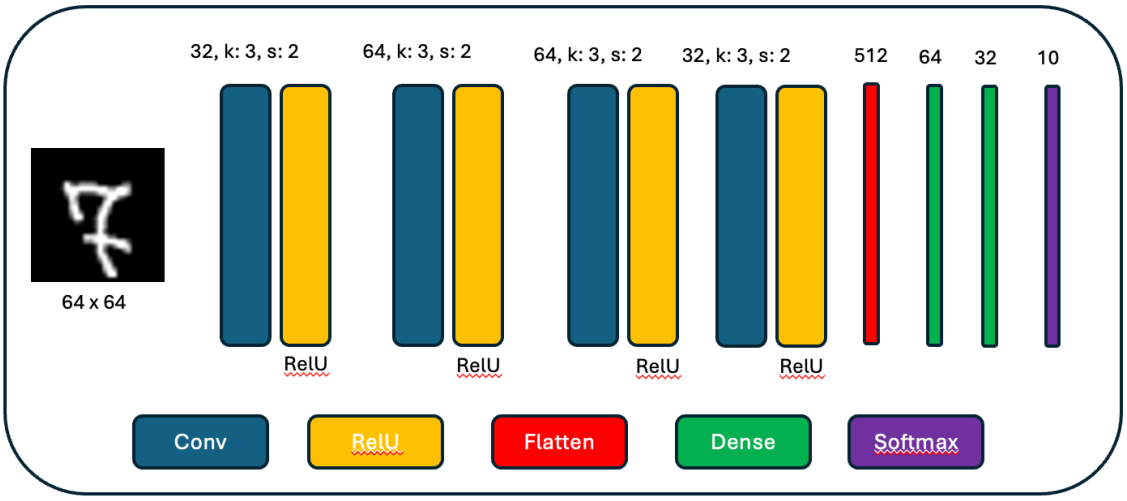
The goal is to perceive the advantages that SSL approaches might represent for particular cases, especially when the amount of labelled data for training purposes is relatively short.

Consider the "MNIST" dataset, available at the course web page in "png" format. This dataset contains grayscale images that regard 10 classes, the digits between "0" to "9".



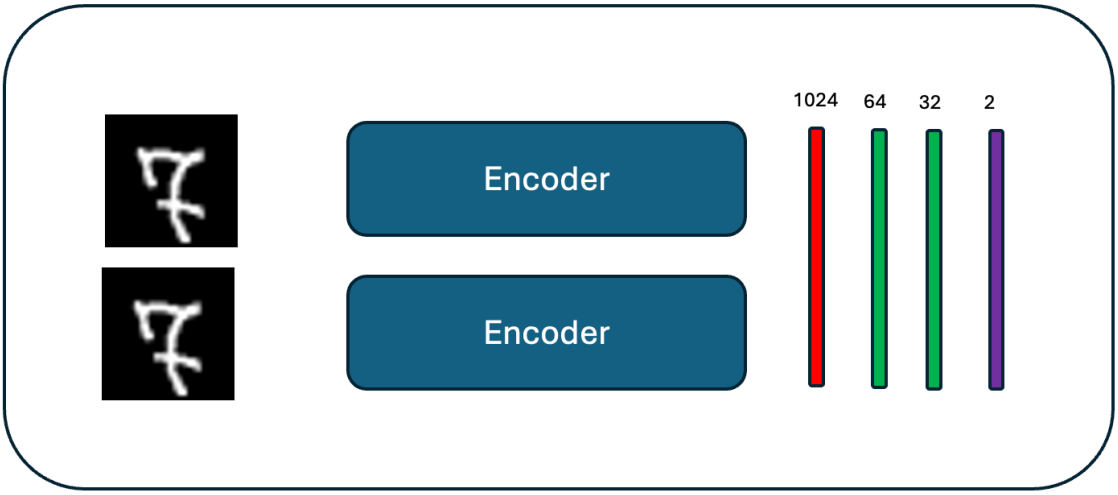
1) Create a python script that randomly divides the available data into three disjoint sets, one for supervised training (with 40% of the data), another one simulating unlabelled data (with 40% of the data) and the final one for testing (with the remaining 20%.

2) Create a simple CNN, according to the following scheme.



3) Train/test the above network, when using 10%, 20%,… 100% of the available learning data. Provide the performance in terms of "classification accuracy".

4) Under a SSL paradigm, create a network with the exact same architecture of the previous model, up to the "Flatten". Let's designate this network as "Encoder"



5) **Pretext Task:** Create a Siamese network that receives pairs of MNIST images. For pairs that regard the same image (rotated, flipped, blurred,…) this network should output "1". Otherwise, it should output "0".

6) Train the above model, using the "unlabelled" data subset.

7) Save the weights of the encoder learned in 6) and use them as initializers of the model created in 2).

8) Fine tune the model created in 7), using the "learning" subset, again when using 10%, 20%,… 100% of the available learning data.

9) Compare the performance obtained by the modelks in 8) and 3), w.r.t. to the amount of labelled data used.