

# Deep Learning - MAI

## Autonomous lab - CNNs

Dario Garcia Gasulla  
[dario.garcia@bsc.es](mailto:dario.garcia@bsc.es)

# The rules

- ❖ **Time:** 3 weeks to delivery (start now!)
- ❖ Done in **couples** (if possible)
- ❖ **Evaluation:** 30 minute interview (Why did you do 'X'? Why not 'Y'?)
  - Individual evaluation, non-transferable questions
  - Bring the material you used to take your decisions to support your work (loss/acc plots, histograms, confusion matrices, etc.)

# The task



- ❖ **MAMe**: Museum Art Medium dataset
  - High-resolution, variable shape images (*LR&FS* 256x256 available)
  - 29 classes of materials and techniques
  - Train: 700 samples/class, Validation: 150 samples/class, Test: varies
- ❖ Get the highest possible accuracy
- ❖ Run the test set once, the last thing before the interview



# The links

## ❖ **MAMe**: Museum Art Medium dataset

- *High-resolution, variable shape images:*

[https://storage.hpai.bsc.es/mame-dataset/MAMe\\_data.zip](https://storage.hpai.bsc.es/mame-dataset/MAMe_data.zip)

- *Labels:*

[https://storage.hpai.bsc.es/mame-dataset/MAMe\\_metadata.zip](https://storage.hpai.bsc.es/mame-dataset/MAMe_metadata.zip)

- *Low resolution, fixed shape:*

[https://storage.hpai.bsc.es/mame-dataset/MAMe\\_data\\_256.zip](https://storage.hpai.bsc.es/mame-dataset/MAMe_data_256.zip)

# The tip

- ❖ Focus on the low resolution version
- ❖ Only if you find you reached the limit... (ca. 75%-80%)
  - Consider using higher resolutions (training time will go UP!)
  - Consider avoiding deformation
- ❖ Try your own design using the methods introduced in theory

# The DO NOTs

- ❖ Do not use pre-trained models
- ❖ Do not replicate well-known designs
- ❖ Do not use external data
- ❖ Do not share code
- ❖ Do not wait until the last week

# The next steps

- ❖ Explore the data. Visualize it. See its distributions
  - ❖ Prepare the data pre-processing pipeline. This is the backbone.
  - ❖ Start with a small design. Underfit.
  - ❖ Grow. Overfit.
  - ❖ Regularize and reduce. Fit.
- Thoroughly document old experiments, and the evidence you use to decide the new ones.



# Take away

- ❖ The goal **is not** to produce the best possible model for the task
- ❖ The goal is to prove that you can
  - Decide coherently **which techniques** are most likely to maximize performance (easy)
  - Design and conduct **conclusive experiments** (medium)
  - **Diagnose the situation** of a training procedure (hardest)

**Dario Garcia-Gasulla (BSC)**  
*[dario.garcia@bsc.es](mailto:dario.garcia@bsc.es)*

