Introduction to Deep Learning

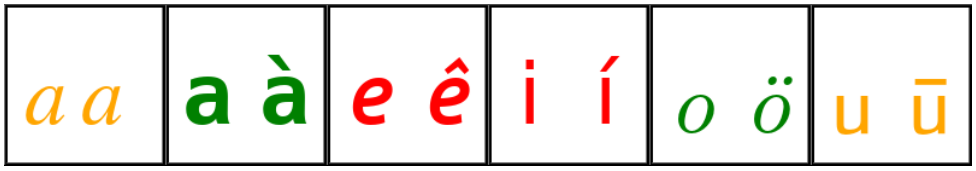
The project assignment consists in classifying different accent marks for letters in an image. You must work in teams of two or three people.

1 Datasets

The trainset contains 2.000 color images of size 128×128 pixels. Each image contains two letters in different fonts and colors. The two letters are always vowels, where the second one can contain an accent mark. The list of possible classes is the following:

$$\mathcal{L} = \{\text{aa, aà, aá, aä, aâ, aā, ee, eè, eé, eë, eê, eē, ii, ìi, íi, ïi, îi, īi,}$$
$$\text{oo, oò, oó, oö, oô, oō, uu, uù, uú, uü, uû, uū}\}$$

The next figure shows an example with 6 images from the trainset:



Each one of the 2.000 images from the trainset is identified with a numerical code, from 00000 to 01999. The combination of letters present in each image can be found in the file `images_train.csv`, which contains two columns: `id` and `values`.

Apart from the trainset, there is a testset with 5.000 images generated with exactly the same procedure used for the training images. The testset will be made public once all teams have trained their models.

2 Models

To classify the images you must train one or more models using Python and Tensorflow in a Jupyter Notebook.

You must consider the following restrictions:

- The models you train must be based on neural networks.
- You can apply preprocessing techniques to the images in order to improve the models performance.
- You are also allowed to create *ensemble* models, but all models in the *ensemble* must be based in neural networks.

3 Deliverables

The following list contains all the items you must deliver for this project.

- A presentation with 8 to 14 slides in PDF format, containing:
  - Name, surnames and NIU of every team member.
  - If you do a descriptive analysis of the dataset, the results and conclusions you obtained.
  - If you apply preprocessing techniques to the images, the detailed description of such process.
  - Description of the experiments you have carried out, together with the results and conclusions obtained.
  - Description of the final model you will use to classify the images in the testset.
  - Description of the hyperparameters you used during the training of the final model.
  - Estimation of the accuracy (in percent) you expect to get when executing the final model on the testset with 5.000 images.
  - Justification for your accuracy estimation.
  - Comments on what your needs would be to improve the accuracy. For example, would you like a larger data set? Or do you need a more powerful CPU? More RAM? More run time? Etc.
  - Any other comments on your development that you deem appropriate.
- Jupyter Notebook showing the construction and training of the final model, from the dataset reading process until the writing of the predictions file, detailing also the training of the model and all the processing you need on the images.
- CSV file with two columns, `id` and `values`, with the predictions of your final model for the images in the testset.

The deliverables must be uploaded to the Virtual Campus as follows:

- Before May 5th at 11:59pm you must submit your presentation and Jupyter notebook.
- On May 6th at 8:00am the testset with 5.000 images will be made public.
- You must submit the CSV file with your predictions for the images in the testset the same day (May 6th) before 11:59pm.

4 Grading

The grading of the project will be done on 10 points, and will consist of three items:

- N_1 (5 points) is the grading of the written presentation. The grade will be based on the clarity of the explanations and the justification of the decisions taken during the development of the project.
- N_2 (2 points) measures the the estimation of the accuracy. In particular, if p_j denotes the real accuracy obtained by one of the teams in the testset and $\hat{p}_j$ denotes the estimation of the accuracy written in the presentation, then,

$$N_2 = \begin{cases} 2 & \text{if} \quad |p_j - \hat{p}_j| \leq 1 \\ 1 & \text{if} \quad 1 < |p_j - \hat{p}_j| \leq 5 \\ 0 & \text{if} \quad 5 < |p_j - \hat{p}_j| \end{cases}$$

- N_3 (3 points) takes into consideration the actual accuracies obtained by all teams. In particular, assume we have G teams and we sort the accuracies from highest to lowest as follows:

  $$p_1 \geq p_2 \geq \cdots \geq p_i \geq 25 > p_{i+1} \geq \cdots \geq p_G$$

  Then, the grade for the team with accuracy p_j is

  $$N_3 = \begin{cases} \frac{2}{p_1 - p_i}(p_j - p_i) + 1 & \text{if } p_j \in \{p_1, \ldots, p_i\} \\ 0 & \text{if } p_j \in \{p_{i+1}, \ldots, p_G\} \end{cases}$$

The final grade will be then computed as

$$N = N_1 + N_2 + N_3$$