Deep and Reinforcement Learning
2024/2025 (M.IA003), FEUP/FCUP

Deep Learning Project
Submission deadline: March 30th, 2025

This work will need to be submitted using the Moodle platform. It will be developed during practical classes, but it is expected that the students will complement this work using extra-class hours.

1 Objective

The objective of this work is to develop deep learning **discriminative** and **generative** models, applied to the context of "deep fakes". The discriminative models will be designed to classify images as "real" vs. "fake", whereas the generative models will be trained to produce new "fake" examples.

2 Dataset

The data that you will be using belongs to the DeepFakeFace (DFF) dataset. You can access the dataset files and description via the Hugging Face link. In addition, you can find a detailed description of the dataset in this paper.

The dataset was generated to assess the ability of deepfake detectors to distinguish AI-generated and authentic images. It contains 30,000 real images of celebrities taken from the IMDB-WIKI dataset. The dataset also contains 90,000 fake images generative with the three following models:

- Stable Diffusion v1.5

- Stable Diffusion Inpainting

- InsightFace

Each model generated 30,000 fake images.

3 Implementation

In order to complete this work, you will need to implement **two different models**:

1. One classifier, which is trained to distinguish between real and fake images

2. One generative model, which is trained to create new fake images

For the first model, you will be free to implement any of the discriminative approaches that will be considered during the theoretical classes (e.g., multilayer perceptrons, convolutional neural networks, visual transformers, etc.).

For the second model, you will be free to implement any of the generative models that will be considered during the theoretical classes (e.g., generative adversarial networks, variational autoencoders, diffusion models, etc.)

For both models, you will need to define a **proper training strategy** as well as the correct way and **metrics** to evaluate the **performance**.

4 Project evaluation

The project will be evaluated by taking in consideration the **suitability** of the proposed model for the specific task, its **correctness**, and **complexity**.

VERY IMPORTANT: The main objective and core of this project will be that of iteratively improving the proposed solutions via a continuous observation of intermediate results and the proposal of adjustments to the algorithm. **Projects that simply present a solution without showcasing the evolution of the proposed model will not be considered as sufficient**.

5 Bonus

The bonus part of the work will be evaluated up to 1.5 pointa (over the 14 points reserved for the work). The maximum grade for this project is, nevertheless, 14.

It will be possible to obtain bonus evaluation by considering the implementation of one or more of the following techniques for work:

- Self-supervised learning

- Uncertainty quantification

- Federated learning

6 Submission of the solution

Your project must be delivered in Moodle by March 30th, 2025, at 23:59:59;
You are required to submit the following material:

- Final code solution, as a notebook or series of files.

- A **report** containing the following information:

    - Brief description of the deep learning solutions considered for the problem, both for the discriminative and generative part.

    - **Most Important**: Description of the different implementation steps considered to improve the proposed models: Motivate your choices in terms of type of approach, model architecture, training strategy, etc.

    - Results:

        * Classification performance obtained by the developed discriminative model. Description of the experimental setup, train/val/test splits, and performance metrics chosen.

        * Data generation performance obtained by the developed generative model. Description of the experimental setup and performance metrics chosen.

    - Discussion and conclusions: Comments on the performance obtained and final remarks.

- Filled auto-evaluation file regarding the contribution of each member of the group.

Further information about the project submission and presentation:

- The code provided as the solution will need to allow to train the considered models and reproduce the results that you reported. **Please do not include dataset files.** You can assume I have local access to the dataset.

- The work must be done by groups of **3 people**. Groups formed by less than 3 people must be justified and approved before starting working.

- Delays in the submission will incur in a grade penalization and eventually in not accepting the work.

- All works must be presented on the practical classes taking place on April 4th and 11th. All group members must be present during the demonstration. **If a member of the group is not present to the work presentation, he will receive a zero grade for this work, thus implying failing to pass.**

- Each member of the group must comment on their contribution to the work, and must know exactly what the other members of the group have done. **Failing to describe in details what your solution is doing and why will determine a penalization in the overall evaluation of the project.**