Processing Big Data 2023/2024

Project

1. Data Description

Portugal went through an election process in March of 2024. In the weeks before the election day, all eight major parties participated in one-on-one debates, in a total of 28 debates across three Portuguese TV channels. The original videos can be found here.

For each frame in these videos, the following information was extracted:

- **detected objects –** an object detector processed each frame to identify objects. Its outputs comprise: a vector with the object position [x, y, width, height]; a string with the object class (up to 80 possible classes), and the class probability' score, which is indicative of the confidence of the detector;
- **detected human poses –** a pose detector extracted the 3D coordinates of 33 keypoints for each detected human, as well as their probability of being visible and of being present, leading to a matrix of 33 x 5 for each pose;
- **detected faces –** a face and emotion detector extracted:
  - **location** of each detected face, expressed by a vector [x1, x2, y1, y2] with coordinates of two corners of a bounding box;
  - **expression/emotion** given by a string, with up to 8 possible classes; and
  - **facial embedding vector** of dimension 128 characterizing the face;
- **image embedding vector** of dimension 1024 characterizing the entire frame. The vector is obtained from a convolutional neural network trained on a classification task (ImageNet).

All this information is stored in a (pandas) dataframe, saved as a pickle file for each video that you can download here**.** As an example, Figure 1 shows one frame overlayed with annotation of the detected objects (red), human poses (white), and faces (green).

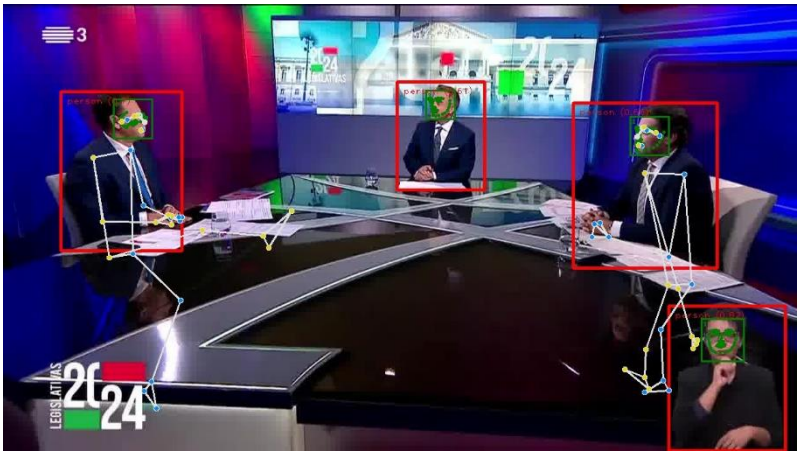

Figure 1: Example of some of the data extracted from the video frames.

2. Objectives

The data described above consists of a typical data science scenario, where we are given large amounts of data of different types and sources. The goal of this project is to learn how to handle this type of real-world data and deal with the challenges of extracting meaningful information from it. The project is structured such that students are required to practice all steps of the data processing "pipeline": Exploratory Data Analysis (EDA), Data Representation, Visualization, Modeling, Algorithm Design and Performance analysis. Furthermore, the project will show the reality of "big data processing": lack of a clear problem structure, heterogeneity of the data, huge dimension, and unreliable data (outliers).

To provide some guidance, the project is divided into three main tasks:

1. **EDA & Visualization** – familiarize yourselves with the data and perform exploratory data analysis: manipulate and normalize variables, compute summary statistics, visualize data.
2. **Single Video Analysis** – Choose one video and process its data separately from the other videos. Then:
    a. "Segment" the video according to the viewpoint/composition of each frame. Use the available information to partition the video into short clips with specific scenes, such as: viewing the two speakers and the moderator (as in Figure 1), camera focusing on speaker 1, etc. (suggestion: analyze the image embedding vectors);
    b. Assign a label ("person 1", "person 2", etc) to each detected face in the video. Using the information from the face detector (and any other you find relevant), find clusters of faces and check if they match unique people in the video.
3. **Multiple Video Analysis** – Find other interesting facts hidden in the data. Combine the information extracted from all the videos (and the video frames, if you find them relevant) and show what they tell us about: the debates, the political parties, the TV channels, etc. This task does not have a single clear solution or approach. Students are encouraged to explore and autonomously test multiple hypothesis/approaches/solutions and will be rewarded for their resourcefulness, creativity, and their outcomes or findings.

3. Evaluation

The project will be done in groups of 3 students. It accounts for a total of 65% of the final grade, and its evaluation is divided into: midterm review (15%), continuous evaluation during practical classes (10%), video report and code evaluation (20%), and presentation and discussion (20%).

Dates and deadlines

- The midterm review will take place in week 4 (May 6-10). This review will be a 15min discussion based on a brief presentation of the work you did in Task 1.
- Submission of your code – May 28 (midnight).
- Submission of a 15min video reporting your work and results – June 1 (midnight).
- Presentation (5min pitch, optional) and discussion (15min) will be in week 8 (June 3-7)