project

Description and Objectives:

The goal of Processing Big Data (PBD) project is to confront students with a typical “data science” problem and apply state-of-the-art “data science” methodologies as well as standard techniques learned in the course. Foundational courses such as linear algebra, programming, calculus or optimization are key so we encourage you to revise those materials.

We will do this training through a specific application of video analytics for cycling races that we will describe below.

The project is structured such that students are required to practice all steps of the data processing “pipeline” : Exploratory Data Analysis, 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).

In short, the project does not have a single clear solution or approach. Students are encouraged to explore and autonomously test multiple approaches/solutions though the context is a bit constrained to make it realizable. The application scenario is extracted from the Italian bicycle competition.

The Giro D’Italia: Analysis of RAI’s Media Coverage

Large outdoor events like the bicycle race “Giro D’Italia” follow complex but very well planned “scripts” by the producer/director of the coverage.

Fig1: skeleton extracted by openpose

In particular the director is managing the viewpoints originated from a multitude of video sources such as motorbikes, fixed cameras, drones or helicopters and at each time instant, he/she selects the best video feed or the most meaningful to the viewer.

Objectives and data

The original data is comprised of a set of videos taped from RAI's media coverage. These videos were processed and the available data is the following:

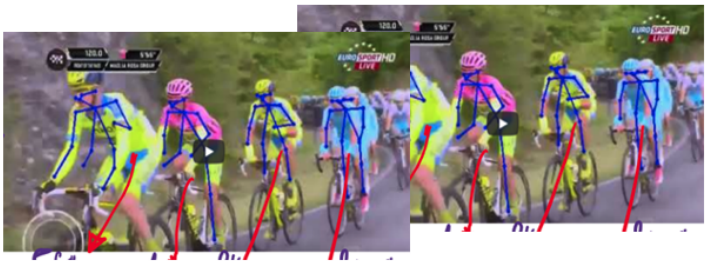
- **Visual embeddings:** A high dimensional vector (1024) for each image of the video. This is a representation provided by a Neuronal Network that encodes each image. This is to say that images can be compared in terms of similarity using these vectors. Details will be provided in the practical classes.
- **Skeletons :** Each image was processed and skeletons of visible people were extracted using the "open pose" algorithm. As fig 1 shows, depending on the number of visible persons and the algorithm reliability, a variable number of skeletons are extracted. A skeleton is a set of 18 2D coordinates encoding the image location of each joint. Details will be provided and the original representation is here https://github.com/CMU-Perceptual-Computing-Lab/openpose.

https://youtu.be/DEZ8yw78Y94

Keypoints of detected poses.

- **The original images** in case you want to extract further information from them.

In short, the minimum set of data is represented by a set of matrices as described below.



$$D = \left[\begin{array}{cccc|cccc} f & f & f & f & f+1 & f+1 & f+1 & f+1 \\ x_{1,1} & \vdots & \vdots & x_{1,N_f} & x_{1,1} & \vdots & \vdots & x_{1,N_{f+1}} \\ y_{1,1} & \vdots & \vdots & y_{1,N_f} & y_{1,1} & \vdots & \vdots & y_{1,N_{f+1}} \\ p_{11} & \vdots & \vdots & p_{1,N_f} & p_{11} & \vdots & \vdots & p_{1,N_{f+1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{18,1} & \vdots & \vdots & x_{18,N_f} & x_{18,1} & \vdots & \vdots & x_{18,N_{f+1}} \\ y_{18,1} & \vdots & \vdots & y_{18,N_f} & y_{18,1} & \vdots & \vdots & y_{18,N_{f+1}} \\ p_{18,1} & \vdots & \vdots & p_{18,N_f} & p_{18,1} & \vdots & \vdots & p_{18,N_{f+1}} \end{array}\right]$$

$$E = \left[\begin{array}{cccccccc} e_1 & \dots & e_f & e_{f+1} & \dots & e_F \end{array}\right]$$

The main data is comprised of a 55 x #detected poses (matrix D) where the first element in each column is the frame number and the remaining elements are triplets (xi,yi,pi) with the x,y coordinate of each point and the score (”probability”) of the detection. E is a 512 x F matrix where each column is a 512-vector representing each of the F images. These embeddings were created with a Neural Network (VGG) trained for object classification. The details of the VGG are irrelevant for the project and it is assumed that similar images have close embeddings. Note also that the length of the vector is not limited or constrained, but that you’ll find while doing EDA.

Main Objective

In general terms, the goal of the project is to segment the whole video by “type of shot” which is indexed by the pose of the cyclists and the contents of

the image. The figure below illustrates the main idea

There will be three levels of difficulty which involve three different types of data.

- 1 - Complete data: the skeleton data is complete, that is, the whole set of skeleton points is available for each detected person. This should be used only if you are not able to estimate the missing points.
- 2 - Incomplete data: skeleton detectors fail and sometimes only part of a body is visible therefore it happens that some points are missing. In fact, the previous dataset (completed) was created from this one using one particular data completion algorithm !
- 3 - Extra data: You can use any extra data extracted from the images, from the skeletons or crawled from the internet or any other source. This is left to each group's initiative. Students are allowed (and encouraged) to extract further data (e.g number of cyclists in images, frontal/side/back view …).

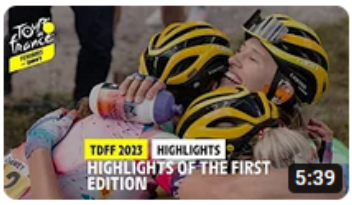
Analytics and application:

Provided you completed the main objective of processing single videos, the typical "datascience" application would be to analise and compare videos. Imagine there are different broadcast companies, different producers, different technical staff or different content. Given several videos of similar events the goal would be to provide analytics about them.

Specifically taking videos from the website https://www.youtube.com/@LeTourDeFrance we will extract the 3 videos of "highlights of xxxx" (2020,2021,2023) and you should do an EDA and explore the data you extract from them to characterize and compare them.

#TDF2020 - Le best-of de la 107e édition !

#TDF2021 - Highlights of the race



The highlights of the 1st edition of the Tour de Franc...

We will make the videos and its skeletons and embeddings available.

Programming and software tools

You can develop your project in Python or Matlab and use any toolbox that you may find useful (pandas, scipy, numpy, opencv, scikitlearn)