

Human action recognition

- **Goal:** Analyzing a sequence to identify the action taking place in the sequence.
- **Motivation:** Human action recognition remains a challenging task with great practical applicability
- **Challenges**
 - There are multiple modalities to represent information: RGB images, depth maps, skeletons
 - The presence of large variations in the execution of action

Human pose estimation

- **Goal:** Detecting 3D coordinates of a specific number of joints (keypoints) on the human body in a multi-person image.
- **Challenges**
 - We need to identify the number of people in the image.
 - Variation of body poses, complicated background or depth ambiguities



Explainable deep learning for video recognition tasks

- **Goal:** Expanding video classification approaches by adding an additional step in which the neural network learns to explain the predicted outcome by providing arguments to validate the network's decision.
- **Motivation:** This explicability property is fundamental for people who design, implement and test the system to enhance system robustness and enable diagnostics to prevent bias, unfairness and discrimination.

Temporal convolutional neural networks for sequence modelling

- **Goal:** Analyzing the results that can be obtained by an architecture that contains TCN type layers for a sequence modelling problem, by highlighting all the characteristics.
- **Motivation:** TCNs afford better control of the model's memory size, and are easy to adapt to different domains.

