

Cross multi-modal learning for routine discovery from egocentric photo- sequences

Human behaviour is determined by lifestyle and is closely related to the health of a person. In this project, we aim to explore the application of multi-modal deep networks for the learning of joint embeddings coming from visual and GPS collected data by a person throughout his or her days. This joint learning will represent a tool for the analysis and understanding of the lifestyle of the wearer. We will work with a novel and labeled egocentric dataset composed of egocentric images, which present a first-person view of the environments of the camera wearer, and GPS data collected by the user mobile phone. The combination of these two types of data will make the description of the days more robust. Therefore, in this project, we will evaluate the joint embeddings of the two types of data for the characterization and understanding of the lifestyle of a person.

Together with the data scientist, we will implement an ablation study of available models for the translation into features of the two types of data, and different decision fusion techniques. This will represent a new baseline in the egocentric vision field for the characterization of routine and lifestyle of the sensors wearer.