

Multimodal Deep Learning for Behavior Analysis

Supervisor: *Tanaya Guha (CS)*

Co-supervisor: *Victor Sanchez (CS)*

Background:

Our experience of the world is inherently multimodal. As humans we are able to relate and integrate information from multimodal sources seamlessly. For example, humans integrate audio-visual information for speech recognition and scene understanding. Deep learning has made tremendous advances in learning from a single modality, especially in the visual domain. This success has been a catalyst to solving larger and more complex problems involving data coming from multiple sensors or modalities (e.g. video and audio). Multimodal deep learning involves developing deep networks for efficiently integrating information from multiple modalities. Application domains include activity recognition, health monitoring, sentiment analysis and autonomous driving.

The project: The primary goal of the MSc. project is to develop a deep network (preferably, *graph convolutional network*, considering each sensor as a graph node) capable of combining information from multiple sensors to identify the human activity being performed with high accuracy. The challenge lies in combining the disparate data and creating a network that can handle missing data (cases when one of the modalities is missing during testing). The MSc project will use smaller databases such as, OPPORTUNITY [1] available from the UCI machine learning repository.

Deliverable: An accurate deep architecture (preferably, graph-based) that can classify different human activities from multimodal data.

Desirable skills: Python, Machine Learning.

Opportunities for a PhD: The mini project can be extended in many ways for a PhD. Multimodal learning is complex and emerging topic. One particular direction would be to develop larger and more sophisticated deep learning models (new architecture, new loss functions and optimization techniques) for multimodal learning from higher dimensional data, such as video, audio and other sensor data. One application domain of interest is autonomous driving. This work has collaboration opportunities with researchers from Intel AI and Ford.

Reference:

[1] <https://archive.ics.uci.edu/ml/datasets/OPPORTUNITY+Activity+Recognition>