Master Thesis: Energy-base Multi-Modal Attention
Name: Aurélien Werenne
Supervisor: Dr. Raphaël Marée
Section: Computer Engineering - Intelligent Systems   Academic Year: 2018-2019

Summary

A multi-modal neural network exploits information from different channels and in different terms (e.g., images, text, sounds, sensor measures) in the hope that the information carried by each mode is complementary, in order to improve the predictions the neural network. Nevertheless, in realistic situations, varying levels of perturbations can occur on the data of the modes, which may decrease the quality of the inference process. An additional difficulty is that these perturbations vary between the modes and on a per-sample basis. This work presents a solution to this problem. The three main contributions are described below.

First, a novel attention module is designed, analysed and implemented. This attention module is constructed to help multi-modal networks handle modes with perturbations.

Secondly, two new regularizers are developed to improve the generalization of the robustness gain on more intensive failing modes (relative to the training set).

Lastly, a unified multi-modal attention module is presented, combining the main types of attention mechanisms in the deep learning literature with our module. We suggest that this unified module could be coupled with a prediction model to enable the latter face unexpected situations, and improve the extraction of the relevant information in the data.