Master Thesis Proposal 2021/2022

Title: Audiovisual processing in the wild by using video-based lip-reading software

Supervisors: Emina Alickovic, Martin Skoglund (Adj. Assoc. Professors at LiU & Senior Researchers at Oticon A/S) & **Johannes Zaar** (Oticon A/S)

Project Description: This project will develop a novel machine learning based lip-reading technique capable of classifying the discrete utterances without having access the acoustic signals (i.e., speech). The aim of this project is to recognize the words coming from a talking face, given access only to the video but not to the audio files. The second aim is to check whether the proposed method generalizes across different speaker variations. Our ultimate is to use methods developed in this project to obtain audiovisual speech from the brain activity recorded with electroencephalography (EEG) instruments.

Program Duration: 20 weeks, 30HP, with a flexible starting date.

Contact: emina.alikovic@liu.se and/or martin.skoglund@liu.se

Relevant Literature:

[1] Noda, Kuniaki, et al. "Lipreading using convolutional neural network." fifteenth annual conference of the international speech communication association. 2014.

[2] Assael, Yannis M., et al. "Lipnet: Sentence-level lipreading." arXiv preprint arXiv:1611.01599 2.4 (2016).

[3] Chung, Joon Son, and Andrew Zisserman. "Lip reading in the wild." Asian conference on computer vision. Springer, Cham, 2016.