NEWSLETTER

Centre for Acoustic Signal Processing Research (CASPR)

January 2023

The Centre for Acoustic Signal Processing Research (CASPR) is a research centre at the Section for Artificial Intelligence & Sound, Department of Electronic Systems, Aalborg University, Denmark. CASPR is supported by the Demant Foundation, Oticon A/S, and Aalborg University.

Contents

Department of Electronic Systems
Section on Artificial Intelligence & Sound
Aalborg University

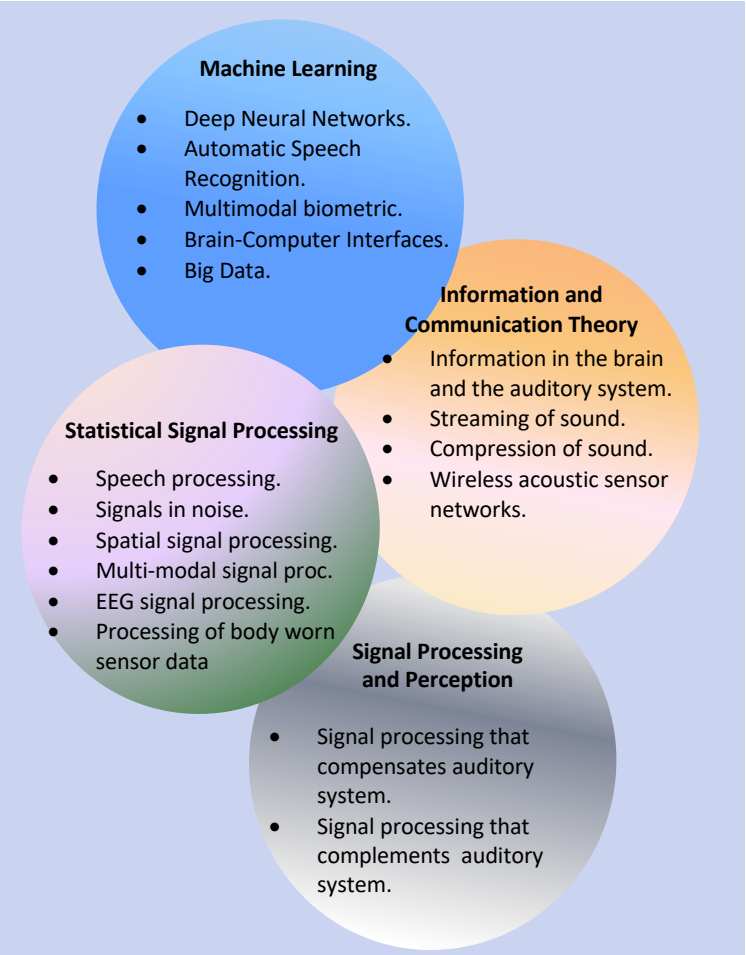
Research in CASPR

CASPR is conducting research related to advanced statistical signal processing solutions for assistive listening. The research finds direct use in communication devices such as hearing aids, helmets, headphones, cochlear implants, ear monitor, ear pieces, diagnostic equipment, etc. However, the envisioned research also finds use in related areas such as robust speech/speaker recognition, brain computer interfaces, acoustic event detection, etc.

CASPR will navigate in a rapidly changing technological landscape: we envision a near future, where the technological landscape allows very different, and better, hearing assistive devices than are known today.

The scientific scope of CASPR encompasses

- Signal processing for hearing assistive devices that use multiple modalities such as sound, vision, EEG, and other body signals.
- Advanced signal processing concepts in closer symbiosis with the user in order to e.g., automatically adapt to the user's needs.
- Beyond audibility. Restore audibility, increase intelligibility, decrease listening effort.



Scientific disciplines of CASPR

Teaching in CASPR

Current courses related to CASPR

CASPR is heavily involved in teaching and education at B.Sc., M.Sc., PhD., and Postdoc level in disciplines that are relevant to the scientific scope of CASPR:

- Control and optimization, PhD course.
- Machine learning, Master and PhD courses.
- Deep learning, PhD course.
- CASPR Seasonal Schools on Signal Processing for Hearing Assistive Devices.
- Information theory, Master course.
- Array and Sensor Signal Processing, Master course.
- Advanced Signal Processing, Master course.
- Platforms and methods for multi-modal system architectures, Master course.

During Autumn 2022, CASPR was involved in 6 student projects

- Self-supervised Learning for Algorithm Unrolling.
  Mathematical Engineering, long Master thesis project.
  Mads Arnløv Jørgensen, Magnus Jónhardsson, and Andreas Larsen.

- Robust Keyword Spotting Using Self-Supervised Deep Learning.
  Signal Processing and Acoustics, long Master thesis project.
  Gergely Kiss.

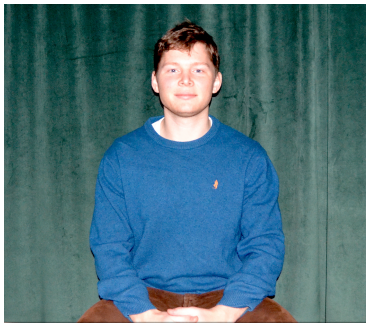
- DNN based Sound Zones.
  Electronic Systems, 7. semester.
  Magnus Munk Jensen, Ujjwal Bagauria, Mikel Jauregui.

- Deep estimation of mutual information for Gaussian signals.
  Mathematical Engineering, 7th semester.
  Anders Højbak Lysgaard Lauridsen, Jakob Olsen, Jacob Mørk, Isabella Quillo, Martin Møller Sørensen.

- Partial Information Decomposition.
  Mathematical Engineering, 7th Semester.
  Jakob Kaltoft Søndergaard, Nicolai Peder Bülow Pedersen, Niklas Keuer Frandsen, Kevin Graversgaard Jepsen.

- A graph theoretical approach to listening effort from EEG signals.
  Mathematical Engineering, long thesis project with Eriksholm.
  Frederik Appel Vardinghus-Nielsen, Magnus Berg Ladefoged, Alexander Djupnes Fuglkjær

PhD project: Self-supervised learning for Speech Source Detection



PhD student: **Holger Severin Bovbjerg**
Start date: **November 1st, 2022.**
Supervisor: **Prof. Zheng-Hua Tan**
Co-supervisors: **Prof. Jan Østergaard and Prof. Jesper Jensen**

In the future, advanced machine learning systems are expected to be able to assist humans in day-to-day life and even extend human ability. Hearing assistive devices such as hearing aids or headphones are examples of technology where machine learning systems can be implemented in order to improve the quality of life for people using this technology, e.g., to do speech enhancement. While current machine learning systems rely heavily on supervised learning methods, this way of learning is limited by the need for labelled data to use for supervision.

Recently, the emerging self-supervised learning paradigm has proved to be able to learn rich representations of data in multiple modalities, including speech, without the need for labels. Self-supervised learning presents a learning paradigm more reminiscent of how humans learn, i.e., through observation of data. The model is first trained to solve a so-called pretext task in which the model learns to capture important information of the data.

As depicted in Figure 1, the model is first trained to solve a so-called pretext task in which the model learns to capture important information of the data. Here, the model predicts labels which are derived from the input data itself without the use of human-annotated labels. An example of a self-supervised pretext task is to predict future speech frames from current ones. After solving this pretext task, the model will have learned to extract information from the data needed to solve the pretext task. The representations learned by the model through solving the self-supervised pretext task can then be used as input to a smaller model, such as a linear classifier, in order to solve a downstream task. Representations learned through self-supervised learning methods are not only less reliant on labelled data but have also been shown to be more general and robust than supervised representations.

This project aims at identifying how self-supervised learning can be used to improve upon two existing technologies used in HA devices known as Voice Activity Detection (VAD) and Direction of Arrival (DoA) estimation. The former is useful for gating the input for a speech processing system such that only speech is sent through, and the latter is used to determine the direction of a sound source, e.g., in order to focus the microphones in the direction of the speaker. Both technologies are important components in improving the quality of speech for HA devices. The main contribution of this project will be an investigation of the use of self-supervised learning for VAD and DoA estimation for improved performance in diverse real-world conditions with background noise.

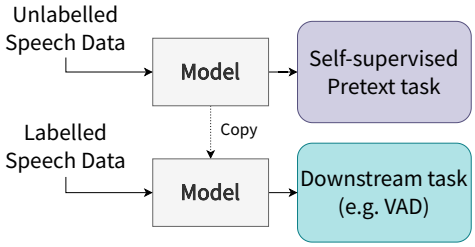


Figure 1 Example of a Self-supervised Learning framework

News

Prof. *Zheng-Hua Tan* had a research stay at Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT, Cambridge, USA, to collaborate on machine learning and speech processing. He gave a talk at MIT on Contrastive Learning for Speech and Multimodal Data Processing.

Zheng-Hua Tan attended Northern Lights Deep Learning Conference 2023 in Tromsø, gave a tutorial on Self-Supervised Learning and participated in a panel.

Zheng-Hua Tan organized a workshop on Self-Supervised Learning for Signal Decoding sponsored by the Pioneer Centre for AI.

Current and former CASPR members received the 2022 IEEE Signal Processing Society Best Paper Award for their paper: "Multitalker speech separation with utterance-level permutation invariant training of deep recurrent neural networks", *IEEE/ACM Trans. Audio, Speech, and Language Processing*, October 2017.

The paper is jointly authored by Morten Kolbæk, Dong Yu, *Zheng-Hua Tan*, and *Jesper Jensen*. For this award, papers in a 5 year window spanning from 2016 - 2021 are considered. The award honors the author(s) of a paper of exceptional merit dealing with a subject related to the Society's technical scope.

Holger Severin Bovbjerg started as a PhD student with CASPR on November 1st, 2022 on the topic of Self-supervised learning for Speech Source Detection.

PhD researcher *Amin Edraki* from Queen's University, Ontario, Canada is visiting CASPR from September to December, 2023. Amin's research is in the area of speech intelligibility prediction and enhancement.

Iván López-Espejo, *Zheng-Hua Tan*, and John H. L. Hansen gave a tutorial on "Deep Spoken Keyword Spotting" at INTERSPEECH 2022, Incheon, Korea, September 2022.

Industrial postdoc researcher *Kateřina Žmolíková* started her industrial postdoc project with CASPR in September, 2022. The project focuses on "brain inspired" signal processing for speech enhancement.

CASPR helped organizing an Open Lab Day within the Section for AI & Sound for Engineering students on November 2nd, 2022. About 30 students enjoyed a guided tour in the labs and experienced several demonstrations including Immersive concerts using Meta Quest 2 headsets, audio visional speech enhancement using deep learning, concurrent EEG and eye tracking, sound zones, and spatial decomposition based sound field construction in an anechoic chamber.

The students got a chance to meet researchers and other students working in the area of AI and sound, and they were getting inspiration for future master projects and/or PhD theses within the field of AI and sound.

Recent CASPR Related Research

Journal papers

1. Utilization of acoustic signals with generative Gaussian and autoencoder modeling for condition-based maintenance of injection moulds. Rønsch, G. Ø., Espejo, I. L., Michelsanti, D., Xie, Y., Popovski, P. & Tan, Z-H., (Accepted/In press) In: International Journal of Computer Integrated Manufacturing.

2. Spectro-temporal modulation glimpsing for speech intelligibility prediction. A. Edraki, W.-Y. Chan, J. Jensen, and D. Fogerty. Hearing Research, vol. 426, Dec. 2022.

3. Incremental Refinements and Multiple Descriptions with Feedback. J. Østergaard, U. Erez, and R. Zamir. IEEE Transactions on Information Theory, vol. 68, no. 10, pp. 6915-6940, Oct. 2022

4. Advanced Dropout: A Model-free Methodology for Bayesian Dropout Optimization. J. Xie, Z. Ma, G. Zhang, J.-H. Xue, Z.-H. Tan and J. Guo. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 9, pp. 4605-4625, Sept. 2022.

5. Speech to noise ratio improvement induces nonlinear parietal phase synchrony in hearing aid users. P.S. Baboukani, C. Graversen, E. Alickovic, J. Østergaard. Frontiers in Neuroscience, vol. 16, August, 2022.

Conference Papers

1. Multiple Description Audio Coding for Wireless Low-Frequency Sound Zones. J. Østergaard, C.S. Pedersen, M. Zhou, N. de Koeijer, M. Møller. Accepted for IEEE Data Compression Conference, March 2023.

2. An Experimental Study on Light Speech Features for Small-Footprint Keyword Spotting. I. López-Espejo, Z.-H. Tan, J. Jensen. Proc. IberSpeech, 2022.

3. Error Reduced Carry Prediction in Approximate Addition for Low-Frequency Sound Zones. P. Koch, J. Østergaard. 18th Biennial Baltic Electronics Conference, BEC October 2022.

4. A linear MMSE filter using delayed remote microphone signals for speech enhancement in hearing aid applications. V. Sathyapriyan, M.S. Pedersen, J. Østergaard, M. Brookes, P. Naylor, J. Jensen. International Workshop on Acoustic Signal Enhancement (IWAENC), September 2022.

5. Distributed Cross-Relation-Based Frequency-Domain Blind System Identification using Online-ADMM. M. Blochberger, F. Elvander, R. Ali, J. Østergaard, M. Moonen, J. Jensen, T. Waterschoot. International Workshop on Acoustic Signal Enhancement (IWAENC), September 2022.

6. Model-Based Estimation of-in-Car-Communication Feedback Applied to Speech Zone Detection. K. Muller, S. Doclo, J. Østergaard, T. Wolf. International Workshop on Acoustic Signal Enhancement (IWAENC), September 2022.

7. Performance of Low Frequency Sound Zones Based on Truncated Room Impulse Responses. J. Cadavid , M. Møller, S. Bech, T. Waterschoot, J. Østergaard. Proceedings of the 17th International Audio Mostly Conference, September 2022.

Contact CASPR

If you are interested in learning more about the research and teaching taking place in CASPR:

Check our webpage at: http://caspr.es.aau.dk

Reach out to Professor Jan Østergaard (jo@es.aau.dk), Professor Zheng-Hua Tan (zt@es.aau.dk), or Professor Jesper Jensen (jje@es.aau.dk).

Aalborg University (http://www.en.aau.dk/) is one of the leading Danish universities with campuses in Aalborg, Esbjerg and Copenhagen. The student population of AAU comprises of around 20.000 regular (both undergraduate and postgraduate) students and 900 PhD students. AAU is famous for its innovative problem and project based learning approach (PBL) where students work on team-based projects solving 'real-life' problems in collaboration with organisations or companies. Aalborg University is acknowledged for collaboration with industry and according to U.S. News & World Report, Aalborg University is the best Engineering University in Europe and the fourth best worldwide.