AALBORG UNIVERSITY
DENMARK

NEWSLETTER

Centre for Acoustic Signal Processing Research (CASPR)

August 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.

CASPR
Centre for Acoustic Signal Processing Research

Contents

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

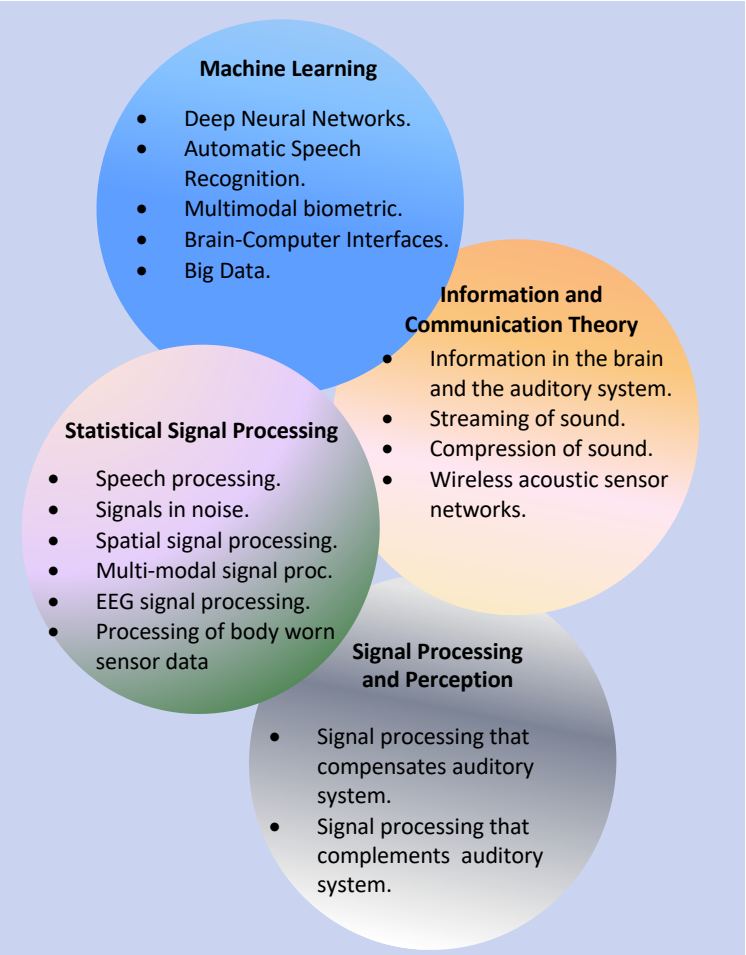
Research in CASPR

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 Spring 2023, 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.

- Automatic Speech Recognition-Driven Speech Enhancement.
  Signal Processing and Acoustics, Master thesis project.
  Michail Kampitakis.

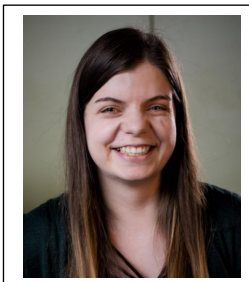
- Speech Enhancement and Deep Learning Speaker Separation.
  Signal Processing and Acoustics, Master thesis project with Oticon.
  Jonas Busk, Morten Andersen and Rasmus Nielsen.

- Dependency Analysis of Electroencephalography Signals. Mathematical Engineering, long thesis project with Eriksholm, Oticon A/S.
  Frederik Appel Vardinghus-Nielsen, Magnus Berg Ladefoged, Alexander Djupnes Fuglkjær

- COINing a New Sound: Compressing Audio with Implicit Neural Representations. Computer Engineering – AI, Vision, and Sound, 8th semester project.
  Christos Kantas, Rasmus Munksø, Mathias Viborg Andersen, Bjørk Antoniussen, Apostolos Ntelopoulos, Shobhit Kotnala.

Research in Focus

Industrial Postdoc project: Self-supervised deep-learning based speech enhancement using interpolation of missing time-frequency features.

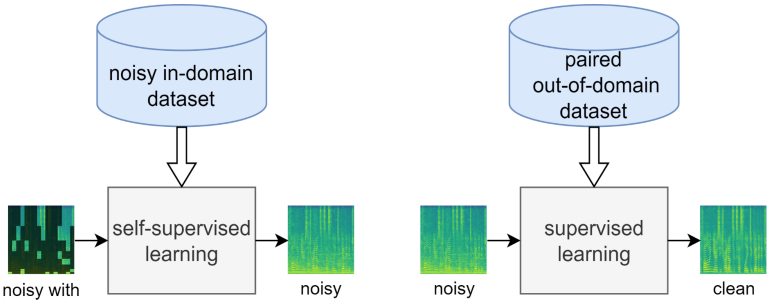


Postdoc researcher: Kateřina Žmolíková
Start date: Aug. 1, 2022.
Supervisor: Prof. Jesper Jensen.
Co-supervisor: Dr. Michael Syskind Pedersen, Oticon A/S.

Speech enhancement is dominated by methods based on deep neural networks (DNNs), which allow for significant improvements in speech intelligibility and quality compared to traditional methods. This remarkable performance has motivated numerous practical applications and commercial products (such as use in hearing aids). However, the performance of DNN based speech processing algorithms in real-world settings, in contrast to controlled laboratory settings, is often somewhat disappointing. This is largely due to discrepancies between the sound data used to train the neural network and the real-world sounds encountered when it is applied. Unseen talkers, noise types and reverberation patterns, as well as mismatches in microphones, etc., are a few examples of such discrepancies.

A straightforward way to overcome this issue, is to train the neural network using data collected in the target domain for the given application (e.g. data collected from hearing aids worn by hearing-aid users in everyday situations). However, in order to train the neural network for speech enhancement, pairs of noisy and corresponding clean recordings are required. Such paired data is extremely difficult to record in the real-world setting. As a result, neural network training methods simulate training data to resemble the target domain, which seldomly captures all of its properties. While paired data are difficult to acquire, target-domain data of noisy mixtures (without corresponding clean ones) is much easier to get. In our project, we aim to find ways to train or adapt DNNs using only the target-domain noisy mixtures.

To train neural networks using target-domain noisy speech signals without corresponding clean speech signals, we aim to make use of self-supervised learning approaches. These approaches have been successfully applied to various tasks in language, image and speech processing, but have not yet been explored for unsupervised domain adaptation in speech enhancement. We hypothesize that self-supervised methods that use interpolation of missing features, e.g., time-frequency regions of spectrograms, can learn salient features and underlying characteristics of noisy mixtures, so that pre-training a model with such self-supervised objective and then transferring it to a supervised learning setting, will lead to a better generalization towards the target domain.



News

In May, 2023, *Mathias Bach Pedersen* successfully defended his phd thesis entitled “Data-Driven Speech Intelligibility Prediction”.

From left: Prof. Fei Chen, Prof. Steven van de Par, Supervisor Dr. Asger Heidemann Andersen, Supervisor Prof. Jesper Jensen, Dr. Mathias Bach Pedersen, Dr. Christian Sejer Pedersen, Supervisor Prof. Zheng-Hua Tan, Prof. Dorte Hammershøi.

In March, 2023, *Poul Hoang* successfully defended his phd thesis entitled “User-Symbiotic Speech Enhancement for Hearing Aids”.

From left: Supervisor Prof. Jesper Jensen, Prof. Søren Bech, Prof. Sharon Gannot, Dr. Poul Hoang, Prof. Reinhold Haeb-Umbach, Supervisor Prof. Zheng-Hua Tan, Supervisor Dr. Jan Mark de Haan.

In June, 2023, the following M.Sc. students supervised by CASPR members graduated from AAU:

Frederik Appel Vardinghus-Nielsen, Magnus Berg Ladefoged, Alexander Djupnes Fuglkjær, Mads Arnløv Jørgensen, Magnus Jónhardsson, and Andreas Larsen, Georgely Kiss, Michail Kampitakis, Jonas Busk, Morten Andersen, and Rasmus Nielsen.

In March, 2023, Payam Shahsavari Baboukani successfully defended his phd thesis entitled “From Global to Local Functional Connectivity: Application to Listening Effort”.

From left: Prof. Maria Chait, Supervisor Dr. Emina Alickovic, Supervisor Assist. Prof. Carina Graversen, Dr. Payam Shahsavari Baboukani, Supervisor Prof. Jan Østergaard, Assoc. Prof. Carles Navarro Manchon, and Prof. Preben Kidmose.

Phd researcher Philippe Gonzales, Technical University of Denmark, visited CASPR from February to June, 2023. Philippe’s research is in the area of generative deep learning models for speech enhancement.

News

Morten Kolbæk, Dong Yu, *Zheng-Hua Tan*, and *Jesper Jensen* 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, Oct. 2017.

From the award ceremony at the IEEE ICASSP conference in Rhodes Island, June 2023: *Zheng-Hua Tan* (left) and *Jesper Jensen* (right) together with the chair for the IEEE Signal Processing Society awards committee, Sergios Theodoridis.

Prof. Geoffrey Chan, Queens University, Canada and Prof. *Jesper Jensen* (CASPR) have received a grant from the William Demant Foundation supporting a project entitled Speaking Style Modifications for Hearing Aid Systems. The project encompasses a phd project for exploring algorithms for speech modifications to enhance speech intelligibility and reduce listening effort, and a phd project for exploring algorithms to predict the perceptual impact of such speech modifications. The project will start during Fall 2023.


Wai-Yip Geoffrey Chan


Jesper Jensen

Recent CASPR Related Research

Journal Papers

1. Data-Driven Non-Intrusive Speech Intelligibility Prediction using Speech Presence Probability. M. B. Pedersen, Z.-H. Tan, S. H. Jensen, and J. Jensen. IEEE Trans. Audio, Speech, Language Process., Accepted. 2023.

2. ACTUAL: Audio Captioning with Caption Feature Space Regularization. Y. Zhang, H. Yu, R. Du, Z.-H. Tan, W. Wang, Z. Ma, and Y. Dong. IEEE/ACM Transactions on Audio, Speech and Language Processing, accepted. 2023.

3. On Training Targets and Activation Functions for Deep Representation Learning in Text-Dependent Speaker Verification. A. K. Sarkar and Z.-H. Tan. Acoustics, accepted. 2023.

4. Cortical Auditory Attention Decoding During Music and Speech Listening. Simon, A., Loquet, G., Østergaard, J. & Bech, S. Accepted for publication in IEEE Transactions on Neural Systems and Rehabilitation Engineering, June 2023.

5. Phase-locking of neural activity to the envelope of speech in the delta frequency band reflects differences between word lists and sentences. Y. Mohammadi, C. Graversen, J. Østergaard, O. K. Andersen and T. Reichenbach, Accepted for publication in Journal of Cognitive Neuroscience, 2023.

6. Leveraging Domain Features for Detecting Adversarial Attacks Against Deep Speech Recognition in Noise. C. Heider and Z.-H. Tan. IEEE Open Journal of Signal Processing, vol. 4, pp. 179-187, 2023.

7. Minimum Processing Near-End Listening Enhancement. A. J. Fuglsig, J. Jensen, Z.-H. Tan, L. S. Bertelsen, J. C. Lindof, J. Østergaard. IEEE Trans. Audio, Speech, Language Process., Vol. 31, June 2023.

8. The Internet of Sounds: Convergent Trends, Insights and Future Directions. L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Backstrom. and C. Fischione. IEEE Internet of Things Journal, Vol. 10., Issue 13, July 2023.

9. On the Deficiency of Intelligibility Metrics as Proxies for Subjective Intelligibility. I. López-Espejo, A. Edraki, W.-Y. Chan, Z.-H. Tan, and J. Jensen. Elsevier Speech Communication, 2023.

Conference Papers

1. Correlation based glimpse proportion index. A. Alghamdi, L. Moen, W.-Y. Chan, D. Fogerty, and J. Jensen. Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2024. Accepted.
2. Adaptive Coding in Wireless Acoustic Sensor Networks for Distributed Blind System. M. Blochberger, F. Elvander, R. Ali, J. Østergaard, J. Jensen, M. Moonen and T. van Waterschoot. Proc. 57th Asilomar Conference on Signals, Systems, and Computers, 2023 (accepted). 2023.
3. Binaural Speech Enhancement using Complex Convolutional Recurrent Networks. V. Tokala, E. Grinstein, M. Brookes, S. Doclo, J. Jensen, P. A. Naylor. Proc. 57th Asilomar Conference on Signals, Systems, and Computers, 2023 (accepted). 2023.
4. Deep Joint Source-Channel Analog Coding for Low-Latency Speech Transmission over Gaussian Channels. M. Bokaei, J. Jensen, S. Doclo, J. Østergaard. Proc. European Signal Processing Conference (EUSIPCO), 2023.
5. Improved Disentangled Speech Representations Using Contrastive Learning in Factorized Hierarchical Variational Autoencoder. Y. Xie, T. Arildsen, and Z.-H. Tan. Proc. European Signal Processing Conference (EUSIPCO), 2023.
6. Speech Inpainting: Context-Based Speech Synthesis Guided by Video. J.F. Montesinos, D. Michelsanti, G. Haro, Z.-H. Tan, and J. Jensen. Interspeech 2023.
7. AR model for low latency packet loss concealment for wireless sound zones at low frequencies. C. S. Pedersen, M. Zhou, M. B. Møller, N. de Koeijer, and J. Østergaard, May 2023, *Audio Engineering Society 154th Convention.* Finland.
8. Distributed Adaptive Norm Estimation for Blind System Identification in Wireless Sensor Networks. M. Blochberger, F. Elvander, R. Ali, J. Østergaard, J. Jensen, M. Moonen and T. van Waterschoot, Proc. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* June 2023.
9. Interpretable Nonnegative Incoherent Deep Dictionary Learning for FMRI Data Analysis. M. Morante, J. Østergaard and S. Theodoridis, Proc. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* June 2023.
10. Filterbank Learning for Noise-Robust Small-Footprint Keyword Spotting. I. López-Espejo, R. C. M. C. Shekar, Z.-H. Tan, J. Jensen and J. H. L. Hansen. Proc. *ICASSP,* June 2023.
11. Robust FIR Filters for Wireless Low-Frequency Sound Zones. M. Zhou, M. Møller, C. S. Pedersen and J. Østergaard, Proc. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* June 2023.
12. Multiple Description Audio Coding for Wireless Low-Frequency Sound Zones. J. Østergaard, C. S. Pedersen, M. Zhou, N. de Koeijer and M. Møller, *IEEE Data Compression Conference.* IEEE Signal Processing Society, March 2023.
13. A Vision-Assisted Hearing Aid System Based on Deep Learning. D. Michelsanti, Z.-H. Tan, S.R. Griful and J. Jensen. IEEE ICASSP 2023 Satellite Workshop: AMHAT.
14. Improving Label-deficient Keyword Spotting Through Self-supervised Pretraining. H.S. Bovbjerg and Z.-H. Tan. ICASSP 2023 Satellite Workshop: SASB 2023.

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.