



AALBORG UNIVERSITY
DENMARK

NEWSLETTER

Centre for Acoustic Signal Processing Research
(CASPR)

August 2024

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

1 CASPR

2-3 Research & Teaching

4 CASPR Research in Focus

5 News

7 Publications

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

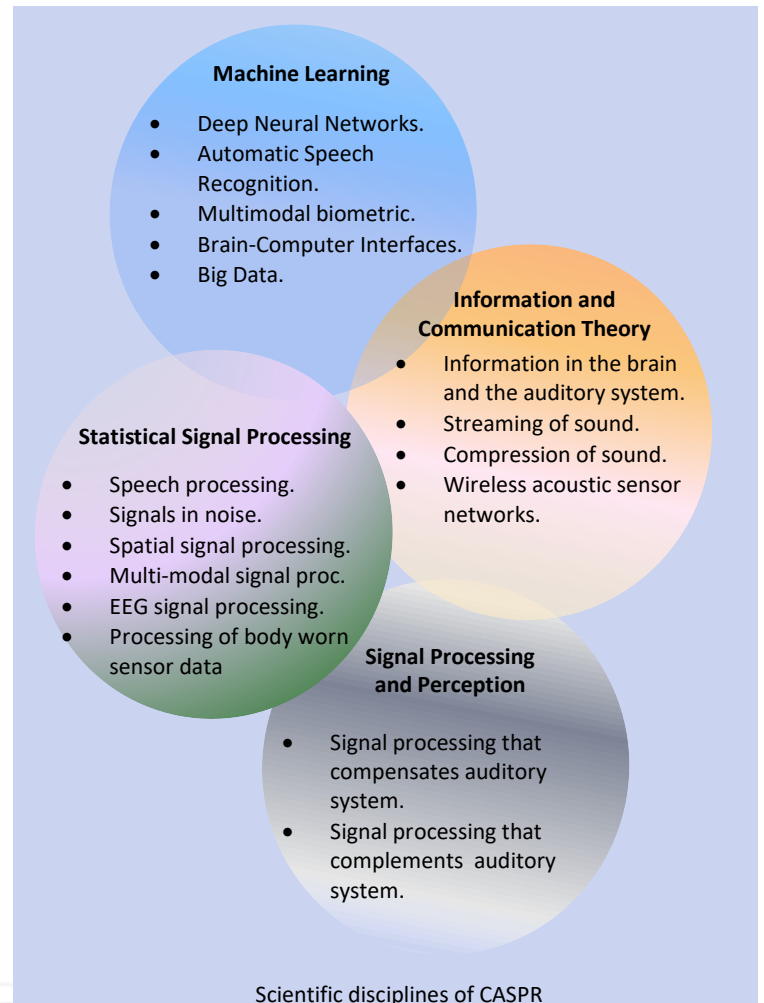
Research in CASPR

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.



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:

- Machine learning, Master and PhD courses.
- Deep learning, Master and PhD courses.
- Self-Supervised Learning, PhD course.
- CASPR Seasonal Schools on Signal Processing for Hearing Assistive Devices.
- Information theory, Master course.

During Spring 2024, CASPR was involved in the following student projects

- DNN based robust compression and communication of speech
Mathematical Engineering, 8th semester project.
Sif B. Lindby, Mads L. Larsen
- VQ-based generative DNN models for low delay speech and audio coding
Computer Engineering, AI, Vision, and Sound.
Lukas Bisgaard
- Approximate computing for resource-constrained DNNs
Electronic Systems, Master thesis project
Simon D. Jepsen, Albert B. Hansen
- Evaluating Multi-Window Multi-Head Attention for Automatic Speech Recognition.
Mathematical Engineering, 8th semester project.
Nikolai Lund Kühne.
- Fourier approximation for local-global Multi-Head Attention.
Computer Engineering – AI, Vision and Sound, 8th semester project.
Mohamad Dalal, Susanne Miranda Aalestrup, Valentina Correa Bejarano, and Vasiliki Ismiroglou.
- Fourier approximation for local-global Multi-Head Attention.
Mathematical Engineering, 8th semester project.
Jakob Olavi Grangaard Olesen.
- A Multimodal Large Language Model for Music Captioning.
Mathematical Engineering, Master thesis project with Bang & Olufsen.
Jakob Olsen, Anders Højbak Lysgaard Lauridsen, Jacob Mørk.
- Indoor Radar Data Analysis for Estimating Room Geometry.
Computer Engineering – AI, Vision and Sound, Master thesis project with Bang & Olufsen.
Máté Tallósi
- Robust Speaker Embeddings through Speech Synthesis.
Computer Engineering – AI, Vision and Sound, 8th semester project.
Tonko Emil W. Bossen, Sameer Aqib Hashmi, Yu-ling Cheng, Marko Putak.

CASPR Winter School

CASPR held a week-long Winter School on Signal Processing for Hearing Assistive Devices (HADs) at the AAU Campus in Copenhagen, Denmark, during 26/2 – 1/3, 2024. The first part of the Winter School gave an overview of fundamental signal processing problems encountered in HADs, and in-depth treatment of state-of-the-art solutions. The second part covered emerging technologies. The participants were from industry and university and worked together in teams to build real-world sound processing algorithms. There were also plenty of opportunities for networking with other participants and experts in sound, signal processing, and machine learning.

We thank the participants for their impressive engagement in the lectures, question rounds, and exercises.



Participants (and some organizers) in CASPR Winter School in Copenhagen, Feb. 2024

We thank the lecturers for their great contributions which made the Winter School exciting and relevant. In particular, thanks to:

- Prof. Steven van de Par, Oldenburg University.
- Dr. Meng Guo, Oticon.
- Dr. Robert Rehr, Oticon.
- Dr. Michael Syskind Pedersen, Oticon.
- Dr. Dorothea Wendt, Eriksholm Research Center.
- PhD student Mohammad Bokaei, Aalborg University
- PhD student Mohammad Bokaei, Aalborg University
- Postdoc Payam Shahsavari, Aalborg University
- PhD student Vasudha Sathyapriyan, Aalborg University and Oticon
- PhD student Andreas Fuglsig, Aalborg University
- PhD student Peter Leer Bysted, Aalborg University
- PhD student Philippe Gonzalez, Technical University of Denmark
- PhD student Asjid Tanveer, Aalborg University
- PhD student Sangeeth G. Jayaprakash, Aalborg University
- PhD student Holger S. Bovbjerg, Aalborg University
- Prof. Jesper Jensen, Aalborg University and Oticon.
- Prof. Zheng-Hua Tan, Aalborg University.
- Prof. Jan Østergaard, Aalborg University.

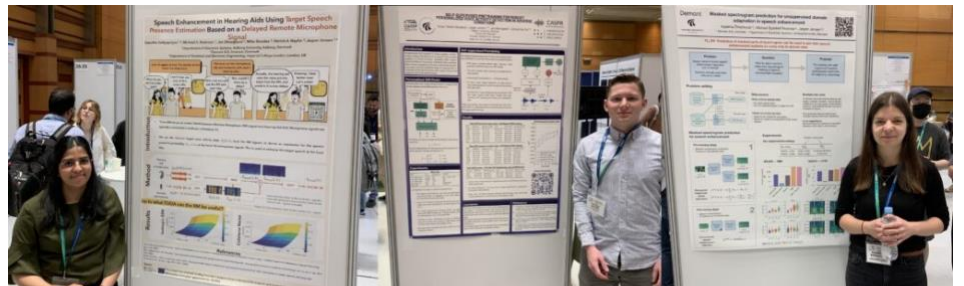
News

On Friday 14th of June, PhD student Holger Severin Bovbjerg represented CASPR at Folkemødet which takes place in Allinge on the Danish Island Bornholm. 50 researchers from Danish universities, including Danish Nobel Prize winner Morten Meldal, were invited to speak at Folkemødets Forskningscene (Folkemødet's Research Scene), to talk about current research. Holger was invited to speak about personalized AI speech models and in his talk, he explained how AI speech models work in layperson's terms. He also presented examples of some of their interesting applications, such as personalized hearing aids and early detection of various illnesses.



CASPR PhD student Holger Severin Bovbjerg, has been granted the prestigious Fulbright scholarship to support his coming research stay in the U.S. at Carnegie Mellon University in Pittsburgh. Fulbright is an international academic exchange program sponsored by the U.S. government to promote scientific and cultural exchange between the U.S. and other countries. On June 20th, Holger visited "Rydhave", the residence of the United States' Ambassador to Denmark, where he was invited to attend a reception for Danish Fulbright grantees. During his stay in the U.S. Holger will serve as a Fulbright ambassador for Denmark.

Vasudha Sathyapriyan, Holger S. Bovbjerg, and Kateřina Žmolíková before their poster sessions at ICASSP 2024.



Zheng-Hua Tan as Technical Program Vice Chair of ICASSP 2024 at the closing ceremony.

News



PhD student José Miguel Cadavid Tobón received the best student paper award at the 156th AES Convention in Madrid for his paper: “Spatial Sampling versus Acquisition Time of Room Impulse Responses for Low-Frequency Sound Zones”.

The co-authors are Martin Møller, Søren Bech, Toon van Waterschoot, and Jan Østergaard.

Along with the paper, a dataset of low-frequency RIRs for sound field control has been released. It is openly available with CC BY 4.0 license at:

<https://lnkd.in/dVU3Hi5n>

PhD student José Miguel Cadavid Tobón held a workshop titled “Basics of Sound Zones Design and Evaluation” at the 156th AES Convention. The workshop included both a theoretical and a practical session. The slides and codes can be found in the following Drive folder: <https://lnkd.in/dan9HHqK>

The following two courses on machine learning and self-supervised learning will be available in 2024 - 2025. Participation is free of charge for PhD students and costs DKK 2000,- per ECTS for industry participants. The courses are organized and lectured by Prof. Zheng-Hua Tan. The courses will take place physically at Aalborg University, Department of Electronic Systems.

Course 1: Machine Learning

ECTS: 3.0

Date/Time: 17, 19, 21, 24, 26 March 2025

Description: This course will give a comprehensive introduction to machine learning both by presenting technologies proven valuable and by addressing specific problems such as pattern recognition, prediction, clustering, generative modeling and anomaly detection. This course covers both theory and practices for machine learning, but with an emphasis on the practical side namely how to effectively apply machine learning to a variety of problems. Topics will include: Supervised learning methods: logistic regression, support vector machines, neural networks, K-nearest neighbors, decision trees, boosting. Unsupervised learning and clustering methods: K-means, Gaussian mixture models, Expectation Maximization algorithm, principal component analysis. Deep learning methods: deep neural networks, long short-term memory recurrent neural networks, convolutional neural networks, generative adversarial networks. Probabilistic graphical models. Reinforcement learning.

Prerequisites: Basic probability and statistics theory, linear algebra and basic programming skills.

Link: <https://phd.moodle.aau.dk/course/index.php?categoryid=308>

Course 2: Self supervised learning

ECTS: 2.0

Date/Time: November 18-20, 2024

Description: The course introduces self-supervised learning methods for learning representations of single- and multiple-modality data, covering deep architectures, training target and loss functions used in state-of-the-art methods, and selected downstream applications. A focus will be given to loss functions including both contrastive and predictive losses.

Prerequisites: Knowledge in machine learning or deep learning and basic skills in Python programming

Link: <https://phd.moodle.aau.dk/course/view.php?id=2373>

Publications

Conference Papers

1. Near-End Listening Enhancement Using a Noise-Robust Linear Time-Invariant Filter. F. Villani, W.-Y. Chan, Z.-H. Tan, J. Østergaard, J. Jensen., Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), September 9 – 12, 2024.
2. Bayesian Sound Field Estimation Using Uncertain Data. J. Brunnstrom, M. B. Møller, J. Østergaard, and M. Moonen. Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), September 9 – 12, 2024.
3. Audio Mamba: Selective State Spaces for Self-Supervised Audio Representations. S. Yadav and Z.-H. Tan, Interspeech 2024, Kos Island, Greece, September 1-5, 2024.
4. No-Reference Speech Intelligibility Prediction Leveraging a Noisy-Speech ASR Pre-Trained Model. H. Wang, J. Jensen, I. Lopez-Espejo, W.-Y. Chan, Proc. Interspeech, 2024.
5. Deep Digital Joint Source-Channel Based Wireless Speech Transmission. M. Bokaei, J. Jensen, S. Doclo, J. Østergaard. Proc. 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.
6. Speaker and Style Disentanglement of Speech Based on Contrastive Predictive Coding Supported Factorized Variational Autoencoder. Y. Xie, M. Kuhlmann, F. Rautenberg, Z.-H. Tan, and R. Haeb-Umbach, The 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.
7. Envelope Based Deep Source Separation and EEG Auditory Decoding for Speech and Music, M. A. Tanveer, J. Jensen, Z.-H. Tan, J. Østergaard, Proc. 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.
8. PAC-Bayesian Error Bound, via Rényi Divergence, for a Class of Linear Time-Invariant State-Space Models. D. Eringis, J. Leth, Z.-H. Tan, R. Wisniewski, and M. Petreczky, The 41st International Conference on Machine Learning (ICML 224), Vienna, Austria, July 21-27, 2024.
9. Synergy and Redundancy Dominated Effects in Time Series via Transfer Entropy Decompositions. J. Østergaard and P.S. Boubakani. IEEE International Symposium on Information Theory Workshops, July, 2024.
10. Directed Redundancy in Time Series. J. Østergaard. IEEE International Symposium on Information Theory, July, 2024.
11. Complex Recurrent Variational Autoencoder for Speech Resynthesis and Enhancement. Y. Xie, T. Arildsen, and Z.-H. Tan, IEEE World Congress on Computational Intelligence (IEEE WCCI 2024), Yokohama, Japan, June 30-July 5, 2024.
12. Masked Autoencoders with Multi-Window Local-Global Attention Are Better Audio Learners. S. Yadav, S. Theodoridis, L. K. Hansen, and Z.-H. Tan, The Twelfth International Conference on Learning Representations (ICLR 2024), Vienna, Austria, May 7-11, 2024.
13. Spatial sampling versus acquisition time of room impulse responses for low-frequency sound zones. J. Cadavid, M. Møller, T. van Waterschoot, S. Bech, and J. Østergaard. Proceedings of the AES International Conference 156, 2024.
14. Masked spectrogram prediction for unsupervised domain adaptation in speech enhancement, K. Žmolíková, M. S. Pedersen, J. Jensen, Proc. ICASSP 2024
15. Self-Supervised Pre-Training for Robust Personalized Voice Activity Detection in Adverse Conditions. H. S. Bovbjerg, J. Jensen, J. Østergaard, Z.-H. Tan. Proc. ICASSP 2024.
16. Diffusion-Based Speech Enhancement in Matched and Mismatched Conditions Using a Heun-Based Sampler. P. Gonzalez, Z.-H. Tan, J. Østergaard, J. Jensen, T. S. Alstrøm, T. May. Proc. ICASSP 2024.
17. Speaker Adaptation for Enhancement of Bone-Conducted Speech. A. Edraki, W.-Y. Chan, J. Jensen, D. Fogerty. Proc. ICASSP 2024.
18. Speech enhancement in hearing aids using target speech presence estimation based on a delayed remote microphone signal. V. Sathyapriyan, M. S. Pedersen, M. Brookes, J. Østergaard, P. A. Naylor, J. Jensen, Proc. ICASSP 2024.
19. Binaural Speech Enhancement using Deep Complex Convolutional Transformer Networks. V. Tokala, E. Grinstein, M. Brookes, S. Doclo, J. Jensen, P. A. Naylor. Proc. ICASSP 2024.
20. Deep Low-Latency Joint Speech Transmission and Enhancement over a Gaussian Channel. M. Bokaei, J. Jensen, S. Doclo, J. Østergaard, Proc. IEEE Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA) 2024.
21. Joint Minimum Processing Beamforming and Near-End Listening Enhancement. A. J. Fuglsig, J. Jensen, Z.-H. Tan, L. S. Bertelsen, J. C. Lindof, J. Østergaard, Proc. IEEE Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA) 2024.
22. Channel-Configurable Deep Wireless Speech Transmission. M. Bokaei, J. Jensen, S. Doclo, J. Østergaard. IEEE Wireless Communications and Networking Conference (WCNC), 2024.
23. PAC-Bayes Generalisation Bounds for Dynamical Systems Including Stable RNNs. D. Eringis, J. Leth, Z.-H. Tan, R. Wisniewski, and M. Petreczky. The 38th Annual AAAI Conference on Artificial Intelligence, 2024.

Publications

Journal Papers

1. The Effect of Training Dataset Size on Discriminative and Diffusion-Based Speech Enhancement Systems. P. Gonzalez, Z.-H. Tan, J. Østergaard, J. Jensen, T. S. Alstrøm, and T. May, *IEEE Signal Processing Letters*, 2024.
2. Joint Far- and Near-end Speech and Listening Enhancement with Minimum Processing. A. J. Fuglsig, Z.-H. Tan, L. S. Bertelsen, J. Jensen, J. C. Lindof, and J. Østergaard, *IEEE Access*, 2024.
3. Identifying principal attributes for evaluating audio quality of reproduction systems with spatially dynamic program material. P. N. P. Moreta, S. Bech, J. Francombe, J. Østergaard, and S. van de Par. *Journal of the Audio Engineering Society*, 2024.
4. Data-Driven Non-Intrusive Speech Intelligibility Prediction using Speech Presence Probability. M. B. Pedersen, Z.-H. Tan, S. H. Jensen, and J. Jensen. *IEEE/ACM Trans. Audio, Speech, Lang. Proc.*, Vol. 32, pp. 55-67, 2024.
5. Generating Accurate and Diverse Audio Captions through Variational Autoencoder Framework. Y. Zhang, R. Du, Z.-H. Tan, W. Wang, and Z. Ma, *IEEE Signal Processing Letters*, June, 2024.
6. How to train your ears: Auditory-model emulation for large-dynamic-range inputs and mild-to-severe hearing losses. P. A. L. Bysted, J. Jensen, Z.-H. Tan, J. Østergaard, and L. Bramsløw, *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 32, pp. 2006-2020, 2024.
7. Reduced complexity for sound zones with subband block adaptive filters and a loudspeaker line array. M.B. Møller, J. Martinez, and J. Østergaard. *The Journal of the Acoustical Society of America* 155 (4), 2314-2326, April, 2024.
8. Performance of low-frequency sound zones with very fast room impulse response measurements. J. Cadavid, M. B. Møller, C. S. Pedersen, S. Bech, T. van Waterschoot, and J. Østergaard. *The Journal of the Acoustical Society of America*, vol. 155, January, 2024.

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)

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.