NEWSLETTER

CENTRE FOR ACOUSTIC SIGNAL PROCESSING RESEARCH (CASPR)

August 2025

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.

AALBORG UNIVERSITY

If you are interested in learning more about the research and teaching taking place in CASPR:
Check our webpage here.

Reach out to:

- Prof. Jan Østergaard
- Prof. Zheng-Hua Tan
- Prof. Jesper Jensen

Contents

CASPR
Centre for Acoustic Signal Processing Research

RESEARCH IN CASPR

ABOUT 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

Machine Learning

- Deep Neural Networks
- Automatic Speech Recognition
- Multimodal Biometric
- Brain-Computer Interfaces
- Big Data

Information and Communication Theory

- Information in the brain and the auditory system
- Steaming of sound
- Compression of sound
- Wireless acoustic sensor networks

Statistical Signal Processing

- Speech processing
- Signals in noise
- Spatial signal processing
- Multimodal signal processing
- EEG signal processing
- Processing of body worn sensor data

Signal Processing and Perception

- Signal processing that compensates auditory system
- Signal processing that complements auditory system

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.
- Reinforcement Learning, PhD course.
- CASPR Seasonal Schools on Signal Processing for Hearing Assistive Devices.
- Information theory, Master course.

DURING SPRING 2025, CASPR WAS INVOLVED IN THE FOLLOWING STUDENT PROJECTS

- Project title: “Quantifying speech comprehension using Large Language Models”
  Student: Sif B. Lindby
  Education: Mathematical Engineering - joint 9+10th semesters long-thesis and 4+4 PhD
  Uni Supervisors: Profs. Jan Østergaard, Zheng-Hua Tan, Jesper Jensen

- Project title: “Speech Enhancement using Sequence Modelling Neural Architectures and Spoken Large Language Models”
  Student: Nikolai Lund Kühne
  Education: Mathematical Engineering - joint 9+10th semesters long-thesis and 4+4 PhD
  Uni Supervisors: Profs. Zheng-Hua Tan, Jan Østergaard, Jesper Jensen

- Project title: “Adaptive Acoustic Noise Cancellation”
  Student: Jacob El-Omar
  Education: Electronic Systems, 9-10 semesters long thesis project.
  Uni Supervisor: Prof. Jan Østergaard

- Project title: “Data-driven self-supervised machine learning using sound data”
  Student: Rune Drongesen
  Education: Electronic Systems, 10th semester
  Uni Supervisors: Prof. Zheng-Hua Tan, Kevin Wilkinghoff
  Company Supervisors: Nicolai Bæk Thomsen, Grundfos

TEACHING IN CASPR

Continued

- Project title: “On the Performance of Sound Bubbles: An Analysis of Directional Performance and Underlying Acoustic Principles”
  Student: Emil Lunau Bentsen
  Education: Electronic Systems, 10th semester
  Uni Supervisor: Prof. Zheng-Hua Tan
  Company Supervisors: Svend Feldt, Jesper Jensen

- Project title: “Study and Test of a Deep Learning-Based Sound Bubble System”
  Students: Magnus Ø. B. Poulsen, Fillip B. Brink, Casper T. Kjeldsen, Nicolai B. Jespersen
  Education: Mathematical Engineering, 8th semester
  Uni Supervisor: Prof. Zheng-Hua Tan
  Company Supervisors: Svend Feldt, Jesper Jensen

- Project title: “Deep Target Conversation Identification for Hearing Aids”
  Students: Jeppe Roden Münster, Andreas Peter Juhl Hansen, Simon Bock Segaard
  Education: Mathematical Engineering, 8th semester
  Uni Supervisor: Prof. Zheng-Hua Tan
  Company Supervisors: Poul Hoang, Jesper Jensen

- Project title: “Single-channel Speech Enhancement with Layer-wise Mixture of Experts”
  Student: Zexu Han
  Education: Vision and Sound, 8th semester
  Uni Supervisor: Prof. Zheng-Hua Tan
  Company Supervisors: Peter Leer, Jesper Jensen

CASPR RESEARCH IN FOCUS

PHD PROJECT: SPEECH ENHANCEMENT USING SEQUENCE MODELLING NEURAL ARCHITECTURES AND AUDIO FOUNDATION MODELS

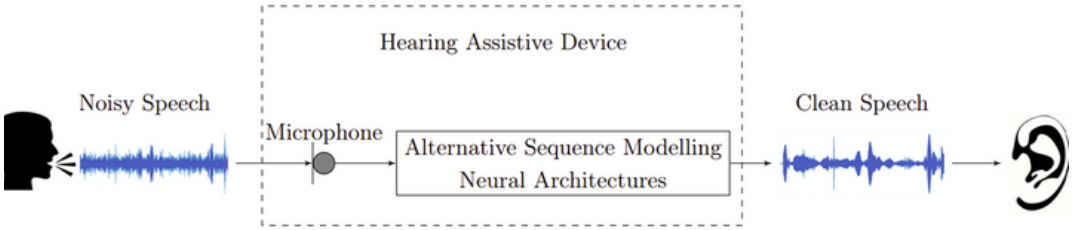
PhD researcher: Nikolai Lund Kühne
Start date: 01.09.2024
Supervisors: Profs. Zheng-Hua Tan, Jan Østergaard, Jesper Jensen

PROJECT SUMMARY

Today, the use of hearing assistive devices for hearing impairment is more widespread than ever. However, one of the biggest issues with these devices is their inadequate performance in noisy conditions. This issue naturally affects the daily lives of hearing-impaired people, driving the development of new and improved algorithms for noise reduction in hearing aids. The task of removing background noise and increasing the quality and intelligibility of the target speech signal is known as speech enhancement.

Recent advances in sequence modelling architectures such as Mamba and extended Long Short-Term Memory (xLSTM) have demonstrated competitive, and sometimes superior performance compared to the current state-of-the-art in various speech-related tasks. Yet, the application of these models in speech enhancement remains underexplored. Moreover, audio foundation models based on Large Language Models (LLMs) have shown promising potential in this area. However, their application in speech enhancement remains significantly underexplored.

This project aims to push the performance of deep learning-based algorithms and tailor them to speech enhancement. This will be carried out by adopting and combining the newest sequence modelling neural architectures for speech enhancement, since we hypothesize that a performance gain can be achieved. Our findings have shown that xLSTM- and even LSTM-based systems can match or outperform state-of-the-art Conformer- and Mamba-based speech enhancement systems. Additionally, our most recent finding revealed that combining Mamba and multi-head attention for speech enhancement yields state-of-the-art performance on out-of-domain datasets. This finding is important, since good generalization performance is paramount for real-world applications in e.g. hearing aids, where speech and noise types may vary significantly from the training data.



NEWS

- Luis Alfredo Moctezuma Pascual started as an industrial postdoc in CASPR and Oticon on March 1, 2025. The research project is on Attention-steered beamformers based on sensor-fusion in multi-talker listening environments. Supervisors are Emina Alickovic (Oticon), Dorothea Wendt (Oticon), Jan Østergaard (CASPR), Svend Feldt (Oticon), and Jesper Jensen (CASPR, Oticon)

- Niels Overby started as an industrial postdoc in CASPR and Oticon on May 1, 2025. The research project is on intent-controlled hearing aid systems with remote microphones. Supervisors are Jan Østergaard (CASPR), Zheng-Hua Tan (CASPR), Svend Feldt (Oticon), and Jesper Jensen (CASPR, Oticon).

- Peter Leer Bysted started as an industrial postdoc in CASPR and Oticon on May 1, 2025. The research project is on self-learning deep neural network based noise reduction for hearing assistive devices. Supervisors are Zheng-Hua Tan (CASPR), Jan Østergaard (CASPR), Svend Feldt (Oticon), and Jesper Jensen (CASPR, Oticon).

- Prof. Zheng-Hua Tan (AAU) and Prof. Lars Kai Hansen (DTU) are organizing the 2nd Workshop on Self-Supervised Learning for Signal Decoding. The event is hosted by the Signals and Decoding Collaboratory of the Pioneer Centre for AI and will take place in Aalborg from August 14-15, 2025. The program features invited talks and posters, covering self-supervised learning methods, large language models, explainable AI and more.

  Invited speakers:
  Jinyu Li, Microsoft
  Sebastian Roland Lapuschkin, Fraunhofer HHI - Heinrich Hertz Institute
  Hung-yi Lee, National Taiwan University
  Sneha Das, Technical University of Denmark
  Lenka Tetková, Technical University of Denmark
  Sarthak Yadav, Aalborg University
  Lars Kai Hansen, Technical University of Denmark
  Zheng-Hua Tan, Aalborg University

NEWS
Continued

- Asjid Tanveer, PhD researcher in CASPR, did a research visit to Oticon during Spring 2025, where he worked on auditory attention decoding for the hearing impaired.

- Nicolaj Boelt Pedersen, a 9th-semester student from Computer Engineering (AI, Vision, and Sound), Aalborg University, will do an internship at Oticon from September 1, 2025. The topic will be hearing assistance with remote microphones.

- Prof. Zheng-Hua Tan will be an invited speaker at the International Symposium on Auditory and Audiological Research (ISAAR) to be held in Nyborg, Denmark, from August 20-22, 2025. His presentation will cover several self-supervised learning (SSL) methods for speech and audio processing. The talk will encompass SSL training techniques, new neural architectures, and a variety of downstream tasks. Zheng-Hua also delivered an invited talk at the Danish Society of Engineers (IDA), Copenhagen, on the topic of Self-Supervised Learning and Foundation Models, in Copenhagen, January 22, 2025.

- Students at AAU supervised by CASPR staff successfully defended their theses in June 2025: Emil Lunau Bentsen, Jacob El-Omar, and Rune Drongesen.

- Pictured on the bottom right: This June, the CASPR PhD student Holger Severin Bovbjerg (left) attended a Fulbright reception hosted by the U.S. Embassy in Charlottenlund, where he received a certificate from Jennifer Hall Godfrey (right), Charge d'Affaires at the U.S. Embassy to the Kingdom of Denmark.

  The certificate acknowledges his participation in the program during his research at Carnegie Mellon University in Pittsburgh, for which he was granted a Fulbright scholarship in 2024.

  The annual reception brings together Danish and U.S. Fulbright students from Fulbright Denmark to celebrate the program and exchange experiences. Since the program began in Denmark in 1951, over 3,500 elite students have participated, with 25-50 grants awarded annually.

  The Fulbright certificate signifies that Bovbjerg now serves as an ambassador for the program's values which include promoting mutual understanding, educational and cultural exchange, and peaceful relations.

NEWS
Continued (ICASSP)

- Prof. Zheng-Hua Tan, as the General Chair, successfully led and won the bid to host the 2029 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) in Copenhagen, scheduled for April 22-27, 2029.

  First held in 1976, ICASSP is the world's largest technical conference dedicated to signal processing, attracting several thousand participants annually. ICASSP 2029 in Copenhagen will mark the first time the conference is held in the Nordic countries. The event will take place at the Bella Center, a premier venue in Denmark.

  Denmark is renowned for its innovation in sound technology and other fields that heavily rely on signal processing. We look forward to welcoming the global signal processing community to Copenhagen for this event and having fruitful discussions and collaborations.

- The following people (left to right) attended ICASSP 2025: Andreas Jonas Fuglsig, Jesper Rindom Jensen, Nikolai Lund Kühne, Jesper Jensen, Eleftheria Lydaki, Marie Saugstrup Jensen, Mikkel Sebastian Lundsgaard Brøndt, and Astrid Holm Filtenborg Kitchen.

PUBLICATIONS

From January 2025 onward

CONFERENCE PAPERS

1. “Rate-Distortion under Neural Tracking of Speech: A Directed Redundancy Approach”, J. Østergaard, S. G. Jayaprakash, R. Ordonez. Accepted for the IEEE Data Compression Conference, March 2025.
2. “Deep Feedback Cancellation for Hearing Aids with Improved System Stability and Sound Quality”, E. Lydika, Z.-H. Tan, J. Jensen, M. Guo. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, April 6-11, 2025.
3. “Detecting and Defending Against Adversarial Attacks on Automatic Speech Recognition via Diffusion Models”, Nikolai L Kühne, Astrid HF Kitchen, Marie S Jensen, Mikkel SL Brøndt, Martin Gonzalez, Christophe Biscio, and Zheng-Hua Tan. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, April 6-11, 2025.
4. “Optimal Sensor Scheduling for Continuous-Discrete Kalman Filtering with Auxiliary Dynamics”, Mohamad Al Ahdab, Zheng-Hua Tan, and John-Josef Leth. The 42nd International Conference on Machine Learning (ICML), Vancouver, Canada, July 13-19, 2025.
5. “Analysis and Extension of a Near-End Listening Enhancement Method Based on Long-Term Fractile Noise Statistics”, Filippo Villani, Wai Yip Geoffrey Chan, Zheng-Hua Tan, Jan Østergaard, and Jesper Jensen. Interspeech, Rotterdam, The Netherlands, 17-21 August 2025.
6. “AxLSTMs: learning self-supervised audio representations with xLSTMs”, Sarthak Yadav, Sergios Theodoridis, and Zheng-Hua Tan. Interspeech, Rotterdam, The Netherlands, 17-21 August 2025.
7. “Intelligibility Prediction for Time-Modified Speech Signals Using Spectro-Temporal Modulation Features”, Aymen Bashir, Haolan Wang, Amin Edraki, Wai-Yip Chan, Jesper Jensen. Interspeech, Rotterdam, The Netherlands, 17-21 August 2025.
8. “xLSTM-SENet: xLSTM for Single-Channel Speech Enhancement”, Nikolai Lund Kühne, Jan Østergaard, Jesper Jensen, and Zheng-Hua Tan. Interspeech, Rotterdam, The Netherlands, 17-21 August 2025.
9. “AudioMAE++: Learning Better Masked Audio Representations with SwiGLU FFNs”, Sarthak Yadav, Sergios Theodoridis and Zheng-Hua Tan. IEEE International Workshop on Machine Learning for Signal Processing (MLSP), August 31 - September 3, 2025, Istanbul, Turkey.
10. "Multivariate Phase Synchrony and Auditory Attention Decoding for Speech and Music", Asjid Tanveer, Jesper Jensen, Zheng-Hua Tan, and Jan Østergaard. ICBEB 2025, September 19-21, 2025, Prague, Czech Republic.
11. “Learning Robust Spatial Representations from Binaural Audio through Feature Distillation”, Holger Severin Bovbjerg, Jan Østergaard, Jesper Jensen, Shini Watanabe, and Zheng-Hua Tan. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2025), October 12-15, 2025, Granlibakken Tahoe, Tahoe City, CA.

JOURNAL PAPERS

1. “Single-microphone Deep Envelope Separation based Auditory Attention Decoding for Competing Speech and Music”, A. Tanveer, J. Jensen, Z.-H. Tan, J. Østergaard. Journal of Neural Engineering, vol 22, 2025.
2. “Effect of Reduced Information in the Performance of Low-Frequency Sound Zones”, J. Cavadid, M. Møller, T. van Waterschoot, S. Bech, J. Østergaard. Journal of the Audio Engineering Society, 2025.
3. “Low-latency Deep Analog Speech Transmission using Joint Source Channel Coding”, Mohammad Bokaei, Jesper Jensen, Simon Doclo, Jan Østergaard, Journal of Selected Topics in Signal Processing, Accepted 2025.
4. “Noise-Robust Hearing Aid Voice Control”, Ivan Lopez-Espejo, Eros Rosello, Amin Edraki, and Jesper Jensen, IEEE Signal Processing Letters, vol. 32, pp. 241-245, 2025
5. “Zero-Shot Audio Captioning Using Soft and Hard Prompts”, Yiming Zhang, Xuenan Xu, Ruoyi Du, Haohe Liu, Yuan Dong, Zheng-Hua Tan, Wenwu Wang, Zhanyu Ma, , IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 33, pp. 2045-2058, 2025.
6. “Hearing-loss Compensation Using Deep Neural Networks: A Framework and Results from a Listening Test”, Peter Asbjørn Leer Bysted, Jesper Jensen, Laurel Carney, Zheng-Hua Tan, Jan Østergaard, and Lars Bramsløw, IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 33, pp. 828-841, 2025.