University of Stuttgart
Germany

TTA for Generative Speech Enhancement

26. May 2026

Thesis	MA
Supervisor	Tobias Raichle
Examiner	Prof. Dr.-Ing. Bin Yang

Motivation

Speech enhancement (SE) studies improving the quality of spoken language and finds applications as a front-end in automatic speech recognition, telecommunication or hearing aids. SE systems have shown remarkable success in controlled environments but often struggle when faced with domain shifts in real-world deployments [1, 2].

Recent advances in test-time adaptation (TTA) for speech denoising have demonstrated promising results by enabling models to adapt on-the-fly to new acoustic environments without requiring labeled data [2, 3]. However, so far they have focused exclusively on discriminative SE models (e.g. CMGAN [1]). The goal of this thesis is to explore the behavior of generative speech enhancement models (e.g. SGMSE [4]) under domain shifts and develop adaptation strategies.

Objectives

- Evaluate generative SE models under domain shifts.
- Explore failure modes of these models.
- Develop a TTA strategy for generative speech enhancement models.

Prerequisites

- Took the Deep Learning exam with good results
- Good programming skills in Python
- Experience in ML-frameworks (Preferably PyTorch)
- *Optional*: Experience in speech processing
- *Optional*: Participated in the ISS Deep Learning Lab

If this topic has sparked your interest, write me an email and we can discuss the proposal in more detail. Please include your current transcript and CV.

References

[1] Ruizhe Cao, Sherif Abdulatif, and Bin Yang. "CMGAN: Conformer-based Metric GAN for Speech Enhancement". In: *INTERSPEECH*. 2022.

[2] Tobias Raichle, Niels Edinger, and Bin Yang. "Test-Time Adaptation for Speech Enhancement Via Domain Invariant Embedding Transformation". In: *IEEE Open Journal of Signal Processing* (2026), pp. 1–10. DOI: 10.1109/OJSP.2026.3656059.

[3] Tobias Raichle, Erfan Amini, and Bin Yang. "Test-Time Adaptation for Speech Enhancement via Mask Polarization". In: *ICASSP 2026 - 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2026, pp. 18882–18886. DOI: 10.1109/ICASSP55912.2026.11464881.

[4] Julius Richter et al. "Speech enhancement and dereverberation with diffusion-based generative models". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 31 (2023), pp. 2351–2364.

Tobias Raichle
tobias.raichle@iss.uni-stuttgart.de