Safe Test-Time Adaptation for Speech Enhancement

October 19, 2026

Thesis	FA/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. Generally, speech enhancement covers multiple types of corruptions (additive noise, reverberation, echo, compression artifacts etc.) but most previous works focus on denoising.

SE models have been shown to perform well, as long as the test data resembles the training data [1]. If there is a domain shift between them, the performance degrades noticeably. Test-time adaptation (TTA) aims at restoring the source performance in previously unseen domains. As TTA happens at test-time, this adaptation must be online and unsupervised. TTA is further restricted by assuming no access to source data out of storage and privacy concerns. Previous works on TTA for SE are presented in [1, 2, 3].

While existing methods have been shown to improve average enhancement performance under domain shifts, they can become unstable over time or fail for certain samples. The goal of this thesis is to identify failure modes of existing methods and mitigating them to ensure no loss of performance.

Objectives

- Evaluate stability and robustness of existing methods
- Identify failure modes and design mitigation strategies
- Evaluate and optimize the approach

Prerequisites

- Took the Deep Learning exam with good results
- Good programming skills in Python
- Experience in ML-frameworks (Preferably PyTorch)
- *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] 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.

[2] Efthymios Tzinis et al. "RemixIT: Continual Self-Training of Speech Enhancement Models via Bootstrapped Remixing". In: *IEEE Journal of Selected Topics in Signal Processing* 16.6 (2022), pp. 1329–1341. DOI: 10.1109/JSTSP.2022.3200911.

[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.

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