

Residual 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, 4].

Previous approaches focus on improved speech reconstruction [1, 5], whereas improved noise reduction is underexplored. While this can be equivalent in some cases, the statistical simplicity of noise compared to speech enables novel adaptation mechanisms by explicitly targeting the residual noise.

Objectives

- Evaluate existing methods on noise reduction
- Design and implement noise-based adaptation techniques.
- Optimize the approach and compare to existing methods.

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.

- [4] Tobias Raichle, Ekaterina Gavrilko, and Bin Yang. *BLINC: Blind Calibration For Training-Free Speech Enhancement Adaptation*. 2026. arXiv: 2609.21898 [eess.AS]. URL: <https://arxiv.org/abs/2609.21898>.
- [5] Mike Qu, Yu-Wen Chen, and Julia Hirschberg. *Mitigating Over-Suppression in Speech Enhancement via Inference-Time Rethink-and-Refine Correction Module*. 2026. arXiv: 2608.07781 [eess.AS]. URL: <https://arxiv.org/abs/2608.07781>.