Mehmet Ali Tuğtekin Turan

Speech Scientist at Fraunhofer Institute for Intelligent Analysis and Information Systems
Address: Schloss Birlinghoven, 53757 Sankt Augustin, Germany
Mobile: +49 1525 2334922 & E-mail: tugtekin.turan@iais.fraunhofer.de
Google Scholar: https://scholar.google.com/citations?user=w3OrQIIAAAAJ

OBJECTIVE / SUMMARY

Experienced speech processing researcher with a background in automatic speech recognition, multi-speaker diarization, on-device modeling, and speech enhancement backed by hands-on expertise on the production-scale corpora. As proficient in formulating and implementing novel algorithms, particularly on transformers-based modeling, I am eager to conduct applied research that pushes the boundaries on real-world data, driving industry-leading innovations.

WORK EXPERIENCE

Research Scientist April 2021 - Present

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Bonn, Germany
www.iais.fraunhofer.de/en/business-areas/speech-technologies.html

- Working for *H2020 SELMA* project: https://selma-project.eu
- Managing the project on Fraunhofer's side, guiding the team in the algorithmic research
- Leveraging neural network architectures and exploring both supervised and unsupervised adaptation techniques for improving ASR performance
- Establishing SOTA speech and NLP technologies for broadcast & newsroom, ensuring optimization for user-feedback, accuracy, and scalability
- Investigating E2E-ASR systems specifically for low-resource setups, focusing on data processing, augmentation, and error-correction strategies

Postdoctoral Researcher June 2019 - March 2021

INRIA, Multispeech Team, Nancy, France, https://team.inria.fr/multispeech

- Worked for *H2020 COMPRISE* project: https://www.compriseh2020.eu
- Developed robust systems that operate effectively under privacy-preserving conditions, ensuring data security without compromising performance
- Achieved significant advancements in speaker-adapted acoustic modeling by leveraging speaker embeddings, leading to enhanced accuracy and user personalization
- Employed language model domain adaptation, uniquely trained on privacy-transformed sentences, resulting in a novel approach to balancing data protection with ASR efficiency

EDUCATION

Doctor of Philosophy October 2013 - March 2019

Koç University, Electrical Engineering, Istanbul, Turkey

- *Thesis:* Use of Transfer Learning for Automatic Dietary Monitoring through Throat Microphone Recordings (http://doi.org/10.5281/zenodo.3841956)
- *Summary:* It proposes a teacher/student learning paradigm for automatic dietary monitoring to improve food intake detection over wearable sensors.
- Funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK)

Master of Science September 2011 - August 2013

Koç University, Electrical Engineering, Istanbul, Turkey

- *Thesis:* Enhancement of Throat Microphone Recordings Using Gaussian Mixture Model Probabilistic Estimator (http://arxiv.org/abs/1804.05937)
- *Summary:* It focuses on improving the quality of band-limited speech and addressing unnatural sound quality using spectral mapping schemes.
- Full scholarship and research funding from Koç University

Undergraduate Degree September 2007 - June 2011

Bilkent University, Electrical Engineering, Ankara, Turkey

LANGUAGES

Turkish, *Native* (*Nationality*: Turkish)
English, *Advanced*
French, *Intermediate*
German, *Pre-intermediate (ongoing learning)*

TECHNICAL SKILLS

Programming Languages: Python, C/C++, R, Perl, Unix/Linux shell scripting

Domain-specific Tools: PyTorch & Lightning, TensorFlow, HuggingFace Transformers, ESPnet, Kaldi & K2, NVIDIA NeMo

Technical Proficiencies: RESTful web apps, Containerization with Kubernetes deployment, Distributed HPC computing, gRPC communications for streaming micro-services

Soft Skills: Project management & coordination, Scrum-based Agile practices, Source code versioning over CI/CD pipelines

JOURNAL PAPERS

- **M. A. Tuğtekin Turan** and Engin Erzin. "Improving Phoneme Recognition of Throat Microphone Speech Recordings Using Transfer Learning." Elsevier Speech Communication, 2021. `https://doi.org/10.1016/j.specom.2021.02.004`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Domain Adaptation for Food Intake Classification with Teacher/Student Learning." IEEE Transactions on Multimedia, 2020. `https://doi.org/10.1109/TMM.2020.3038315`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Source and Filter Estimation for Throat-Microphone Speech Enhancement." IEEE Transactions on Audio, Speech and Language Processing, 2016. `https://doi.org/10.1109/TASLP.2015.2499040`
- Can Yağlı, **M. A. Tuğtekin Turan** and Engin Erzin. "Artificial Bandwidth Extension of Spectral Envelope Along a Viterbi Path." Elsevier Speech Communication, 2013. `https://doi.org/10.1016/j.specom.2012.07.003`

CONFERENCE PAPERS

- Moncef Benaicha, David Thulke and **M. A. Tuğtekin Turan**. "Exploring Spoken Named Entity Recognition: A Cross-Lingual Perspective." *to be submitted for* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024.
  Pre-print available at `https://arxiv.org/abs/2307.01310`
- **M. A. Tuğtekin Turan**, Emmanuel Vincent and Dietrich Klakow. "Adapting Language Models When Training on Privacy-Transformed Data." Language Resources and Evaluation (LREC), 2022. `https://aclanthology.org/2022.lrec-1.465`
- **M. A. Tuğtekin Turan**, Emmanuel Vincent and Denis Jouvet. "Achieving Multi-accent ASR via Unsupervised Acoustic Model Adaptation." INTERSPEECH, 2020. `http://dx.doi.org/10.21437/Interspeech.2020-2742`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Monitoring Infant's Emotional Cry in Domestic Environments Using the Capsule Network Architecture." INTERSPEECH, 2018. `https://doi.org/10.21437/Interspeech.2018-2187`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Detection of Food Intake Events From Throat Microphone Recordings Using Convolutional Neural Networks." IEEE International Conference on Multimedia, 2018. `https://doi.org/10.1109/ICMEW.2018.8551492`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Empirical Mode Decomposition of Throat Microphone Recordings for Intake Classification." ACM Multimedia Workshop on Health Care, 2017. `https://doi.org/10.1145/3132635.3132640`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Synchronous Overlap and Add of Spectra for Enhancement of Excitation in Artificial Bandwidth Extension of Speech." INTERSPEECH, 2015. `https://www.isca-speech.org/archive/interspeech_2015/i15_2588.html`
- **M. A. Tuğtekin Turan** and Engin Erzin. "Enhancement of Throat Microphone Recordings by Learning Phone-Dependent Mappings of Speech Spectra." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013. `https://doi.org/10.1109/ICASSP.2013.6639029`
- **M. A. Tuğtekin Turan** and Engin Erzin. "A New Statistical Excitation Mapping for Enhancement of Throat Microphone Recordings." INTERSPEECH, 2013. `https://www.isca-speech.org/archive/interspeech_2013/i13_3244.html`