HAO ZHANG

108 W. 39th St., Baltimore, MD 21210, USA | haozhang1639@gmail.com
https://github.com/TedZhangHao | Google Scholar | Homepage https://tedzhanghao.github.io

EDUCATION

Johns Hopkins University, USA, MSE in Computer Science, Overall GPA: **3.90**/4 01/2025-12/2026 (expected)
- Completed Courses: Intro to Algorithms, Information Retrieval and Web Agents, Databases, Machine Intelligence, Introduction to Human Language Technology, Natural Language Processing, Machine Learning for Signal Processing, Information Extraction, Generative AI.

Shandong University, China, B.Eng in Intelligent Medical Engineering 09/2020-07/2024
- Overall GPA: **88.16**/100; WES Calculator Overall GPA: **3.74**/4; Jr./Sr. GPA: **92.06**/100 (**3.93**/4)
- Honors Graduate Student in Shandong University (Top 5 in Major), Academic Scholarship in 2022, 2023

PAPER & PATENT

Journal & Conference (Google Scholar Citation - 34; *indicates equal contribution; ** indicates corresponding author)
- S. Napa Ugandhar*, **H. Zhang***, ..., L. Moro-Velázquez**, *TRACE: Temporal Relationship-Aware Conversational Entrainment Detection in Dyadic Speech*, INTERSPEECH, ISCA, 2026.3 (**Under Review**)
- **H. Zhang**, A. Jing, S. Yang, T. Li**, *Sequence-Spectrogram Fusion Network for Wind Turbine Diagnosis through Few-Shot Time Series Classification*, Advanced Engineering Informatics (AEI), SCI Journal (Impact Factor: 9.9), 2025.1 (**Published**)
- **H. Zhang**, Z. Pang, J. Wang, T. Li**, *Few-Shot Learning Using Data Augmentation and Time-Frequency Transformation for Time Series Classification*, 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), IEEE, 2023.12 (**Published**)
- Z. Pang, **H. Zhang**, T. Li**, *Hybrid Fine-Tuning in Large Language Model Learning for Machinery Fault Diagnosis*, 22nd IEEE Internati- onal Conference on Industrial Informatics (INDIN), IEEE, 2024.6 (**Published**)

Invention Patent *A Method and System for Wind Turbine Condition Monitoring Based on Multimodal Feature Fusion,* 2024.8 (**Published**)

ACADEMIC RESEARCH

***Research Assistant,* Center for Language and Speech Processing** 11/2025 - Present
Supervisor: Prof. Laureono Moro-Velázquez, Johns Hopkins University

❖ **TRACE: Temporal Relationship-Aware Conversational Entrainment Detection in Dyadic Speech** 12/2025-03/2026
- Co-built DyadEE, a dyadic speech benchmark for emotional entrainment detection, totaling 21,596 dyads samples (1,143.81 h).
- Designed and implemented artifact-robust augmentation pipelines for entrained dyads via voice conversion (VC) and speech denoising, producing high-quality VC samples (105.58 h) and denoised samples (63.82 h) to encourage models to learn interaction dynamics rather than synthetic shortcuts.
- Contributed to the controlled construction of non-entrained dyads, including partner swapping and emotion-contradicting resynthesis, as well as dataset quality control and balance over relationship and conversational context using stratified sampling.
- Played a major role in strengthening both the empirical study and the manuscript, including baseline selection, ablation design, results analysis, and development of clear pipeline figures/tables, helping turn the project into an **INTERSPEECH** submission.

❖ **Natural Human-AI Verbal Conversation Dataset Construction Pipeline** 11/2025 - Present
- Led the design and early deployment of a naturalistic huma-AI speech interaction pipeline, integrating human-subject workflow, participant recruitment, recording infrastructure, pilot validation, and evaluation design.
- Designed controlled topic and assignment protocols and evaluated SoTA real-time speechLLMs (e.g. gpt-realtime, Nova 2 Sonic, and cascaded baselines) while minimizing confounds from topic repetition, ordering effects, and speaker-specific variability.
- Developed structured human-evaluation rubrics to assess conversational naturalness, paralinguistic cue understanding, emotion entrainment, and other fine-grained interaction quality.
- Designed a scalable platform concept for automated real-time human-AI benchmarking and built a reliable single-device recording pipeline with VB-Audio and Zencastr; Conducted a pilot study with 7 participants averaging 70 minutes per session.

Supervisor: Prof. Philipp Koehn & Prof. Berrak Sisman, Johns Hopkins University

❖ **Speaker-disentangled Expressive Speech-to-speech Translation (S2ST) Using Aligned Dubbing Data** 01/2026 - Present
- Established a data-driven hypothesis that aligned dubbing pairs can serve as expressive supervision for S2ST by systematically validating prosodic preservation with stopes metrics (AutoPCP, locol prosody, vocal similarity), duration and speed compliance.
- Curated an aligned expressive S2ST corpus from Netflix de-en dubbing data with 369,137 sample pairs totaling 377.67 hours for training and evaluation.
- Designed a speaker-disentangled expressive S2ST framework that separates semantic translation from speaker/expressivity modeling, using unit-based encoder, CAMPPlus speaker embeddings, DiT+CFM acoustic generation, and BigVGAN vocoding.

Independent Research

❖ **Learning Sentiment Dynamics: Knowledge-Augmented Sequence Modeling for Document-Level Aspect-Based Sentiment Analysis (DL-ABSA)** 05/2025 - Present
- Designed an ABSA framework integrating FLAN-T5 to improve document-level aspect-based sentiment classification (ASC).
- Built domain-specific knowledge bases via Chain-of-Thought (CoT) prompting to generate aspect candidates with GPT-5.2.
- Introduced the concept of aspect triggers, enabling less ambiguous and more fine-grained identification of aspect occurrences across long documents.
- Devised Next Polarity Prediction to explicitly modeling sentimental intensity sequence among reviews using the decoder with teacher forcing supervised by knowledge base matched pseudo label, addressed cross-sentence sentimental conflicts.

***Research Assistant,* Key Laboratory of Machine Intelligence & System Control, Ministry of Education** 07/2023-06/2024
Supervisor: Prof. Teng Li, Shandong University

❖ **Sequence-Spectrogram Fusion Network for Machinery Diagnosis through Few-Shot Time-Series Classification** 09/2023-06/2024
- Proposed a novel network consisting of scalable blocks for hierarchical multimodal fusion and joint representation learning.
- Devised an embedded hybrid encoder combining multi-head self-attention (MSA) and convolution module to enhance feature extraction, optimize the alignment of latent spaces, and allocate modalities weight adaptively for quality feature fusion.
- Conducted thorough comparative and ablation studies, demonstrating the outperformance and interpretability of our network.
- Yielded 0.996 macro AUC on 2 multi-class few-shot wind turbine fault datasets, composing a paper submitted to **AEI** journal.

❖ **Hybrid Fine-Tuning in Large Language Model (LLM) Learning for Machinery Fault Diagnosis (INDIN)** 03/2024-05/2024
- Configured LLaMA2-13b using Hugging Face, implementing layer pruning to enhance training efficiency.
- Co-designed a parameter-efficient fine-tuning (PEFT) method to tailor LLaMa2 for time-series analysis: integrating Freeze and LoRA fine-tuning techniques to differentiate LLM downstream, generic, and embedding layers.
- Conducted supervised fine-tuning (SFT) using adaptive learning rate adjustment, achieving 96%+ accuracy and maximally 16.6% improvement on 6 bearing-fault datasets compared to other algorithms and PEFT methods. (Based on Alibaba One-Fits-All frame).

WORK EXPERIENCE

***Algorithm Engineer Intern,* iFLYTEK CO.LTD.** 07/2024-11/2024
Business Unit: Xunfei Healthcare, Research Institute

❖ ***Xiaoyi* Medical Interrogation LLM Optimization: Data-Centric LLM Fine-Tuning for Patient Persona Slot Extraction**
- Designed Chain-of-Thought prompting strategies to integrate GPT-4o into the data pipeline, including: (1) generating synthetic dialogues to compensate for missing labels or under-represented slot data, and (2) performing efficient quality screening, correction, and standardization of dialogues pre-annotated by the deployed *Xiaoyi* LLM.
- Developed a human-in-the-loop annotation framework, streamlining collaboration with doctors and crowdsource annotators through automated task routing and data recollection. Reduced rework ratio by 14% and saved 110 person-hours per month.
- Constructed a high-quality supervised fine-tuning corpus from 10GB+ online dialogues, applied instruction-tuning to post-train *Xiaoyi*, and achieved 10%+ F1 improvement in average across 7 critical patient slots.

❖ ***Xiaoyi* Backend Optimization**
- Developed RESTful services in Spring Boot for task dispatching and data collection, enabling smooth interaction with data preprocessing modules.

SELECTED COURSE PROJECTS

IMDb Movie Review Multifunctional Engine (Information Retrieval and Web Agent) 03/2025-05/2025
- Designed a hybrid web scraping pipeline combining static parsing and JavaScript-rendered dynamic operations with BeautifulSoup and Selenium, enabling full extraction of metadata and 100k+ user reviews from IMDb.
- Constructed a user-item rating matrix and applied KNN imputation with truncated SVD to predict missing ratings, enabling personalized movie recommendations.
- Developed a MobileBERT-based regression model to predict both the mean and variance of sentiment distributions from multi-review inputs, capturing audience consensus and controversy, and eventually achieved a RMSE of 1.91 on a 0-10 scale rating score.

SKILLS

Computer Language: Python/Matlab (advanced), SQL/R (intermediate).
IDE and Software: PyCharm, Visual Studio, Jupyter, MySQL Workbench.
Framework, Package, and Tool: Pytorch, Pytorch Lightning, TensorFlow, Keras, HuggingFace, transformers, Scikit-learn, Pandas, Numpy, PyWavelets, SciPy, NLTK, OpenCV, sktime, Matplotlib, Seaborn, EEGLAB, WFDB, Latex, Git.
Data Processing Skillset: Management, Cleaning, Filtering, Resampling, Augmentation, Visualization, and Statistical Analysis.

EXTRACURRICULAR HIGHLIGHTS

Student Admission Reviewer, Department of Computer Science, Johns Hopkins University
Academic Reviewer
- Journal: Knowledge-based Systems, MethodsX
- Conference: 2024 International Conference on Biomimetic Intelligence and Robotics & Medical Robot Forum (ICBIR).

Oral Presentation at 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI).