Arindam Ghosh

Research Specialist, Solventum (3M Health)
Machine Learning for Speech and Healthcare AI

Email:arindam.gm@gmail.com Google Scholar ORCID: 0000-0002-4545-087
Homepage:https://arndmghosh.github.io/ LinkedIn

SUMMARY

My current research focuses on healthcare AI solutions that leverage speech and natural language processing (NLP) to automate clinical workflows, enabling doctors to focus more on patient care. This involves developing robust wake-word detection, automatic speech recognition (ASR), and speaker-role diarization systems for real-world doctor–patient conversations in noisy clinical environments, as well as automatic summarization of these conversations using NLP techniques. I also work on calibration of neural networks to improve system reliability. These efforts have resulted in production deployments used daily by doctors in U.S. healthcare systems, as well as publications at top-tier venues including NeurIPS, INTERSPEECH, and EMNLP.

Previously, my research was in wireless communications, underwater sensor networks, and free-space optical communication, leading to publications in IEEE Communications Letters, Sensors, and IEEE SPCOM. I have also worked as a hardware engineer for development of high-speed network switches and router hardware deployed across multiple secure and mission-critical networks for the Government of India.

EDUCATION

- **Carnegie Mellon University**, Pittsburgh, USA — Jan 2019 - Dec 2019
  Master of Science in Electrical and Computer Engineering
  Graduated early within one year by completing the three-semester coursework in only two semesters.
- **National Institute of Technology Durgapur**, Durgapur, India — July 2009 - May 2013
  Bachelor of Technology in Electronics and Communication Engineering

PROFESSIONAL APPOINTMENTS

- **Research Specialist, Solventum (formerly 3M Health)**, Pittsburgh, USA — June 2020 - Present
  - Developed in-house transducer-based (RNN-T) automatic speech recognition (ASR) systems for doctor–patient conversations, improving accuracy, reliability, and robustness, and deployed in real-world clinical systems used by doctors in U.S. healthcare environments.
  - Designed a parallel-branch ASR-synchronous speaker-role diarization system, enhancing speaker-role prediction without affecting ASR accuracy, and implemented a novel blank-suppression decoding heuristic that reduced ASR word error rate by 2% (relative) in a production-grade system (INTERSPEECH 2025).
  - Developed low-resource, low-footprint on-device wake-word detection models for clinical workflows, exceeding internal performance targets (INTERSPEECH 2022).
  - Proposed a novel neural network model calibration method called AdaFocal that significantly improved calibration without loss of accuracy, rapidly becoming a widely used benchmark in the model-calibration community (NeurIPS 2022).
  - Developed a multi-stage summarization system for clinical dialogues, fine-tuning large pretrained language models to generate fluent, clinically meaningful summaries from noisy doctor–patient conversation transcripts (EMNLP 2021).
- **Research Assistant, Carnegie Mellon University**, Pittsburgh, USA — Jan 2020 - May 2020
  - Worked in the Department of Electrical and Computer Engineering, under the supervision of Prof. Ian Lane, focusing on end-to-end automatic speech recognition (ASR) for conversational speech.
  - Designed and trained dialogue-context–aware ASR models by incorporating conversational history into sequence-to-sequence architectures to improve recognition robustness in multi-turn dialogue settings.
  - Implemented and benchmarked state-of-the-art ASR systems using Kaldi and wav2letter, training on the Switchboard and CallHome corpora and evaluating performance on the Hub5 2000 test set.

- **Research Engineer, Centre for Development of Telematics**, Bangalore, India Aug 2013 - Dec 2018
  - Designed and developed high-throughput network processor and Packet-over-SONET (Synchronous Optical Network) hardware modules for multi-terabit-capacity routers.
  - Led the complete product lifecycle for these modules, spanning system conceptualization, hardware architecture and circuit schematic design, development, and system-level testing, including module bring-up, performance validation, and cross-functional integration.
  - Conducted signal integrity simulations using HyperLynx to ensure reliable high-speed PCB routing; implemented VHDL logic for on-board programmable devices; and performed power-up validation and JTAG-based functional testing.
  - Developed a mixture-based mathematical framework for predicting wireless network coverage under spatially correlated interference, enabling the derivation of outage probabilities for MRC receivers that were previously intractable with existing tools (IEEE Communications Letters 2017 and 2018).

PUBLICATIONS

Journal Papers:

- **A. Ghosh**, J-W. Lee, and H-S. Cho, "Throughput and Energy Efficiency of a Cooperative Hybrid ARQ Protocol for Underwater Acoustic Sensor Networks," **Sensors 2013** (39 citations).
- **A. Ghosh** and H. S. Dhillon, "Performance Analysis of MRC Under Spatially Correlated Interference Using Mixture-Based Method," **IEEE Communication Letters 2018** (4 citations).
- **A. Ghosh**, "Mixture-Based Modeling of Spatially Correlated Interference in a Poisson Field of Interferers," **IEEE Communication Letters 2017** (3 citations).

Conference Papers:

- **A. Ghosh**, T. Schaaf, and M. Gormley, "Adafocal: Calibration-aware adaptive focal loss," **NeurIPS 2022** (74 citations).
- **A. Ghosh**, M. Fuhs, D. Bagchi, B. Farahani, M. Woszczyna, "Low-resource Low-footprint Wake-word Detection using Knowledge Distillation," **Interspeech 2022** (8 citations).
- **A. Ghosh**, G. Ghatak, and A. Chandra, "SEP of dual-ring star-QAM over FSO channels with atmospheric turbulence," **IEEE SPCOM 2014** (5 citations).
- L. Zhang, R. Negrinho, **A. Ghosh**, V. Jagannathan, H. Hassanzadeh, T. Schaaf, and M. R. Gormley, "Leveraging Pretrained Models for Automatic Summarization of Doctor-Patient Conversations," **EMNLP 2021** (85 citations).
- Kim, B., **Ghosh, A.**, Fuhs, M.C., Chowdhury, A., Bagchi, D., Woszczyna, M., "A Hybrid Approach to Combining Role Diarization with ASR for Professional Conversations," **Interspeech 2025** (2 citations).

Citations:

Total: 220	**First Author: 133**	h-index: 5	i10-index: 3

REVIEWING ACTIVITIES

- Reviewer for IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2026
- Reviewer for Neural Information Processing Systems (NeurIPS) 2025 (**Top Reviewer award**)
- Reviewer for INTERSPEECH 2025
- Reviewer for IEEE Spoken Language Technology Workshop (SLT) 2024