Mohammad Nur Hossain Khan

Graduate Research Assistant, ECE, Worcester Polytechnic Institute
mkhan@wpi.edu mohammad-nur-hossain-khan nurhossain2301 +1-774-519-0452

Summary

› Passion for research in audio-LLM, context-aware LLM, time-series analysis, health data processing, audio signal processing, multi-modal analysis.
› Published several conference papers showing impactful research skills.
› Proficient in LLM Agent and deep learning model development.

Educational Background

August 2022-Present | **PhD in Electrical and Computer Engineering**
Worcester Polytechnic Institute, Worcester, MA, USA.
› CGPA : **3.92** on a scale of 4.00

August 2022-May 2025 | **Masters of Science in Electrical and Computer Engineering**
Worcester Polytechnic Institute, Worcester, MA, USA.
› CGPA : **3.87** on a scale of 4.00

March 2011-March 2016 | **Bachelor of Science in Electrical and Electronic Engineering**
Bangladesh University of Engineering & Technology (BUET), Dhaka, Bangladesh.
› CGPA **3.73** on a scale of 4.00

Technical Skills

Languages	Python, MATLAB, C
Libraries	Scikit-learn, OpenCV, NLTK, PyTorch, TensorFlow, Keras, Matplotlib, Seaborn
Web Base Tools	Git, GitHub, Latex
Micro Controllers	Arduino, Raspberry-pi

Research and Work Experiences

August 2022-Present | **Graduate Research Assistant (GRA)**, Worcester Polytechnic Institute, USA
Supervisor : Dr. Bashima Islam

LittleBeats-LLM : Leveraging Audio LLM for Accurate Target Speaker Labeling in Real-World Naturalistic Home Recording
We use LittleBeatsTM (LB), an infant wearable multi-modal device, to collect fam- ily audio and enhance its audio pipeline by lever- aging audio large language models (LLMs). Our model, LittleBeats-LLM (LB-LLM), can reliably and precisely label speaker types and vocalizations for target family members, including infants, par- ents, and siblings, at a precise temporal resolution of 0.1 seconds, while detecting and distinguish- ing non-family members as a secondary goal. We conduct extensive experiments demonstrating that LB-LLM outperforms traditional audio-only mod- els across multiple synthetic datasets with varying signal-to-noise ratios (SNRs). We open-source LB-LLM to facilitate relevant research on infants' vo- cal interactions with family members in the home environment. To the best of our knowledge, this is one of the first studies using audio LLMs for labeling family vocal interactions in real-world naturalistic home recordings.

› Developed an audio-LLM (LB-LLM) to reliably label speaker types and vocalizations for target family members—including infants, parents, and siblings—at 0.1-second resolution, while also detecting non-family members as a secondary objective.
› Developed LB-LLM with multi-task capabilities of audio captioning, event describing with time duration, speaker counting, and speaker diarization.
› Created an audio captioning dataset to describe the audio event for various-duration audio.

Infant behaviour and interaction analysis

Infant behavior and interaction with parents/caregivers are crucial to determining a child's physiological and psychological development. In this study, we determine the baby's position, movement, interaction with parents, and vocalization by using motion, ECG, and audio data that are collected using *Littlebeats*^{TM}. This study is conducted in collaboration with the University of Illinois.

- Developed an unsupervised deep learning model using contrastive loss to get embedding of baby position and motion from unlabeled IMU data.
- Used audio spectrogram transformer to classify environmental noises collected from an Infant-centric soundscape.

Measuring mindfulness skills during meditation
We assess the potential of using acoustic and motion sensors to sense the improvements in mindfulness skills passively during self-paced meditation. We study the impact of respiratory signal feedback on enhancing mindfulness skills and propose a respiratory rate estimation model. We develop and evaluate our model and perform a user study on data collected using a smartphone application for mindfulness training. This study is conducted in collaboration with Carnegie Mellon University.

- Developed an algorithm using statistical signal processing to get the respiration rate from accelerometer data with an average error of 1.6 respiration rate per minute.
- Defined slow-paced breathing in terms of breathing rate to quantify the relationship between slow-paced breathing rate and mindfulness and developed a machine learning algorithm to measure mindfulness skills from raw respiration rate.
- Developed a deep learning model to measure changes in mindfulness skills during meditation from motion signals.

Context-aware Large Language Model (LLM) we address the underexplored area of sensor-based question answering (SQA), where traditional LLMs struggle due to the dense, noisy, and semantically sparse nature of IMU signals. To tackle this challenge, we introduce two large-scale datasets : SensorCap, comprising 35,960 IMU–caption pairs, and OpenSQA, containing 199,701 question–answer pairs designed for causal and explanatory reasoning. We also curate a 19,440-example tuning fold to enhance scientific accuracy and diagnostic relevance. Building on these resources, we develop LLaSA-7B and LLaSA-13B, multimodal LLMs that incorporate a sensor-aware encoder and projection module to enable open-ended, context-aware reasoning over IMU data.

- Developed a multi-modal large language model that can include sensor information and is capable of interpreting and responding to activity and motion analysis queries.
- Developed SensorCaps, a dataset of 26,288 IMU-derived activity narrations, and OpenSQA, an instruction-following dataset with 257,562 question-answer pairs.

August 2023-May 2024 **Graduate Teaching Assistant, Worcester Polytechnic Institute, USA**

- Courses : On-Device Deep Learning, Discrete-Time Signal And System Analysis, and Introduction To Communications And Networks
- Prepared course materials, demo videos, project ideas, assignments, and homework.
- Assessed and graded projects, assignments, presentations, and paper reviews.

August 2017-May 2022 **Assistant Engineer, Electricity Generation Company of Bangladesh, Bangladesh**

Publications

2025 **Mohammad Nur Hossain Khan** [1], Subrata Biswas[1], and Bashima Islam, "RAVEN : Query-Guided Representation Alignment for Question Answering over Audio, Video, Embedded Sensors, and Natural Language." *arXiv preprint arXiv :2505.17114 (Accepted at EMNLP main conference, 2025)*

2025 **Mohammad Nur Hossain Khan**, Mohammad Nur Hossain Khan, David creswell, Jordan Albert, Patrick O'Connell, Shawn Fallon, Mathew Polowitz, Xuhai "orson" Xu, and Bashima islam. 2025. Mindfulness Meditation and Respiration : Accelerometer-based Respiration Rate and Mindfulness Progress Estimation to Enhance App Engagement and Mindfulness Skills. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 9, 3, Article 95 (September 2025). https ://doi.org/10.1145/3749498*

2025 **Mohammad Nur Hossain Khan**[1], Subrata Biswas[1], and Bashima Islam, "QUADS : QUAntized Distillation Framework for Efficient Speech Language Understanding." *Proc. Interspeech 2025, 4098-4102, doi : 10.21437/Interspeech.2025-532*

2025 Subrata Biswas, **Mohammad Nur Hossain Khan**, and Bashima Islam, "LOCUS – LOcalization with Channel Uncertainty and Sporadic Energy" *Accepted at EWSN, 2025*

2025 Sheikh Asif Imran, **Mohammad Nur Hossain Khan**, Subrata Biswas, and Bashima Islam, "LLaSA : A Sensor-Aware LLM for Natural Language Reasoning of Human Activity from IMU Data" *In Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25), October 12–16, 2025, Espoo, Finland. ACM, New York, NY, USA, 7 pages. https ://doi.org/10.1145/3714394.3756187*

2024 **Mohammad Nur Hossain Khan**, Jialu Li, Nancy L. McElwain, Mark Hasegawa-Johnson, and Bashima Islam, "Sound Tagging in Infant-centric Home Soundscapes" *In Proceedings of the IEEE/ACM Conference on Connected Health : Applications, Systems and Engineering Technologies (CHASE), IEEE, 2024)*

2024 **Mohammad Nur Hossain Khan**, Nancy L. McElwain, Mark Hasegawa-Johnson, and Bashima Islam, "InfantMotion2Vec : Unlabeled Data-Driven Infant Pose Estimation Using a Single Chest IMU" *In 2024 IEEE 20th International Conference on Body Sensor Networks (BSN), pp. 1-4. IEEE, 2024.*

2023 Jannatul Ferdaous Progga, **Mohammad Nur Hossain** Khan, and MG Mostofa Amin. "Meteorological Parameters–Soil Temperature Relations in a Sub-Tropical Summer Grassland : Physically-Based and Data-Driven Modeling." *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, **54(2)**, 2023, pp. 48-56

2019 **Mohammad Nur Hossain Khan**, Shaikh Anowarul Fattah, "Surface EMG Based Basic Hand Movement Detection using Singular Value Decomposition." *In 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), 2019,* pp. 97-102

References

Available on Request

1. Equal Contribution