Md Kamrujjaman Mobin

Software Engineer I, Samsung R&D Institute Bangladesh

kamrujjamanmobin123@gmail.com · +880-1749-500-360 · kmobin555.github.io · Google Scholar · LinkedIn · GitHub

Research Interests

Multimodal Learning, Vision-Language Models, Event-Based Vision, Long-Horizon Video Understanding, Spatial Reasoning, Embodied AI, Tool-Use LLM Agents, Health AI, LLM Safety & Robustness.

Education

Shahjalal University of Science and Technology (SUST) *Sylhet, Bangladesh*
B.Sc. in Computer Science and Engineering *CGPA: **3.83 / 4.00*** Feb 2020 – Jul 2025

Undergraduate Thesis: *AI-Driven Health System for Mortality Prediction Using Electrocardiograms (ECGs) with Transformers and Selective State-Space Models* [PDF]
Focused on deep sequence modeling, representation learning, and AI-driven healthcare analytics.

Professional Experience

Samsung R&D Institute Bangladesh *Dhaka, Bangladesh*
Software Engineer I July 2025 – Present

- **Personal Health Agent (PHA)**: On-device Android (Kotlin) agent that unifies data from multiple health platforms (Samsung Health, MyFitnessPal, Strava, etc.) via Google Health Connect and an MCP-style tool architecture, answering user queries with personalized insights and cross-app correlations. Fine-tuned **Gemma 4 2B-IT** for tool calling and compact system prompts, and deployed it with **LiteRT-LM** using on-device GPU acceleration to substantially reduce inference time.
- **Brain Pulse App** (launching soon): Android app for longitudinal cognitive health monitoring, correlating passive device-usage patterns with brain health biomarkers via on-device ML. Built multi-module Clean Architecture (Hilt DI, Room DB, WorkManager); passive pipelines for app usage, screen events, sleep, and call logs; on-device ONNX/TFLite inference over a 10+ biomarker feature extraction pipeline.
- **RSDS Mobile SDK** (Remote Sensor Data Service): Clinical-grade cross-device SDK enabling large-scale health research by collecting **40+ sensor modalities** from smartphone and Wear OS (PPG, ECG, EDA, BIA, accelerometer, sleep stages, app usage). KMP multi-module architecture with compile-time DI, SQLCipher-encrypted local storage, GZIP sync, and AES file encryption for HIPAA-compliant data handling.
- **Compass Watch App** (Galaxy Smartwatch 7 & 8): Maintained a system utility for wearable sensor calibration and validation. Improved codebase via unit testing and clean-code refactoring: **7.5% SAM score improvement**, **20% CBO reduction**, **90%+ unit test coverage**.
- Conducted applied R&D on **non-invasive blood glucose estimation** from PPG signals and multimodal **Alzheimer's risk modeling** — both targeting Samsung Health's **450M+ active user** platform.

Publications

Conference Papers · 2025

[1] **Mobin, M. K.**, Islam, M. S., Al Barid, S., & Masum, M.
Cardioformer: Advancing AI in ECG Analysis with Multi-Granularity Patching and ResNet.
38th Canadian Conference on Artificial Intelligence **(Canadian AI 2025)**, Calgary, Canada. arXiv:2505.05538

[2] **Mobin, M. K.** & Islam, M. S.
LuxVeri at GenAI Detection Task 1: Inverse Perplexity-Weighted Ensemble for Robust Detection of AI-Generated Text across English and Multilingual Contexts.
31st International Conference on Computational Linguistics **(COLING 2025)**, Dubai, UAE. arXiv:2501.11914

[3] **Mobin, M. K.** & Islam, M. S.
LuxVeri at GenAI Detection Task 3: Cross-Domain Detection of AI-Generated Text Using Inverse Perplexity-Weighted Ensembles.
31st International Conference on Computational Linguistics **(COLING 2025)**, Dubai, UAE. arXiv:2501.11918

2024

[4] **Mobin, M. K.**
Transforming ECG Images to Waveforms with UNet-LSTM and Multilabel Classification Using ResNet-LSTM: A Multimodal Approach.
51st Computing in Cardiology **(CinC 2024)**, Karlsruhe, Germany. [PDF]

[5] **Mobin, M. K.**, Al Barid, S., & Rahman, M. S.
CricPose360: A Comprehensive Dataset for Real-Time Pose-Enhanced Cricket Shot Classification Using YOLOv8.

27th International Conference on Computer and Information Technology **(ICCIT 2024)**, Bangladesh.
DOI: 10.1109/ICCIT64611.2024.11022536

Manuscripts Under Review

[U1] **Mobin, M. K.**, *et al.*
The Serialization Bottleneck: Separating State Preservation from Property Derivation in Tool-Using Agents.
Submitted to the 15th International Conference on Learning Representations **(ICLR 2027)**, under review.
Diagnostic benchmark (3 domains, 25 properties, 7 LLMs, 43.7K queries) separating two agent failure points: *deriving* a branch property from serialized state vs. *re-emitting* that state as a tool argument. Passing object references instead of payloads cuts aborts from 99.7% to 1.2%; runtime summaries turn derivation into lookup.

Technical Reports

[T1] **Mobin, M. K.**, Shamil, M. S., & Arefeen, M. A.
Training-Free Representation and Prompting for Event-Based Multimodal Understanding.
Technical Report, EventBench General Understanding Track, Workshop on Event-Based & Multimodal Vision **(EBMV @ ECCV 2026)**. [PDF] [Competition]
Training-free system (no challenge training data) rendering event streams as polarity-coded, temporally ordered image grids for a frozen VLM (Qwen3.5-9B) with task-specific bin selection, cropping, and prompting. Ranked **9th of 27 teams** (62.28% avg. across 5 tasks).

Ongoing Research

- **Event-Based Vision for Long-Horizon Video Understanding** Jan 2026 – Now
  Investigating event-stream representations and temporal models for long-horizon egocentric video understanding on the newly released **EventKitchen** dataset (ECCV 2026). Builds on earlier hands-on work on N-EPIC Kitchens voxel-grid generation, E2VID reconstruction [NB1][NB2], and SSM+SNN hybrid temporal modeling [NB3] toward spike-native state-space models and noise-aware ANN–SNN hybrids.
- **Spatial Reasoning in VLMs (MindCube)** · Challenge · EvalAI 2026 – Now
  Fine-tuned Qwen2.5-VL-3B-Instruct (QLoRA) with structured cognitive-map + chain-of-thought outputs that force explicit spatial reasoning before prediction; ensembled it with the official GRPO-trained model via self-consistency. Now extending this into a full research study.

Research Projects & Competitions

- **GSA-VLN: General Scene Adaptation for VLN** · *Embodied AI, ICLR 2025 Replication* · GitHub 2026
  Replicated GR-DUET (ICLR 2025) on R2R/Matterport3D. Implemented two novel improvements: **Semantic-Aware Navigation** with InfoNCE contrastive loss grounding instructions to room-type-matching graph nodes (**+1.4% SR, +1.6% SPL**), and **Experience Replay** for continual learning across scenes (**+0.7% late-episode SR**).
- **Kaggle Competitions** 2022 – Present
  - **Image Matching Challenge 2024** — **Bronze Medal** (Top 6.3%): Built a multi-model 3D reconstruction pipeline integrating COLMAP with deep feature matchers. [Notebook]
  - **ISIC 2024: Skin Cancer Detection** — **Bronze Medal** (Top 5.7%): Developed CNN-based multimodal models with advanced augmentations for melanoma classification. [Notebook]

Technical Skills

Programming	Python, C++, Java, Kotlin, JavaScript, C
ML / AI Frameworks	PyTorch, Hugging Face Transformers, Scikit-learn, OpenCV, NumPy, Pandas
LLM / VLM	QLoRA fine-tuning, SFT/GRPO models, self-consistency ensembling, batch inference
Research Areas	CV, NLP, Multimodal & Event-Based Vision, Embodied AI, Signal Processing
Academic	Research Writing, Experimental Design, Scientific Presentation, Peer Review

Language Proficiency

IELTS Academic: 7.0 (Listening 7.5 | Reading 7.0 | Speaking 6.5 | Writing 6.0) November 2025

References

Md Saiful Islam
Data Scientist, Dept. of CS
Athabasca University, Canada
saiful-cse@sust.edu
Google Scholar | LinkedIn

Prof. Mohammad Shahidur Rahman
Professor, Dept. of CSE
SUST, Bangladesh
rahmanms@sust.edu
Google Scholar | LinkedIn

Prof. Md Masum
Professor & Head, Dept. of CSE
SUST, Bangladesh
masum-cse@sust.edu
Personal Page | IEEE Xplore