SHADMAN SAKIB

Ph.D. Researcher, Information Systems (AI/ML), UMBC
Agentic AI · Vision-Language Reasoning · Multimodal Learning · Trustworthy Health AI
Open to Summer 2027 research internships

ssakib1@umbc.edu
Personal Website
in/shadman-sakib15
Google Scholar

RESEARCH SUMMARY

Ph.D. researcher building reliable, evidence-grounded **vision-language and multimodal AI** for health: **agentic AI**, **multimodal reasoning**, and **physics-informed** models for physiological sensing. **35+ peer-reviewed papers** at **WACV**, **CIKM**, **NeurIPS**, **CVPR**, **ICML**, **MICCAI**, **BMVC**, **AAAI**, **IEEE/ACM CHASE**, and **IEEE ICMLA** (incl. workshops); **2× Best Paper Award**.

EDUCATION

Ph.D. in Information Systems (AI/ML Track), *University of Maryland, Baltimore County (UMBC)* Jan 2024 – Present
GPA **4.00/4.00** · M.Sc. in Information Systems (AI/ML), UMBC, Aug 2025, GPA 4.00/4.00 · Invited to Phi Kappa Phi

B.Sc. in Electrical and Electronic Engineering, *IUBAT, Dhaka, Bangladesh* Aug 2015 – Aug 2019
GPA 3.96/4.00 · **Ranked 1st of 66** · Dean's List, every term · University Merit Scholarship

RESEARCH EXPERIENCE

Graduate Research Assistant, *MPSC Lab, UMBC* Jan 2024 – Present

- Built **Pixels-to-Phase**, a reliability-aware latent state-space model for respiratory estimation from video (**WACV 2027**), and reliability-guided video token learning for vital-sign estimation (**CIKM 2026**).
- Proposed **ViTaL-Diff**, a video-conditioned latent diffusion transformer that generates respiratory waveforms with uncertainty estimates, cutting MAE by up to **28.9%** over deterministic deep baselines on 3 datasets.
- Designed agentic VLM systems that decouple perception from reasoning and verify evidence before predicting: **VLS-RR** (**IEEE/ACM CHASE 2026**), **CALIBRA** multi-agent verification, and **EviCal-RR**.
- Benchmarked VLM reliability on physiological video: active evidence acquisition and quantitative temporal reasoning (2 papers, **NeurIPS 2026 Med-Reasoner Workshop**) and counterfactual auditing of health agents (**AAAI 2026 Fall Symposium Series**).
- Developed **RespFormer** (ICMLA 2025) and **E2RespUNet** (SmartComp 2025); curated multimodal RGB/NIR/IR datasets.

Graduate Research Assistant, *CARDS, UMBC, in collaboration with DEVCOM Army Research Laboratory* May 2025 – Aug 2025

- Deployed an end-to-end on-device respiratory-rate system on a **Boston Dynamics Spot** robot (**Jetson Orin**, ∼22 W); field-tested at ARL R2C2 in the 2025 ArtIAMAS Experimentation Event using the DVPG.
- Hardened motion extraction against viewpoint, lighting, occlusion, and robot motion (optical flow, distance compensation, SQI gating); generalized it to **3 robots** and **4 camera modalities**, with RGB range up to **8 m**.

Earlier: Graduate RA, *QUEST Lab, UNC Charlotte* (2023; deep learning for quantitative ultrasound) · Research Assistant, *AMERI, University of Hyogo, Japan* (2019–20; real-time surgical-video understanding, HUMAP Fellow)

SELECTED PUBLICATIONS

35+ peer-reviewed · first author unless marked † · Google Scholar

WACV 2027	Pixels-to-Phase: Reliability-Aware Latent State-Space Modeling for Contactless Respiratory Estimation
CIKM 2026	Reliability-Guided Video Token Learning for Contactless Vital Sign Estimation
NeurIPS 2026	**2 papers:** Measure What Matters: Active Evidence Acquisition for Reliable Physiological Reasoning; Can VLMs Keep Time? Probing Quantitative Temporal Reasoning in Physiological Videos *(Med-Reasoner Workshop)*
CVPR 2026	CALIBRA: Calibration-Aware Multi-Agent Verification for Contactless Physiological Monitoring *(GRAIL-V Workshop)*
ICML 2026	ViTaL-Diff: Video-Token Latent Diffusion for Contactless Respiratory Monitoring *(SD4H Workshop)*
MICCAI 2026	EviCal-RR: Evidence-Calibrated Reasoning for Contactless Respiratory Monitoring *(SAFER Workshop)*
BMVC 2026	Topology-Reliable Retinal Artery–Vein Segmentation via Spatially-Aware Differentiable Graph Pruning † *(BD-MedVis, **Oral**)*
AAAI 2026	Does the Agent Follow the Evidence? Counterfactual Auditing for Trustworthy Multimodal Health Agents *(Fall Symp., **Oral**)*
CHASE 2026	Decoupling Perception and Reasoning for Contactless Respiratory Rate with Vision Language and Small Language Models
ICMLA 2025	RespFormer: A Motion-Guided Temporal-Frequency Multimodal Fusion Transformer for Contactless Respiratory Monitoring

HONORS, SERVICE & TEACHING

- **2× Best Paper Award** (IEEE IEMCON 2021; IEEE R10 HTC 2021) · Doctoral Research Poster Award, UMBC (2025, 2026) · NSF Student Travel Grant (CHASE 2026) · Student Travel Grant (ICMLA 2025)
- **Reviewer** for 13 journals and 19 conferences/workshops (IEEE ISBI, MICCAI, IEEE BigData, NeurIPS/ICML/CVPR workshops) · Program Committee, AAAI FSS-26 · Organizing team, NSF TRACE Workshop 2026 · Local organizer, IEEE PerCom 2025
- **Teaching:** Lecturer in CSE, *Leading University* (2020–22; supervised published theses) · Graduate TA, *UMBC* (80+ students)
- **Field demos:** NSF TRACE Testbed (UMBC, 2026; gallery, workshop) · DVPG Remote Robotics Experiment (ARL, 2025)

TECHNICAL SKILLS

- **Languages & ML:** Python, C/C++, MATLAB, SQL · PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn
- **Agentic & Multimodal AI:** VLMs/MLLMs, multi-agent reasoning, RAG, diffusion, physics-informed & state-space models, PEFT
- **Vision & Signals:** OpenCV, Detectron2, optical flow, point tracking · spectral analysis, SQI, RGB/NIR/thermal fusion
- **Edge & Systems:** NVIDIA Jetson, CUDA, quantization, distillation · Boston Dynamics Spot · Docker, MLflow, FastAPI