Arman Nik Khah

Dallas, TX | arman.nikkhah@utdallas.edu | (682) 226-5082 | linkedin.com/in/arman-nik-khah | github.com/IamArmanNikkha

Summary. Graduate researcher (M.S. '25; PhD in progress) in **audio/vision ML**, **XR/360° streaming**, and **low-latency media systems**. Published **IMX Workshops '24** paper on **unsupervised spatial-audio surprise detection** (CVAE–LSTM); two submissions on **geometry-correct viewport prediction (SAE)** and **risk-aware ABR (MARC)**. Strengths: **PyTorch**, **C/C++ & Python**, **WebRTC/FFmpeg/GStreamer**, **QoE/ABR**, **CUDA/NVENC (foundational)**, **calibration** (ECE/NLL), **risk (CVaR)**. Interested in **Speech/Audio**, **Realtime ML**, and **Systems/Networking** research internships.

Education

PhD in Intelligent Systems (in progress), University of Texas at Dallas — Computer Science *Expected Aug 2028*
GPA: 3.8/4.0 Advisor: Dr. Ravi Prakash Research: viewport prediction, uncertainty calibration, risk-aware streaming, AI & VR (medical)

M.S. in Intelligent Systems, University of Texas at Dallas *May 2025*
Graduate focus across computer networks, immersive media, and ML systems.

Research Experience

Unsupervised Bayesian Surprise Detection in Spatial Audio (Published) — *CVAE–LSTM on FOA; TEA-aware* *IMX Workshops '24*

- Built an **unsupervised CVAE–LSTM** pipeline for **Bayesian surprise** in FOA (4-ch B-format), modeling **Temporal Evolution of Attention (TEA)**.
- Pipeline: 16 kHz, STFT + pre-emphasis + Hamming → log mel-spectrogram; **Audio Energy Maps (AEM)** via circular harmonics for source localization.
- Surprise metric: weighted NELL (reconstruction) + KL between LSTM-predicted next-latent and CVAE posterior; scenario achieved **F1=100%**.
- *Artifacts:* plan to release training/eval scripts, Docker env, and a small FOA dataset sample.

Spatiotemporal Attentive Entropy (SAE) — *SO(3)-invariant cross-entropy on S^2; O(1) streaming estimator*

- Formulated **spherical, geometry-correct** cross-entropy between vMF attention and prediction for real-time **viewport forecasting**; lightweight LSTM forecaster.
- Set new accuracy–latency Pareto frontier; validated **calibration** (ECE-sphere, NLL) and robustness to sampling-rate shifts and synthetic dropouts.
- *Metrics (to confirm):* ∼**X%** great-circle MAE reduction at ≤**Y ms** E2E latency (CPU/GPU prototypes); reproducibility via deterministic splits + Docker/LUTs.

MARC: Meta-Cognitively Adaptive Risk Control for 360° ABR — *Video + Spatial Audio; UKF entropy → CVaR*

- Coupled UKF posterior entropy to **CVaR** to hedge **tail-QoE**; fused video and **spatial audio** within the control loop.
- Vectorized coarse-to-fine candidate search; entropy-gated saliency; optional per-tile risk micro-actuation.
- *Metrics (to confirm):* ∼**A%** 1st-pct QoE improvement; ∼**B%** bandwidth efficiency; decision latency ≤**C ms**.

Teaching & Mentoring

Teaching Assistant — Computer Networks University of Texas at Dallas *Fall 2025*

- Delivered **multiple lectures**; ran labs/recitations on RTP/RTCP/WebRTC; coached **latency/jitter** measurement & **QoE** evaluation for ∼**50** students.

Teaching Assistant — Advanced Operating Systems University of Texas at Dallas *Fall 2024*

- Mentored distributed systems projects (synchronization, kernel subsystems); profiling & debugging at scale.

Teaching Assistant — Programming Language Paradigms University of Texas at Dallas *Spring 2024; Fall 2023*

– Reinforced functional/OO paradigms, type systems, and runtime models; office hours and assessments.

Supervisor — Undergraduate Research Project Course University of Texas at Dallas *Ongoing*

– Supervise undergraduate research; mentor on experimental design, implementation, evaluation, and technical writing.

Skills

ML/AI: PyTorch, TensorFlow, VAEs/CVAEs, LSTM/Seq2Seq, Transformers; calibration (ECE/NLL); CVaR risk
Audio/Signal Proc.: Ambisonics (FOA/B-format), STFT, mel-spectrogram, AEM, circular harmonics
Realtime/Streaming: WebRTC, RTP/RTCP, HLS/DASH, MPEG-OMAF (360°), ABR, QoE (VMAF/SSIM/PSNR)
Systems/Perf: C/C++, Python, Linux, concurrency, Docker, Git; profiling/latency (p95/p99)
Codecs/Accel: H.264/AVC, H.265/HEVC, AV1, CUDA/NVENC (foundational); FFmpeg, GStreamer

Publications

– **Nik Khah, A.** *Spatiotemporal Attentive Entropy: A Geometry-Correct Cross-Entropy on S^2 for Calibrated, Low-Latency Viewport Prediction.* Manuscript submitted, 2025.

– **Nik Khah, A.** *Self-Aware Streaming: Using Model Entropy to Control Tail-Risk in 360° ABR.* Manuscript submitted, 2025.

– **Nik Khah, A.**, Htun, C., Prakash, R. *Unsupervised Bayesian Surprise Detection in Spatial Audio with Convolutional Variational Autoencoder and LSTM Model. IMX Workshops '24.* DOI: 10.1145/3672406.3672422

Selected Open Source

– Spatial-audio surprise detection — planned artifact release; GitHub.

– SAE reference implementation — to align with review cycle.

– MARC (risk-aware ABR) controller — uncertainty-coupled CVaR policy with audio-visual fusion.

Selected Coursework

Advanced Machine Learning, Deep Learning, Computer Vision, Natural Language Processing, Probabilistic Modeling, **Computer Networks**