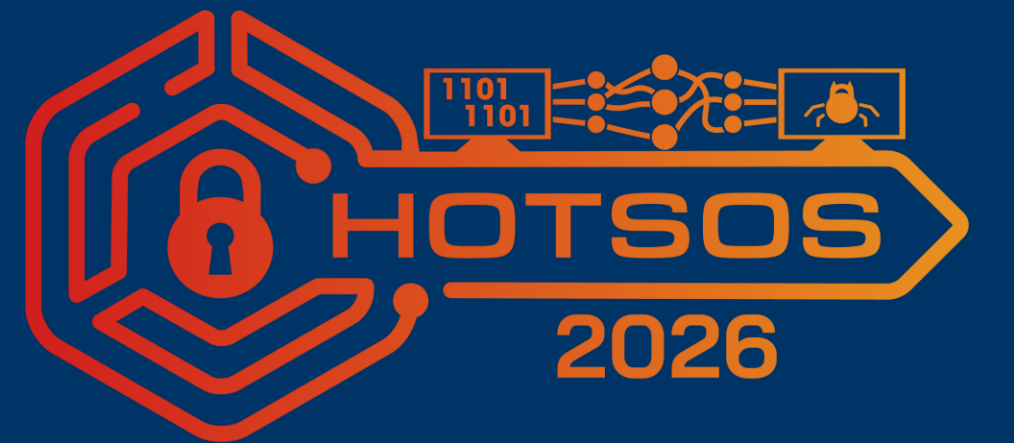


Towards Practical Edge Security: Energy-Efficient RF Anomaly Detection on the NVIDIA Jetson Orin Nano

Nicholas D. Redmond, Mohd Hasan Ali, Dipankar Dasgupta, Myounggyu Won

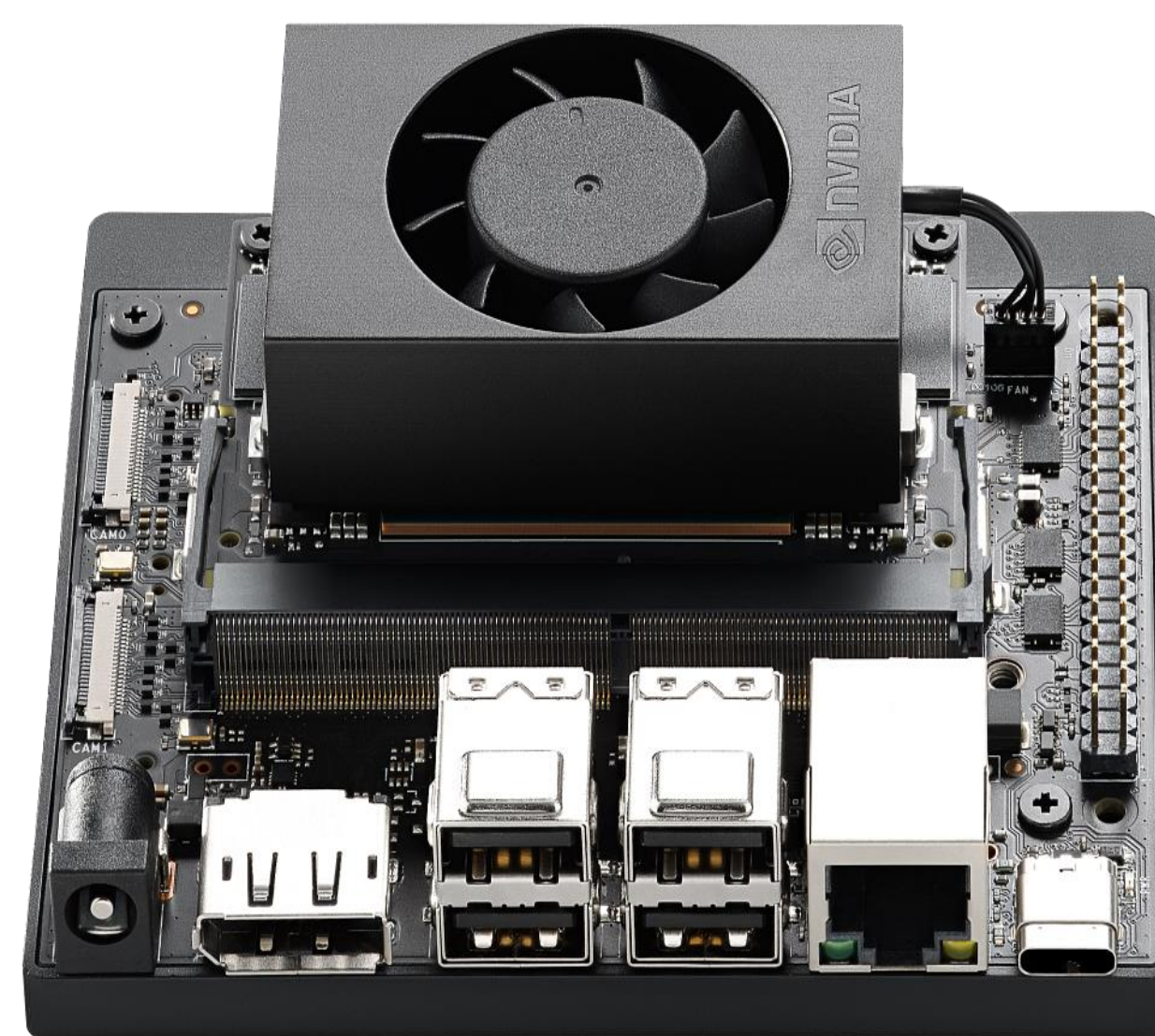
University of Memphis – Department of Engineering, Department of Computer Science



Motivation

- Radio Frequency (RF) communication is **critical for infrastructure**: healthcare, transportation, government services.
- Existing ML models for RF security often **depend on cloud computing**, causing latency and privacy issues.
- **AI edge computing** enables deploying models **directly on edge devices**, reducing these concerns.
- This study provides the **first benchmark of ML models for RF anomaly detection** on the **NVIDIA Jetson Orin Nano** platform.

Hardware



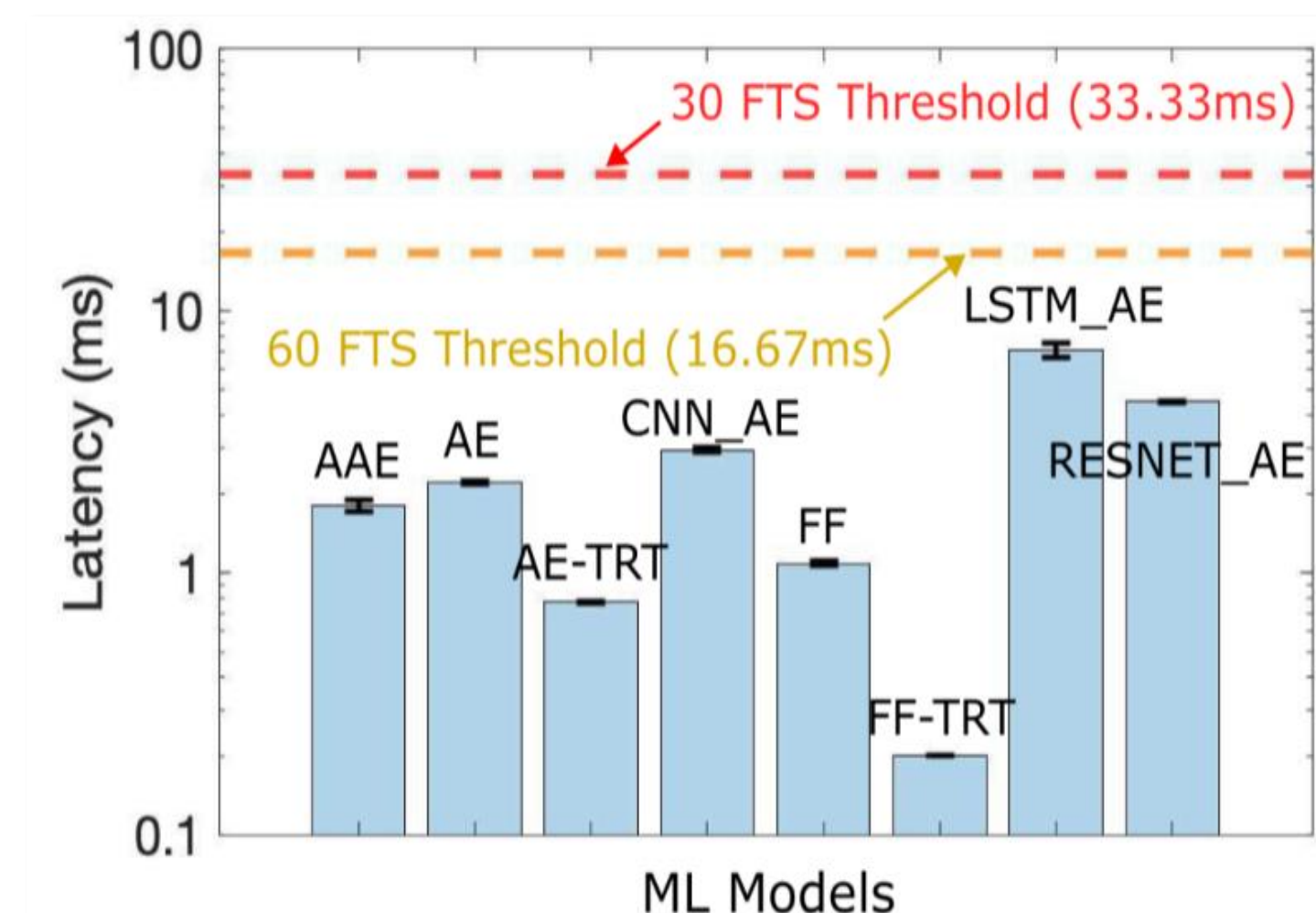
- **GPU**: Ampere architecture
- **CPU**: Arm Cortex-A78AE cluster
- **Memory**: LPDDR5
- **AI performance**: Up to 40 TOPS (INT8 precision)
- Developer Kit features: High-speed I/O, storage expansion, multimedia support
- Practical platform for deploying and benchmarking ML workloads

Models

- Autoencoder (AE)
- Adversarial Autoencoder (AAE)
- Convolutional Neural Network–based Autoencoder (CNN_AE)
- Long Short-Term Memory–based Autoencoder (LSTM-AE)
- Residual Network–based Autoencoder (ResNet-AE)

These models match SoTa time-series and spectrum-based anomaly detection literature.

Latency & Throughput



- **All models <10 ms** (AE 2.2 ms; TensorRT AE 0.76 ms; LSTM-AE 7.1 ms). **Real-time thresholds satisfied.**

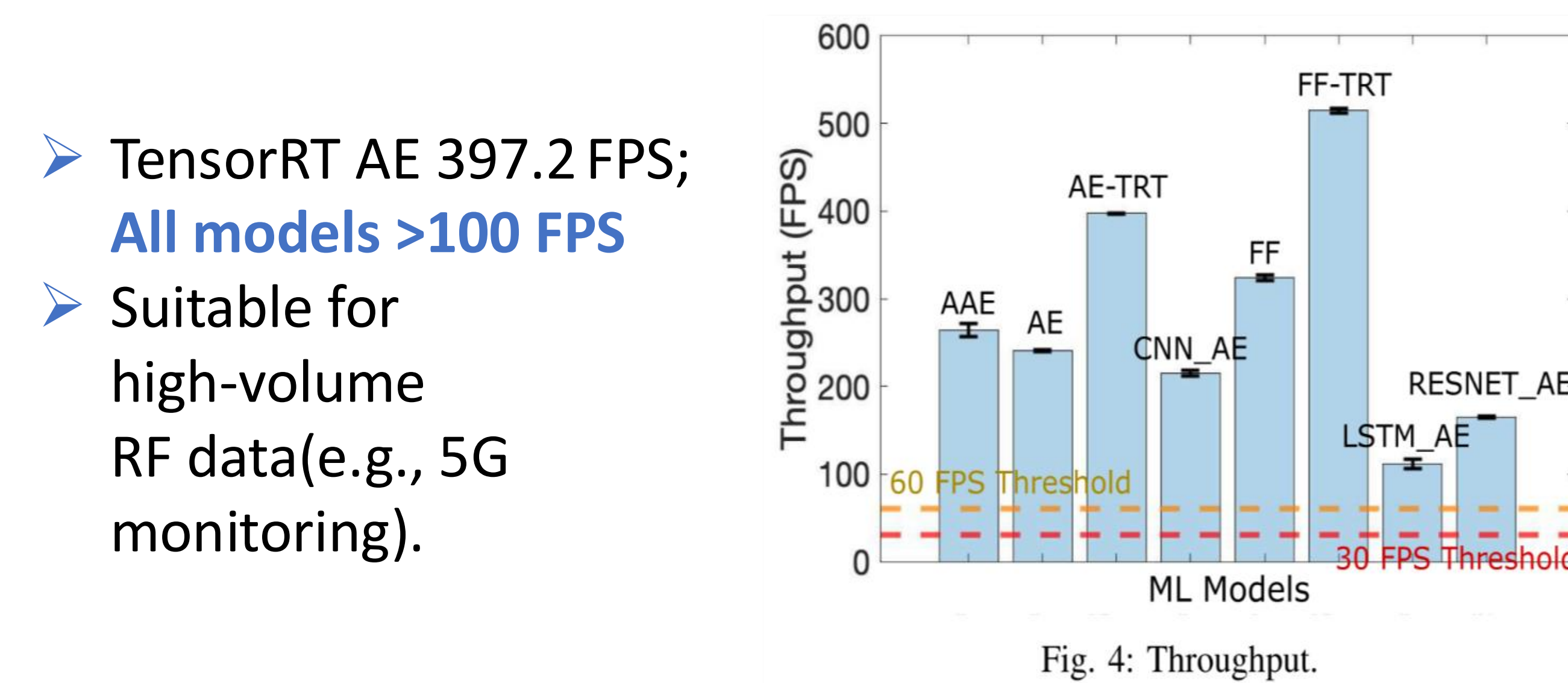
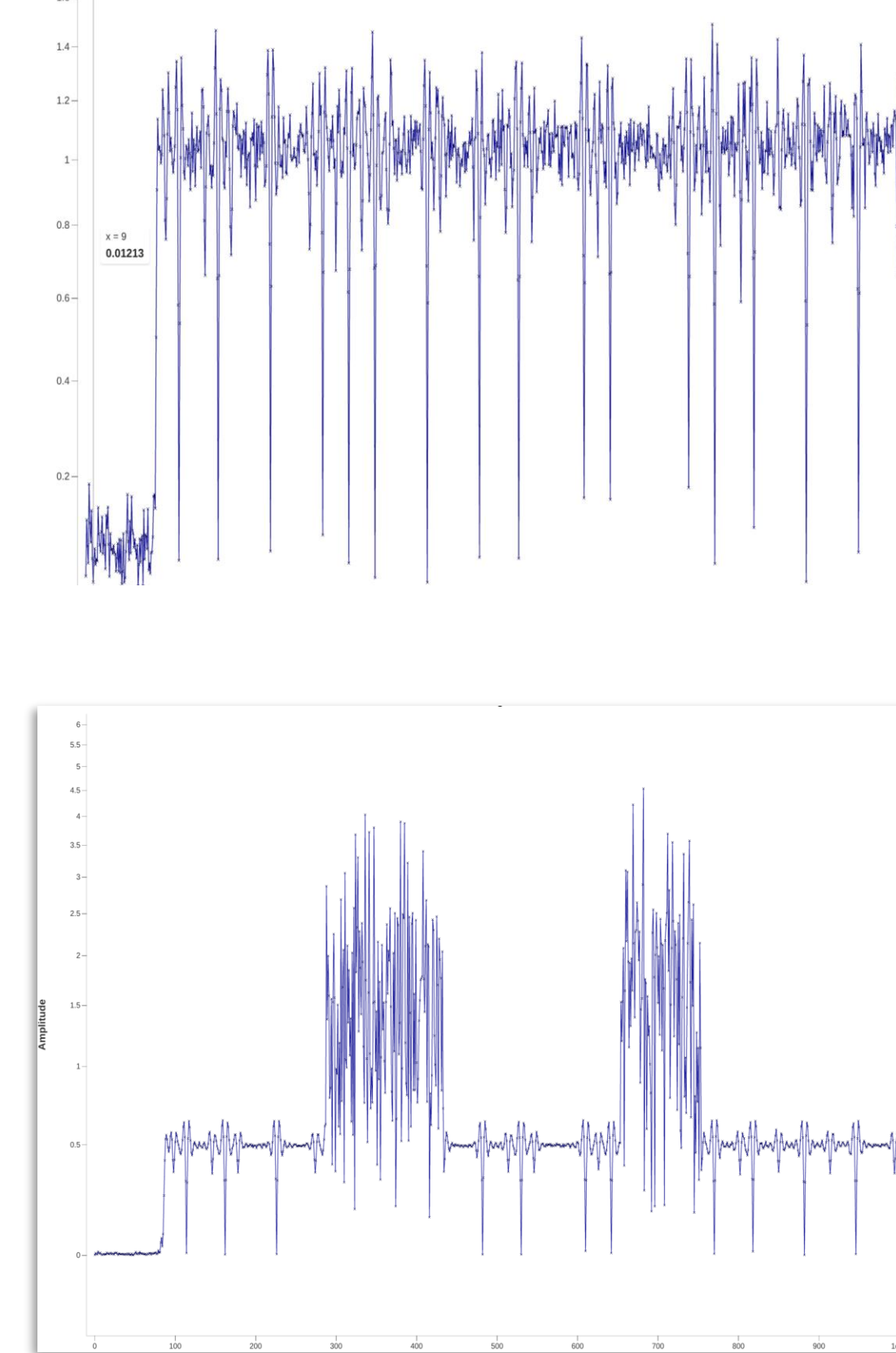


Fig. 4: Throughput.

- TensorRT AE 397.2 FPS; **All models >100 FPS**
- Suitable for high-volume RF data(e.g., 5G monitoring).

Dataset

- **1,024 complex I/Q** samples per record
- **Sampling rate**: 1 MHz → ≈1.024 ms duration
- **Baseband generation**: 0 Hz carrier, unit power
- **Preprocessing**: Normalization for model input
- **Why baseband?** Improves computational efficiency in simulation workflows



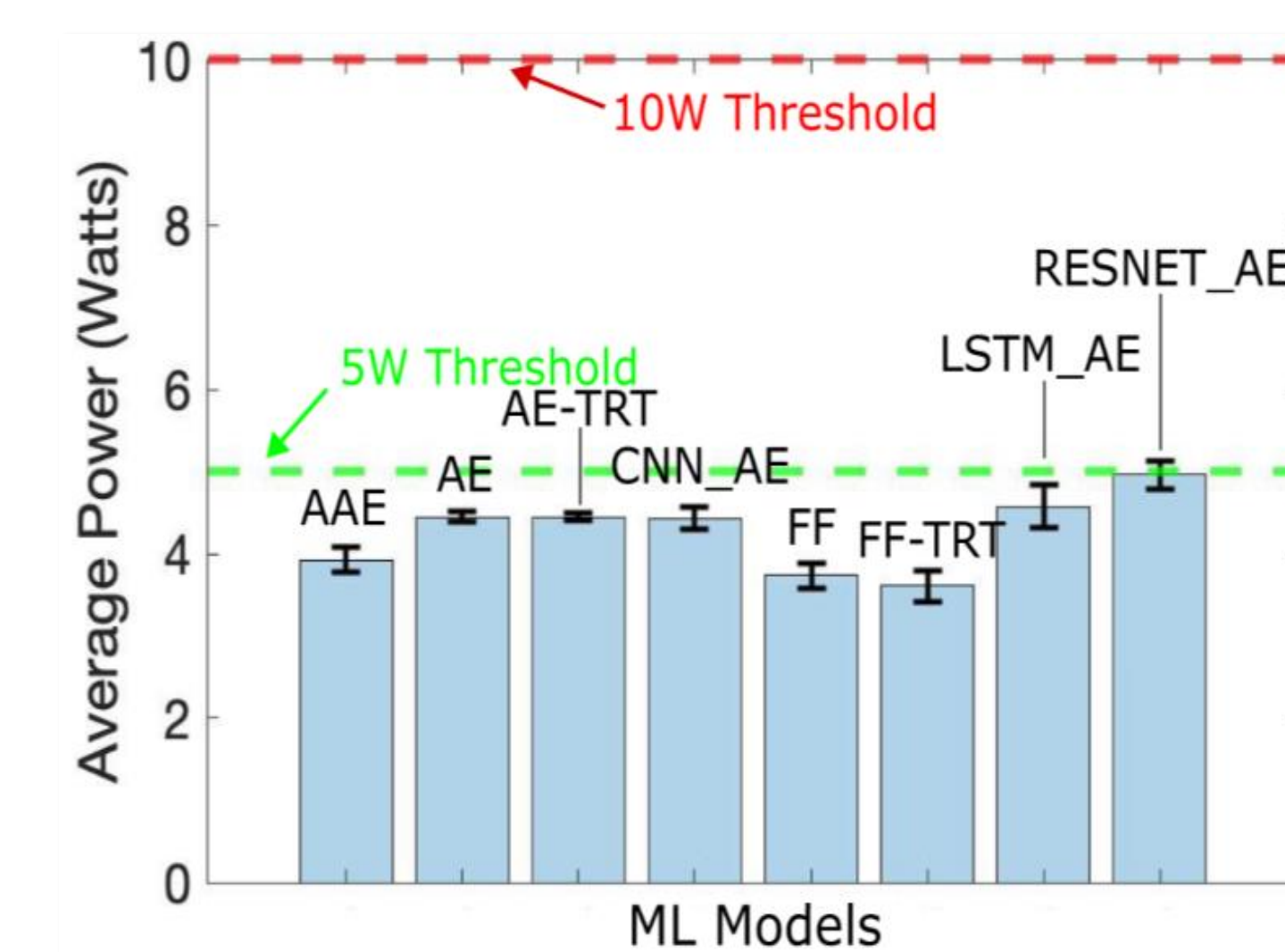
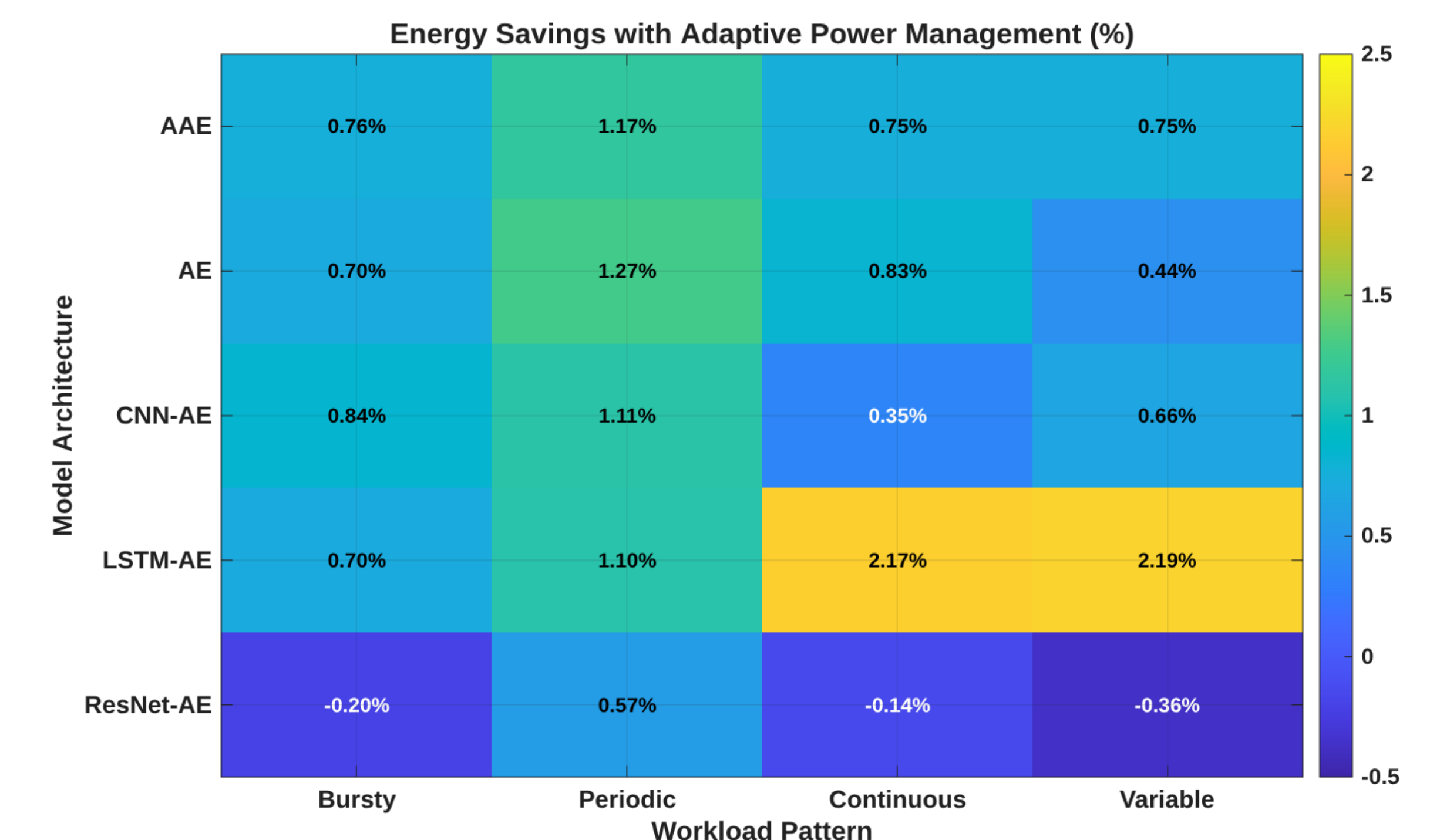
Comparison of Related Works

An overhead **performance comparison** against recent ML works benchmarked on the Jetson Orin Nano

TABLE I: Performance Comparison of Jetson Orin Nano Across Selected Studies

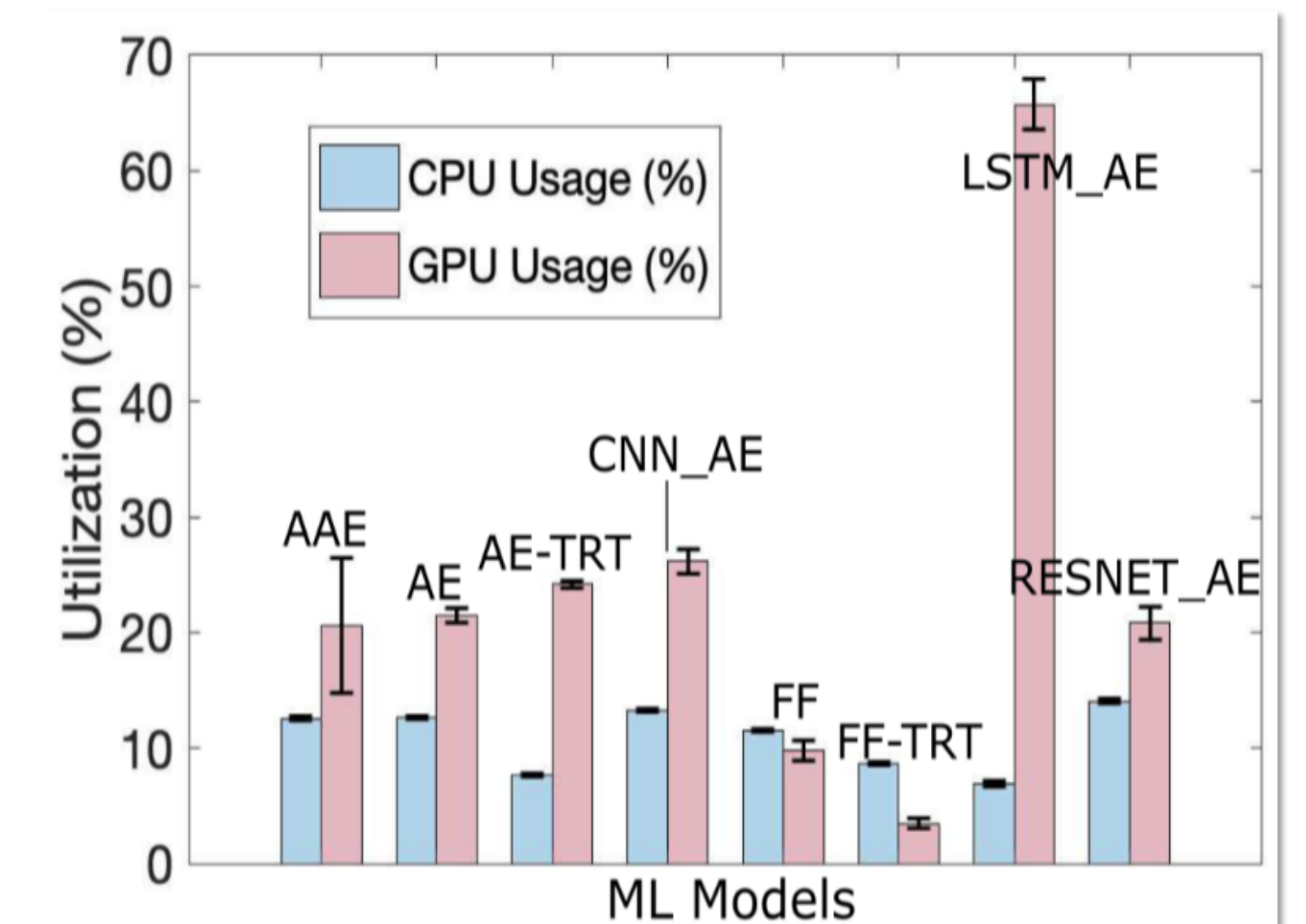
Authors	Field of Study	Model / Task	Quantization / Precision	FPS / Inference Time	Power / Energy	Latency
Rey et al. (2025)[10]	Computer Vision	YOLOv8s	FP32	27.0 FPS	8.79 W / 0.375 J/infer	RTT = 73.18 ms, Proc = 47.18 ms
		YOLOv8s	FP16	37.9 FPS	7.84 W / 0.302 J/infer	RTT = 49.44 ms, Proc = 28.90 ms
		YOLOv8s	INT8	41.2 FPS	7.26 W / 0.223 J/infer	RTT = 44.27 ms, Proc = 20.62 ms
		YOLOv8n	FP16	40.0 FPS	7.42 W / 0.208 J/infer	RTT = 49.70 ms, Proc = 23.81 ms
Chakravarty et al. (2024)[11]	Speech Recognition	Conformer ASR	FP16 (TensorRT)	54.8 FPS	9.2 W avg	18.3 ms processing
		QuartzNet 15x5	FP32	32.1 FPS	8.5 W avg	29.7 ms processing
		CitriNet S12	INT8 (PTQ)	61.3 FPS	7.1 W avg	16.3 ms processing
		AE	FP32	244.6 FPS	4.5 W avg	2.2 ms
Redmond et al. (2025)	RF Anomaly Detection	TensorRT	FP32	397.2 FPS	4.8 W avg	0.76 ms
		AAE	FP32	276.4 FPS	4.1 W avg	1.9 ms
		CNN-AE	FP32	221.4 FPS	4.7 W avg	2.5 ms
		FF	FP32	330.8 FPS	4.4 W avg	1.7 ms
		TensorRT	FP32	523.9 FPS	3.9 W avg	0.51 ms
		LSTM-AE	FP32	111.4 FPS	6.8 W avg	7.1 ms
		ResNet-AE	FP32	177.3 FPS	7.7 W avg	4.8 ms
		AE-TRT	FP32	397.2 FPS	4.5 W avg	0.76 ms

Energy Efficiency



- **Adversarial Autoencoder (AAE)** showed **highest energy efficiency**.
- **Complex models** (LSTM-AE, ResNet-AE) exhibited **lower energy efficiency**.

- **GPU utilization** was moderate (20–25% typical; LSTM-AE up to 65%), CPU remained low (7–14%), enabling concurrent edge tasks.



Conclusion

- Benchmarking ML models for RF anomaly detection on edge devices, remains a gap.
- This study **addresses that gap** with a comprehensive benchmark analysis.
- Evaluated **multiple ML models** deployed on NVIDIA Jetson Orin Nano.
- Results show the device is **well equipped for RF anomaly detection**, achieving strong performance, across key metrics.