

Energy-Aware Efficient Federated Learning over Mobile AI Systems

2025 CSR PI
MEETING



Hao Wang | Stevens Institute of Technology | 2534286

Miao Pan | University of Houston | 2403249

Seung-Jong Park | Missouri University of Science and Technology | 2403248

STEVENS
INSTITUTE OF TECHNOLOGY™

IntelliSys Lab

Project Overview

Existing research efforts largely ignore the fact that *FL applications are shifting from wall-plug-powered AI devices to battery-powered mobile AI systems (e.g., smartphones, tablets, laptops, AR/VR glasses/helmets, and various IoT devices)* due to their integral roles in daily life, while battery energy awareness plays a crucial role in efficient FL training over mobile AI systems.

- 1) How can energy awareness be effectively incorporated across diverse DNN models and heterogeneous mobile devices?
- 2) With improved energy awareness, how can the communication and computation strategies of each participating mobile device be optimized?
- 3) How can we select participant devices that not only have adequate residual energy but also can contribute efficiently to FL?

Results/Findings

Ampere: A Generic Energy Estimation Approach for On-Device Training, ACM SIGMETRICS 2025 Workshop: AI for Crossroad: Systems, Energy, and Applications 2025

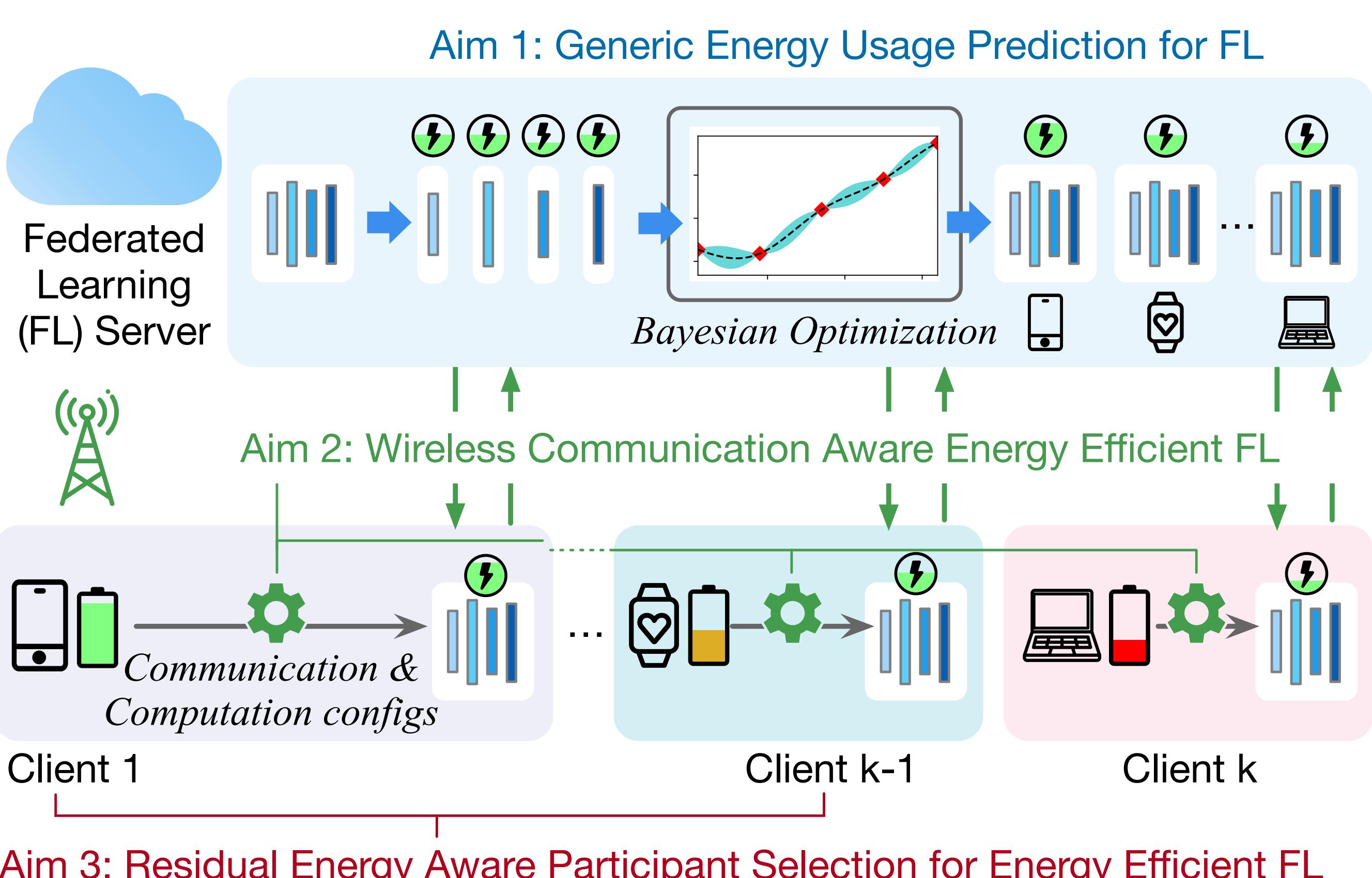
WHALE-FL: Wireless and Heterogeneity Aware Latency Efficient Federated Learning over Mobile Devices via Adaptive Subnetwork Scheduling, AAAI Conference on Artificial Intelligence (AAAI) 2025

pFedGPT: Hierarchically Optimizing LoRA Aggregation Weights for Personalized Federated GPT Models, accepted by Empirical Methods in Natural Language Processing (EMNLP) 2025

An Efficient On-Device Federated Learning System through the Interplay of Client Selection and Batch Size with Watermarked Data, Accepted by IEEE Transactions on Mobile Computing (TMC)

Research Objectives

Exploring the *efficiency, quality, and robustness* of Federated Learning (FL) systems from the perspective of energy with three synergistic aims.

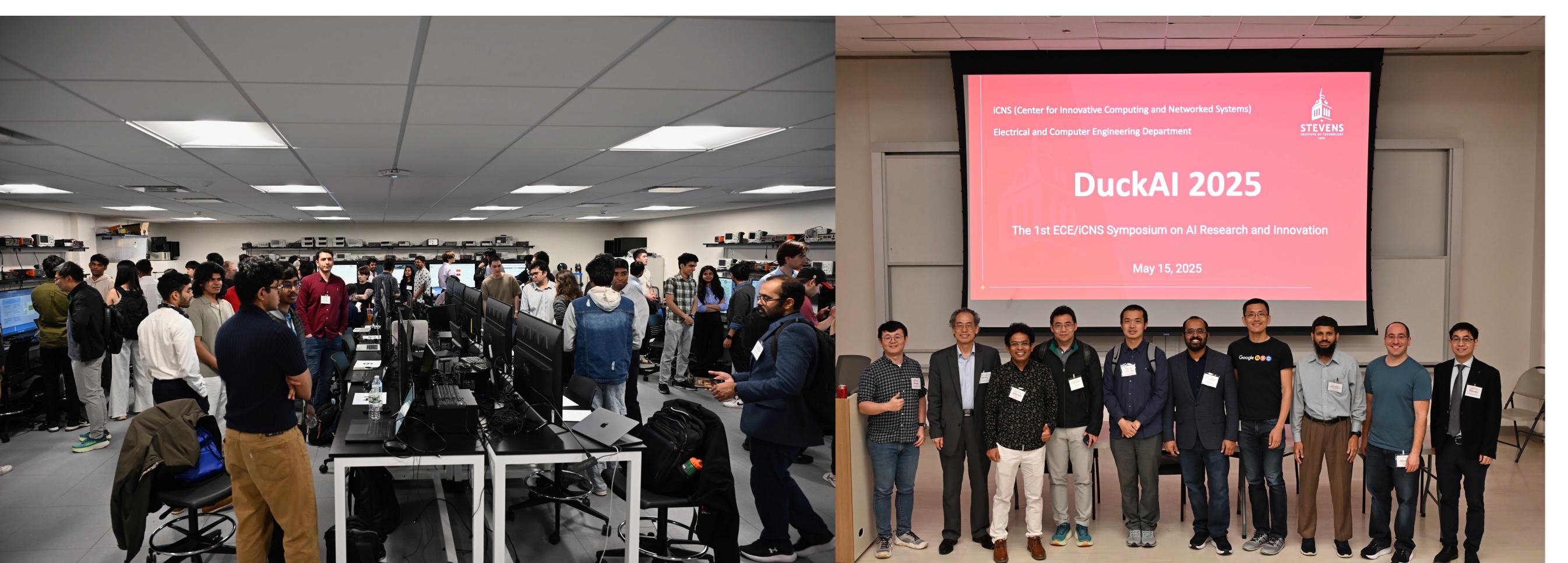


Broader Impacts

Host ACM SIGMETRICS 2025 Workshop:
AI for Crossroad: Systems, Energy, and Applications

Stevens DuckAI Symposium 2025:

- 4 talks + poster session (25 posters) + awards
- Bringing together industry leaders (e.g., Google, IBM, hedge funds), faculty, and over 150 students
- Topics cover efficient AI, AI security, AI for gaming, AI for healthcare, and Fintech



Methods/Approach

Aim 1: Generic Energy Usage Prediction for FL over Mobile AI Systems.

Aim 2: Wireless Communication Aware Energy Efficient FL over Mobile AI Systems.

Aim 3: Residual Energy Aware Participant Selection for Energy Efficient FL over Mobile AI Systems.

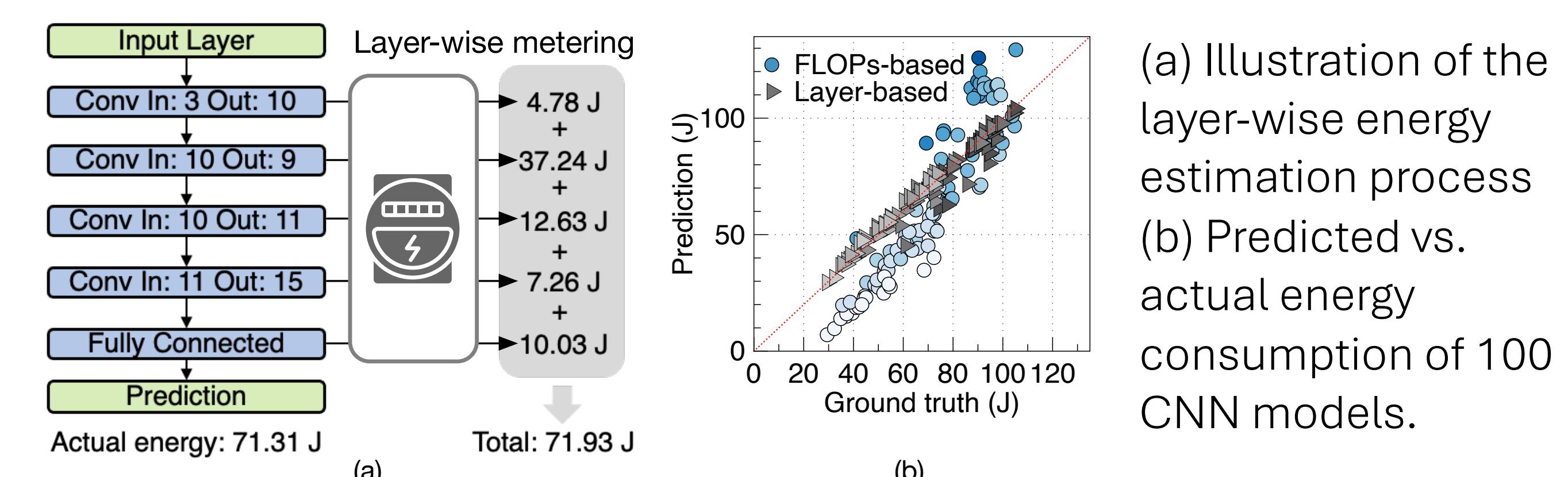


Table 1: Energy efficiency (EE) and accuracy.

Methods	CNN@MNIST		ResNet@CIFAR10	
	EE	Acc.	EE	Acc.
FedAvg [1]	1x	98.37	1x	89.41
LUPA [58]	0.81x	98.60	0.90x	90.12
FEDL [60]	1.17x	98.33	1.15x	87.25
SparFL [49]	1.31x	98.30	1.76x	89.32
E-AdaH	1.40x	98.41	2.12x	89.82

Future Work

From Model's Perspective:

- Exploring LLM serving/fine-tuning over Mobile AI systems, e.g., multi-modal LLMs, multi-agent LLMs, MoE-based LLMs etc.
- Building benchmarks, leaderboard, and datasets for AI models' fine-grained energy consumption and efficiency

From System's Perspective:

- Building a prototype mobile FL testbed for energy-aware efficiency with open interfaces, allowing new hardware/software to be integrated
- Developing a library on enabling energy-awareness and energy-aware scheduling on mobile AI systems

Broader Impacts:

- Organizing workshops and tutorials disseminating our studies and solutions

STEVENS
INSTITUTE OF TECHNOLOGY™

UNIVERSITY OF
HOUSTON

MISSOURI S&T™