

# ECE Distinguished Lecture

## AI Models for Edge Computing: Hardware-aware Optimizations for Efficiency



### Hai Li

Marie Foote Reel E'46 Distinguished Professor and  
Department Chair of the Electrical and Computer  
Engineering Department at  
Duke University

**Friday, October 11**  
10:30 A.M. • Potter 234 Fu Rm

Zoom Link: <https://purdue-edu.zoom.us/j/99210875121>  
Meeting ID: 992 1087 5121

**Abstract:** As artificial intelligence (AI) transforms various industries, state-of-the-art models have exploded in size and capability. The growth in AI model complexity is rapidly outstripping hardware evolution, making the deployment of these models on edge devices remain challenging. To enable advanced AI locally, models must be optimized for fitting into the hardware constraints. In this presentation, we will first discuss how computing hardware designs impact the effectiveness of commonly used AI model optimizations for efficiency, including techniques like quantization and pruning. Additionally, we will present several methods, such as hardware-aware quantization and structured pruning, to demonstrate the significance of software/hardware co-design. We will also demonstrate how these methods can be understood via a straightforward theoretical framework, facilitating their seamless integration in practical applications and their straightforward extension to distributed edge computing. At the conclusion of our presentation, we will share our insights and vision for achieving efficient and robust AI at the edge.

**Bio:** Hai “Helen” Li is the Marie Foote Reel E’46 Distinguished Professor and Department Chair of the Electrical and Computer Engineering Department at Duke University. She received her B.S. and M.S. from Tsinghua University and Ph.D. from Purdue University. Her research interests include neuromorphic systems, machine learning acceleration and trustworthy AI, conventional and emerging memory design and architecture, and software and hardware co-design. Dr. Li served/serves as the Associate Editor-in-Chief and Associate Editor for multiple IEEE and ACM journals. She was the General Chair or Technical Program Chair of multiple IEEE/ACM conferences and the Technical Program Committee member of over 30 international conference series. Dr. Li is a Distinguished Lecturer of the IEEE CAS Society (2018-2019) and a distinguished speaker of ACM (2017-2020). Dr. Li is a recipient of the NSF Career Award, DARPA Young Faculty Award, TUM-IAS Hans Fischer Fellowship from Germany, ELATE Fellowship, Ten Year Retrospective Influential Paper Award from ICCAD’2023, nine best paper awards, and another nine best paper nominations. Dr. Li is a fellow of ACM and IEEE.

**Host:** Maggie Zhu ~ [zhu0@purdue.edu](mailto:zhu0@purdue.edu)