

4th International Workshop on Embedded and Mobile Deep Learning

WORKSHOP co-located with ACM MobiCom 2020
London, UK – Monday, 21 SEPTEMBER 2020
<https://emdl20.github.io>



In recent years, breakthroughs from the field of deep learning have transformed how sensor data (e.g., images, audio, and even accelerometers and GPS) can be interpreted to extract the high-level information needed by bleeding-edge sensor-driven systems like smartphone apps and wearable devices. Today, the state-of-the-art in computational models that, for example, recognize a face, track user emotions, or monitor physical activities are increasingly based on deep learning principles and algorithms. Unfortunately, deep models typically exert severe demands on local device resources and this conventionally limits their adoption within mobile and embedded platforms. As a result, in far too many cases existing systems process sensor data with machine learning methods that have been superseded by deep learning years ago.

Because the robustness and quality of sensory perception and reasoning is so critical to mobile computing, it is critical for this community to begin the careful study of two core technical questions. First, how should deep learning principles and algorithms be applied to sensor inference problems that are central to this class of computing? This includes a combination of applications of learning some of which are familiar to other domains (such as the processing image and audio), in addition to those more uniquely tied to wearable and mobile systems (e.g., activity recognition). Second, what is required for current -- and future -- deep learning innovations to be either simplified or efficiently integrated into a variety of mobile resource-constrained systems? At heart, this MobiCom 2020 co-located workshop aims to consider these two broad themes; more specific topics of interest, include, but are not limited to:

- Compression of Deep Model Architectures
- Neural-based Approaches for Modeling User Activities and Behavior
- Quantized and Low-precision Neural Networks (including Binary Networks)
- Resource-efficient Federated Learning
- Mobile Vision/AR/VR supported by Convolutional and Deep Networks
- Optimizing Commodity Processors (GPUs, DSPs etc.) for Deep Models
- Audio Analysis and Understanding through Recurrent and Deep Architectures
- Hardware Accelerators for Deep Neural Networks
- Distributed Deep Model Training Approaches
- Applications of Deep Neural Networks with Real-time Requirements
- Deep Models of Speech and Dialog Interaction on Mobile Devices
- Partitioned Networks for Improved Cloud- and Processor-Offloading
- Operating System Support for Resource Management at Inference-time

Workshop Organizers

PC CHAIRS

Stylianios I. Venieris (*Samsung AI Center, Cambridge*)
Nicholas D. Lane (*Samsung AI Center, Univ. of Oxford*)
Marco Gruteser (*Google, Rutdgers University*)
Koen Helwegen (*Plumerai*)

FULL PAPER SUBMISSIONS

Solicited submissions include both full technical workshop papers and white position papers. Maximum length of such submissions is 6 pages, and if accepted they will be published by ACM and appear in the ACM Digital Library.

Submission Deadline: May 15th, 2020 – 11:59pm AOE

WORK-IN-PROGRESS AND DEMO SUBMISSIONS

Abstracts describing work-in-progress and demonstrations are also welcome and warmly encouraged. Submissions are limited to 2 pages, and if accepted included in the program as a short oral presentation – but will only be published on the workshop website (not the ACM DL). **Deadlines for this informal track remain open** even past the early registration deadline of MobiCom 2020; author notifications will be rolling (i.e., max. of 4 days after submission) to enable early authors to take advantage of available discounts.

