

CALL FOR PAPERS

IEEE Journal of Selected Topics in Signal Processing Special Issue on Compact Deep Neural Networks with Industrial Applications

Artificial neural networks have been adopted for a broad range of tasks in areas like multimedia analysis and processing, media coding, data analytics, etc. Their recent success is based on the feasibility of processing much larger and complex deep neural networks (DNNs) than in the past, and the availability of large-scale training data sets. As a consequence, the large memory footprint of trained neural networks and the high computational complexity of performing inference cannot be neglected. Many applications require the deployment of a particular trained network instance, potentially to a larger number of devices, which may have limitations in terms of processing power and memory e.g., for mobile devices or Internet of Things (IoT) devices. For such applications, compact representations of neural networks are of increasing relevance.

This special issue aims to feature recent work related to techniques and applications of compact and efficient neural network representations. It is expected that these works will be of interest to both academic researchers and industrial practitioners, in the fields of machine learning, computer vision and pattern recognition, media data processing, as well as fields such as AI hardware design etc. In spite of active research in the area, there are still open questions to be clarified concerning, for example, how to train neural networks with optimal performance while achieving compact representations, and how to achieve representations that do not only allow for compact transmission, but also for efficient inference. This special issue therefore solicits original and innovative works to address these open questions in, but not limited to, following topics:

- Sparsification, binarization, quantization, pruning, thresholding and coding of neural networks
- Efficient computation and acceleration of deep convolutional neural networks
- Deep neural network computation for low power consumption applications
- Exchange formats and industrial standardization of compact & efficient neural networks
- Applications e.g. video & media compression methods using compressed DNNs
- Performance evaluation and benchmarking of compressed DNNs

Prospective authors should follow the instructions given on the IEEE JSTSP webpage:

<https://signalprocessingsociety.org/publications-resources/ieee-journal-selected-topics-signal-processing>,

and submit their manuscript through the web submission system at: <https://mc.manuscriptcentral.com/jstsp-ieee>.

Dates: **(extended)**

Submission deadline: ~~01-Jun-2019~~ 01-July-2019

First Review: ~~01-Aug-2019~~ 1-Sept-2019

Revisions due: ~~01-Oct-2019~~ 15-Oct-2019

Second Review: 15-Nov-2019

Final Manuscripts: 10-Jan-2020

Publication: March 2020

Guest Editors:

Diana Marculescu, Carnegie Mellon University, USA

Lixin Fan, JD.COM, Silicon Valley Labs, USA (Lead GE)

Werner Bailer, Joanneum Research, Austria

Yurong Chen, Intel Labs China, China