Combining Compressive Sensing with Deep Learning for Efficient Ubiquitous Computing Inference

ALINA MACHIDON, Faculty of Computer and Information Science, University of Ljubljana, Slovenia
VELJKO PEJOVIĆ, Faculty of Computer and Information Science, University of Ljubljana, Slovenia and Computer Science Department, Lancaster University, UK

The ability to sense the environment and harness machine learning to extract knowledge from the sensed data represents the foundation of awareness and intelligence that we expect from ubiquitous computing. However, both sensing and inference require significant energy resources. This, having in mind that our devices are expected to handle numerous other tasks as well, often hampers the realisation of the full potential of mobile sensing.

Approximate mobile computing, a philosophy that drives our research, states that in a range of situations mobile computers can take fewer samples and conduct less computation, yet still deliver satisfactory results. The benefits of such a way of designing mobile systems is evident as we hit the physical limits of hardware scaling, and our devices' computational and energy resources cannot support the increasingly complex tasks we wish them to complete.

Compressive sensing (CS) is a mathematical tool that enables successful reconstruction of signals sampled at rates far lower than the Nyquist rate, in case these signals are characterised by certain mathematical properties. Since many real-world signals indeed satisfy these restrictions, CS has a tremendous potential to reduce the resource burden of mobile sensing. Realising this potential, however, is not straighforward. First, in mobile computing, signal reconstruction is seldom the goal. Instead, higher-level inferences are made from the collected samples, and the most informative samples with respect to these inferences, not reconstruction, should be attained. Second, the properties of the signal of interest vary with device mobility. From the CS standpoint, this means that the optimal sampling strategy is unlikely to remain the same over time.

In this work we address the above issues and develop a *compressive sensing – deep learning (CS-DL) pipeline for ubiquitous computing*. Our solution includes an optimal subsampling scheme for a given task, where the sparsest selection of the most informative samples is obtained. The solution also includes a sampling rate-adaptable neural network layer, which enables the execution of the downstream classification task with variable size inputs. Finally, the pipeline is trained to allow for a graceful degradation of performance, allowing the dynamic trade-off between the inference accuracy and resource usage (Figure 1).

The proposed pipeline builds upon our previous work[1], yet in this presentation we introduce a novel adaptation of the sampling layer that further minimises the amount of computation, and we also present recent results of sparse sampling and learning in various domains: sleep stage classification from EEG signals, keyword recognition from audio signals, and image classification, demonstrating the general applicability of our solution to different tasks and domains.

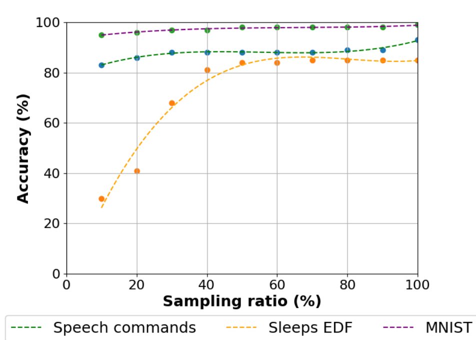

Fig. 1. Inference accuracy vs sampling (indirectly *resource usage*) trade-offs enabled by our CS-DL pipeline in three domains: speech command recognition, sleep stage detection, and image recognition on the MNIST dataset.

[1] A. Machidon and V. Pejović, Enabling Resource-Efficient Edge Intelligence with Compressive Sensing-Based Deep Learning, ACM Computing Frontiers, May 2022