BatchBALD: Efficient and Diverse Batch Acquisition for Deep Bayesian Active Learning

In this paper, we have been introduced to a new batch acquisition function, BatchBALD, for Deep Bayesian Active Learning and a greedy algorithm that selects good candidate batches compared to the intractable optimal solution. Applying BALD naively to a dataset that contains many replicated data points leads to poor performance. To address this, BatchBALD was introduced. This data-efficient active learning method acquires sets of high-dimensional image data, leading to improved data efficiency and reduced total run time. Acquisitions show increased diversity of data points and improved performance over BALD and other methods. While our approach comes with additional computational cost during purchase, BatchBALD can significantly reduce the number of data points that need to be labelled and the number of times the model has to be retrained, potentially saving considerable costs and filling an essential gap in practical Deep Bayesian Active Learning.