

The paper highlights the problem with traditional deep learning. The problem is the traditional models, being data hungry, rendering the deployment of deep learning in the real-world challenging for many tasks. One solution for this is using active learning(AL). In AL we iteratively acquire labels from an expert only for the most informative data points from a pool of available unlabelled data. After each acquisition step, the newly labelled points are added to the training set, and the model is retrained. This process is repeated until a suitable level of accuracy is achieved. The goal of AL is to minimise the amount of data that needs to be labelled. The informativeness of new points in AL is assessed by an acquisition function. The method being talked about in this paper is called BALD. The method employs the use of bayesian learning:

- Method uses an acquisition function, which estimates mutual information between the model predictions and model parameters.
- It captures how strongly the model predicts for the given data point and model parameters are coupled.

Paper focuses on BNNs rather than the traditional CNNs, because they scale well to high dimensional inputs. BNNs maintain distributions over their weights instead of point estimates.