[The power of ensembles in active learning]

- computationally costly → snapshot ensemble

<Methodology>

√ Pool-Based Active Learning

- U : Large unlabelled pool
- L : Small labelled set of data
- M : Model
- a(U,M) → s
  - a : acquisition function
  - s : points to be labeled, then added

regression : a(U,M) is based on the predictive variance of the model output

classification :
- softmax output vector of model as an input to acquisition function
- outputs of the last FC layer are used as feature vectors to calculate image similarities for a density-based approach.

< Approaches for Uncertainty estimation >

① Monte Carlo Dropout

$$p(y=c \mid x, D_{train}) = \frac{1}{T}\sum_{t=1}^{T} p(y=c \mid x, w_t)$$

② Deep ensembles

$$p(y=c \mid x, D_{train}) = \frac{1}{N}\sum_{i=1}^{N} p(y=c \mid x, w_{init})$$

< Approaches for Acquisition >

uncer-tainty based

① Highest entropy of predicted class entropy

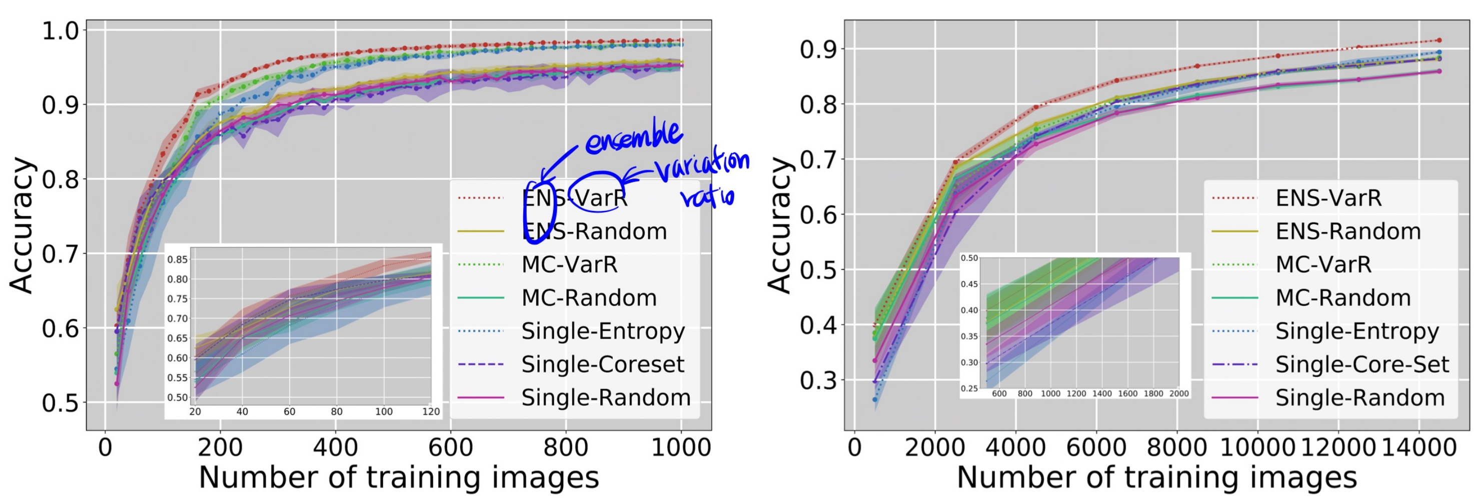
② BALD : the mutual information between data points and weights

③ Variation ratio : the proportion of predicted class labels that are not the modal class prediction.

best performance when married with deep ensemble

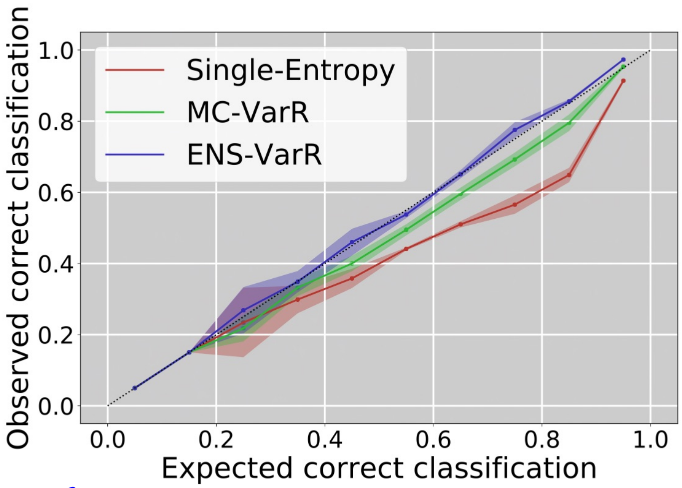
④ Core-set

⑤ REPR

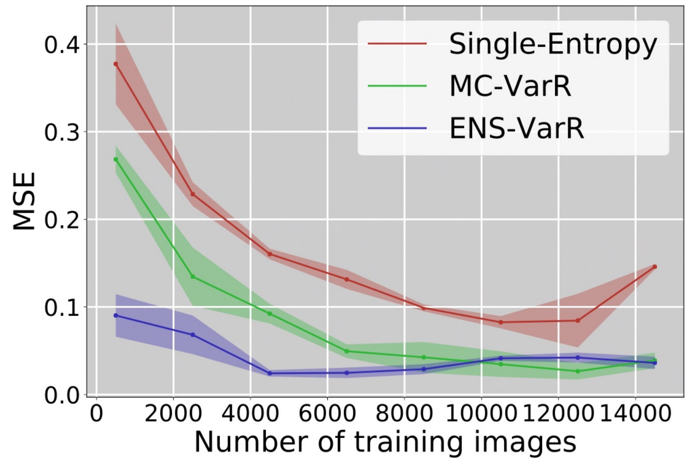


(a) MNIST on S-CNN

(b) CIFAR-10 on DenseNet



(a)



(b)

Uncertainty
Calibration