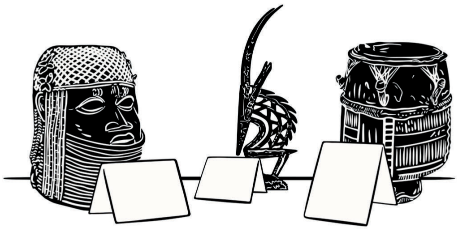
A STUDY OF ACQUISITION FUNCTIONS FOR MEDICAL IMAGING DEEP ACTIVE LEARNING

SCAN ME

Bonaventure F. P. Dossou
Lelapa AI, McGill University, Mila Quebec AI Institute, Masakhane NLP

ABSTRACT



In the real world, especially in the medical imaging context, data scarcity and limited labeled data are recurrent and frequent problems. This is very often a bottleneck to high-performance of recent Deep Learning approaches that are very data-hungry.

In this work, we show that active learning could be very effective in data scarcity situations, where obtaining labeled data is expensive. We compare several acquisition functions (**AF**) such as BALD, MeanSTD, and MaxEntropy on the ISIC 2016 Melanoma detection dataset, explore the impact of selecting either the most or least uncertain samples, and leverage the effect of acquired pool sizes on the performance of the model.

Our results on the Melanoma detection test set, demonstrate that uncertainty is useful to the Melanoma detection task and that it is more beneficial to select the most uncertain pool samples. These results suggest that active learning could be very useful for medical imaging tasks (in particular) and more generally in low-resource settings.

THE FORMULA

Maximum Entropy This acquisition function aims at selecting the data points which maximize the entropy of the model over each unlabelled data sample and known labels (classes). With the entropy defined as

$$\mathcal{H}[y|x, D_T] = -\sum_c p(y=c|x, D_T) logp(y=c|x, D_T)$$

where D_T is the training set, which is augmented by the set of newly acquired samples at each active learning round.

Mean Standard Deviation The Mean Standard Deviation (for short MeanSTD) is the most commonly used acquisition function. It leverages the variance of the model over classes, given an input x and the parameters w of the model. It is mathematically defined as follow:

$$\sigma_c = \sqrt{\mathcal{E}_{q(w)}[p(y=c|x,w)^2] - \mathcal{E}_{q(w)}[p(y=c|x,w)]^2}$$

$$\sigma_x = \frac{1}{C}\sum_c \sigma_c$$

As with the Maximum Entropy, in this scheme, we are also selecting points that maximize the MeanSTD.

BALD is based on mutual information. By definition, the mutual information denoted $\mathcal{I}$ between two random variables X, Y is telling us how much uncertainty we observe in X if we observe Y. BALD focuses on maximizing the mutual information between the predictions of the model and its posterior. BALD is mathematically defined as

$$\begin{aligned}\mathcal{I}(y, w|x, D_T) &= \mathcal{H}[y|x, D_T] - \mathcal{E}_{p(w|D_T)}[\mathcal{H}[y|x,w]] \\ &= -\sum_c p(y=c|x, D_T) logp(y=c|x, D_T) \\ &+ \mathcal{E}_{p(w|D_T)}[\sum_c p(y=c|x,w) logp(y=c|x,w)]\end{aligned}$$

where w are the parameters of the model. In other words, BALD chooses points that are expected to maximize the information gained about the parameters of the model w. These points are points on which the model is uncertain on average, but about which some parameters produce disagreeing predictions with high certainty.

CONCLUSION & FUTURE WORKS

We showed the efficiency of Active Learning in the context of the Melanoma Detection Task and is strongly viable in the context of low-resource settings (e.g. in medical domain)

1. Incorporating uncertainty with AFs in model training is useful and increases performance
2. Query size impacts performance of the model the test set
3. As future work, leverage EPIG (*Smith et al, 2023*) a recent **AF** which offers better performance than Bald

DATA & TASK DESCRIPTION

- The task is to detect if a given image is cancerous or not
- The ISIC 2016 is a dataset of dermoscopic created to enable automated diagnosis of melanoma

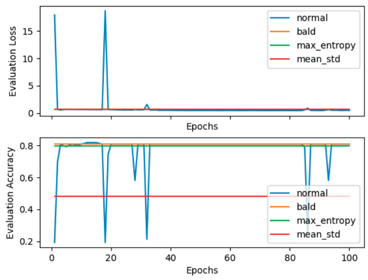
IMPORTANCE OF EPISTEMIC UNCERTAINTY

Method	Testing Loss	Testing Accuracy
normal	0.01538	0.8021
bald	0.0077	**0.8047**
max_entropy	0.0075	0.7784
mean_std	**0.0072**	0.4670

Table 1: Results on the testing set for both with and without uncertainty Bayesian CNN

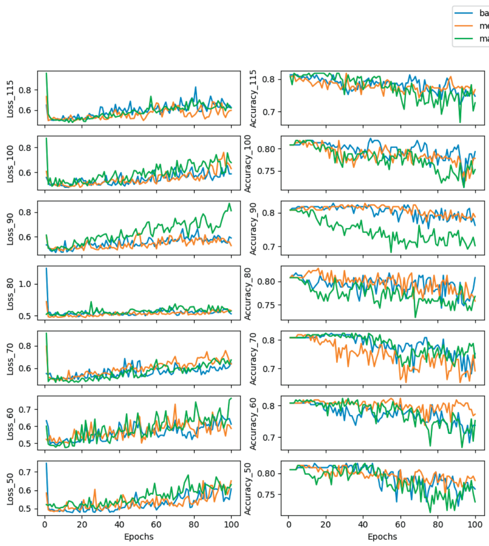
Method	Testing Loss	Testing Accuracy
bald	**0.009**	0.20
max_entropy	0.0099	0.5876
mean_std	0.0094	**0.8012**

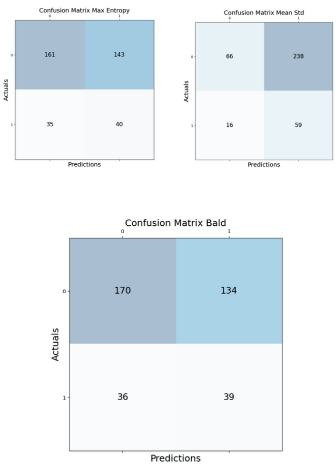
Table 2: Test Results of Bayesian CNNs using the Least Uncertain Samples



Our experiments have shown that **uncertainty is beneficial** to our Melanoma Detection task. Additionally, ablation studies revealed that, in the context of our Melanoma Detection task, **max_entropy** has been proven to be agnostic of the acquisition function, offering more **robustness and flexibility**.

IMPORTANCE OF QUERY SIZE





Method	Metric	Query=115	Query=100	Query=90	Query=80	Query=70	Query=60	Query=50
bald	loss	0.0177	0.0174	0.0183	**0.0169**	0.0208	0.0201	0.0185
bald	accuracy	0.8047	**0.8047**	0.7994	0.7942	0.7994	0.8021	0.8047
max_entropy	loss	0.0173	0.0235	0.0192	0.0202	**0.0157**	0.0167	0.0170
max_entropy	accuracy	0.7995	**0.8047**	0.8021	0.7863	0.8021	0.8021	0.7863
mean_std	loss	0.0185	**0.0164**	0.0191	0.0177	0.0164	0.0202	0.0186
mean_std	accuracy	0.7916	0.8021	0.8021	0.8047	0.7889	**0.8074**	0.7968

Table 3: Report of Testing Loss and Testing Accuracy on ISIC 2016 dataset as a function of the different query sizes. For each method and for each metric, the number in bold represents the best value achieved for a given query size.

1. For the loss metric, we can observe that generally, all acquisition functions are impacted by the query size: on average, a higher query size leads to a lower loss value
2. For accuracy, even though the scale doesn't change that much, On the accuracy scale, we can see that max entropy and mean std vary a lot compared to bald, which consequently offers more stability. This is because BALD avoided selecting noisy points: nearby images for which there exist multiple noisy labels of different classes (points for which the aleatoric uncertainty is large)