

Supplementary Materials for Class-Balanced Active Learning for Image Classification

G. Performance on Tiny ImageNet dataset

Fig. 7 illustrates the performance of class balanced (CB) methods and AL baselines. As can be seen, both Entropy-CB and BALD-CB outperform the corresponding baselines. Notably in Tiny ImageNet, Random sampling serve as a competitive baseline. Nevertheless the addition of class balancing made Entropy-CB superior in almost all active learning cycles across different imbalance factors.

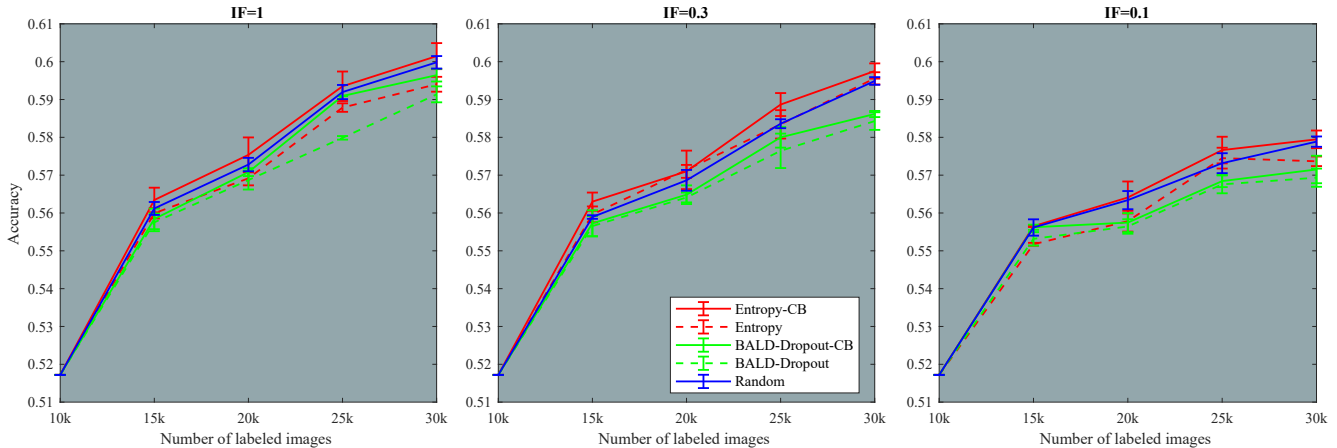


Figure 7. **Performance evaluation.** Results for active learning methods on Tiny ImageNet with different imbalance factors (IF).

H. Pseudo Label balancing

Fig. 8 presents the performance of another Entropy variation on CIFAR100 for comparison. Among them, "Entropy L1 Pseudo Label" benefits from "pseudo labels" defined as the most probable labels that the model assigns to unlabeled samples (the prediction of the model is then converted to a one-hot vector). This method utilizes the pseudo labels to balance the distribution of samples and select certain number of samples (specified by Ω in Eq.3) from each class with maximum entropy. The experiments show that Entropy-CB outperforms Entropy L1 Pseudo Label both in terms of active learning performance (see Fig. 8) and the ability of class balancing (see λ tuning in Section 4).

I. CoreSet performance

The performance of CoreSet on CIFAR10 and CIFAR100 is shown in Fig. 9 and Fig. 10 respectively. In our experiments CoreSet and KCenterGreedy-CB perform similarly on the balanced dataset (IF=1). However, when the dataset is imbalanced (IF=0.3 and IF=0.1) the performance of CoreSet degrades compared to KCenterGreedy-CB. As CoreSet is a MIP (Mixed Integer Programming) problem, our technique cannot be applied to this method.

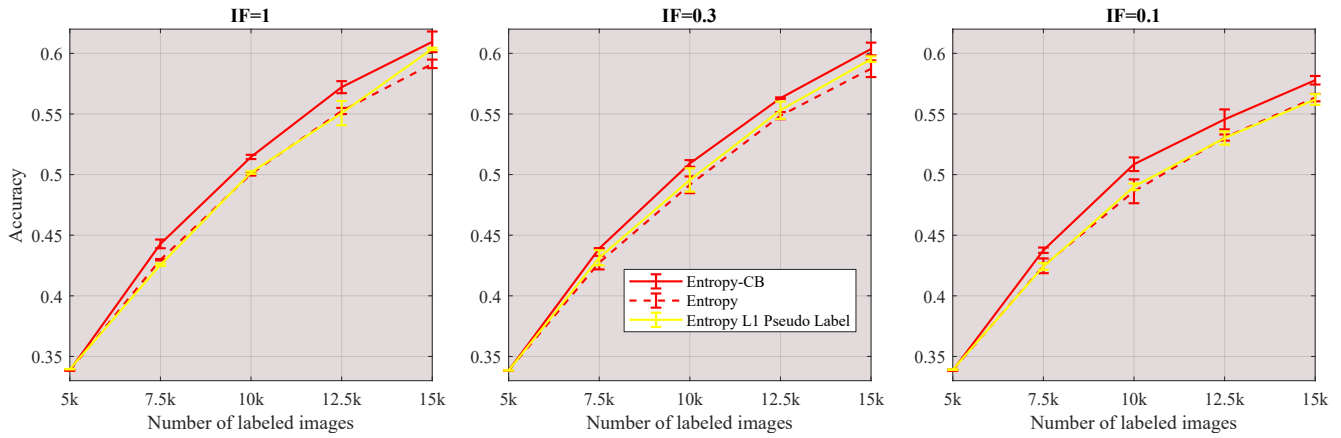


Figure 8. **Performance evaluation.** Comparing Entropy standard, Entropy balanced by Pseudo Labels against the proposed Entropy CB.

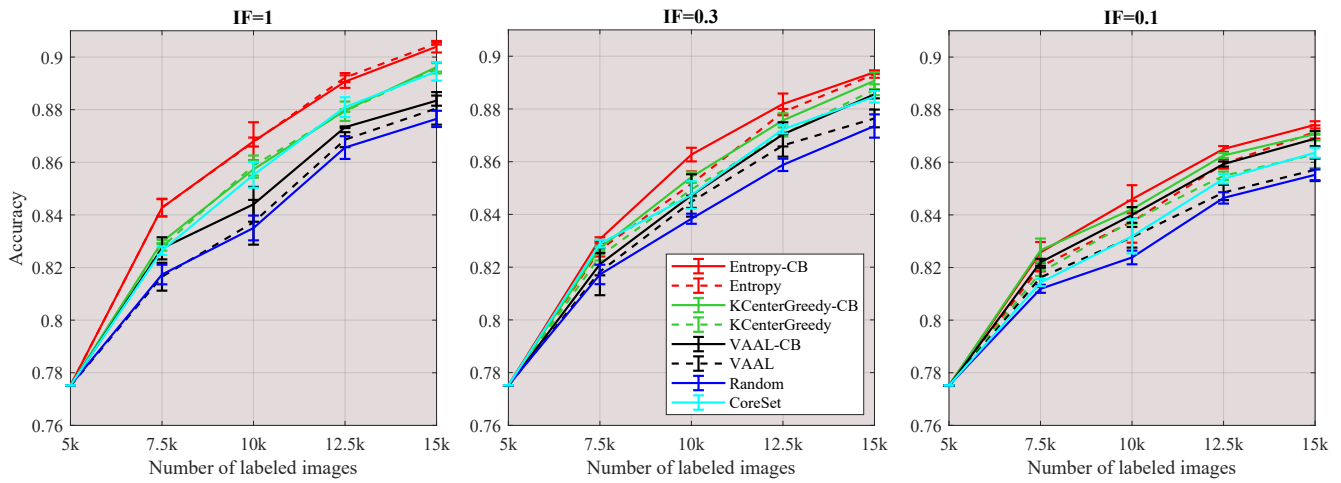


Figure 9. **Performance evaluation.** CoreSet compared to active learning methods on CIFAR10 with different imbalance factors (IF).

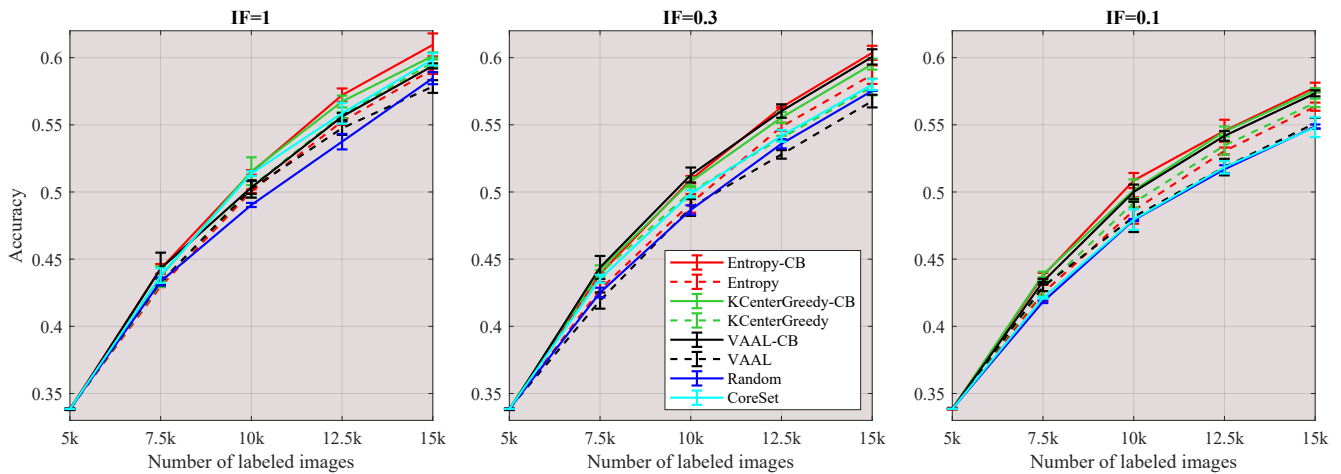


Figure 10. **Performance evaluation.** CoreSet compared to active learning methods on CIFAR100 with different imbalance factors (IF).

J. Distribution of selected samples in CIFAR100

Fig. 11, 12 and 13 show the distribution of samples selected by AL methods on original (IF=1), imbalanced (IF=0.3) and (IF=0.1) respectively. The L1 score above the distributions (introduced in Section 5.1) measures the ℓ_1 distance from uniform distribution in the corresponding cycle. As can be seen, CB methods are remarkably effective in balancing the distribution of selected samples regardless of imbalance factor. It is worth mentioning in Fig 11 although the dataset is balanced, AL baselines (Entropy and KCenterGreedy) result in biased sampling. In contrast, CB methods provide more balanced samples across all cycles and imbalance factors.

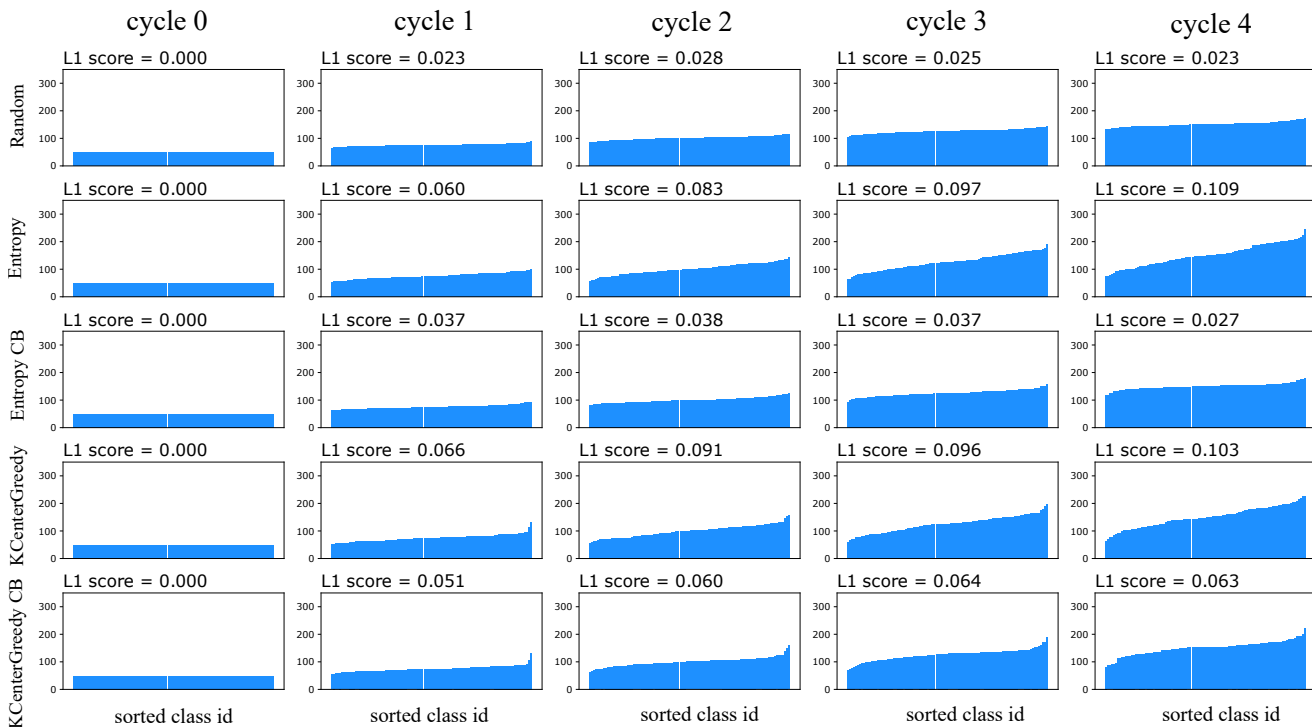


Figure 11. Distribution of samples selected by our proposed method (CB) compared to baselines on CIFAR100 with IF=1.

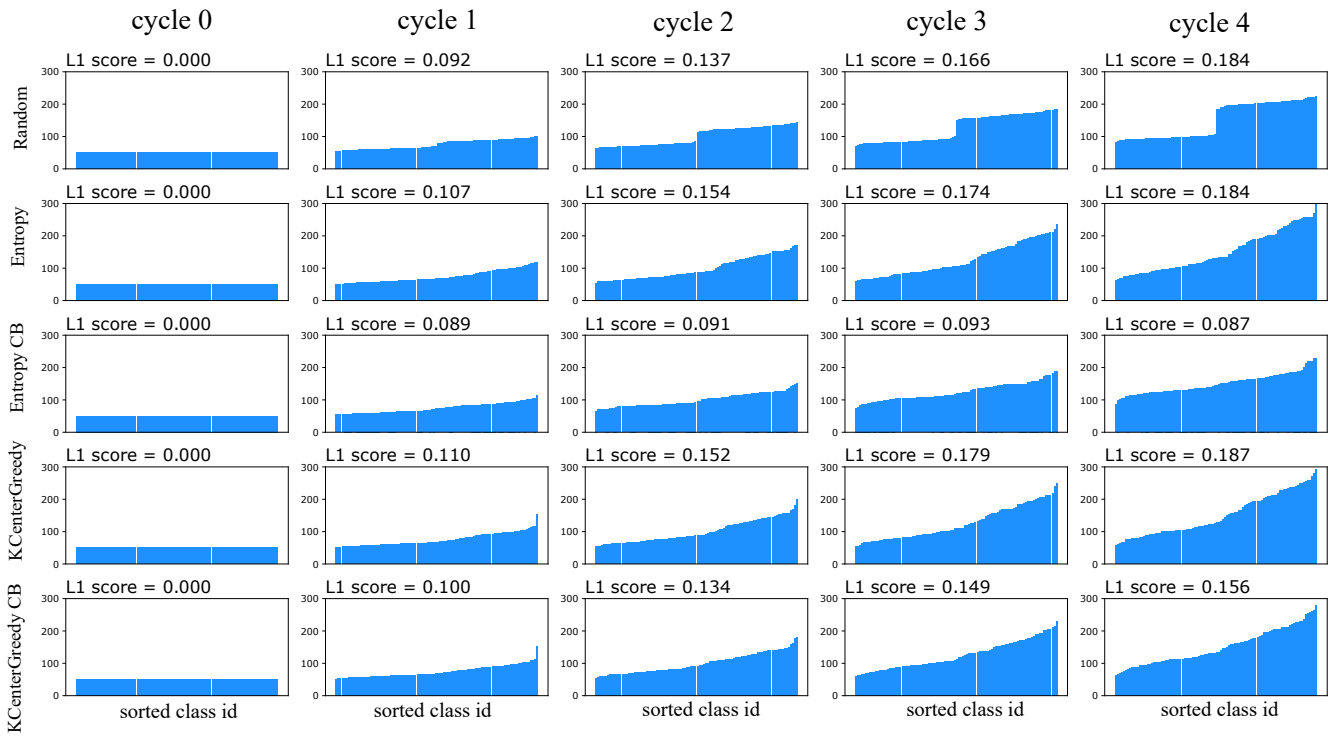


Figure 12. Distribution of samples selected by our proposed method (CB) compared to baselines on CIFAR100 with IF=0.3.

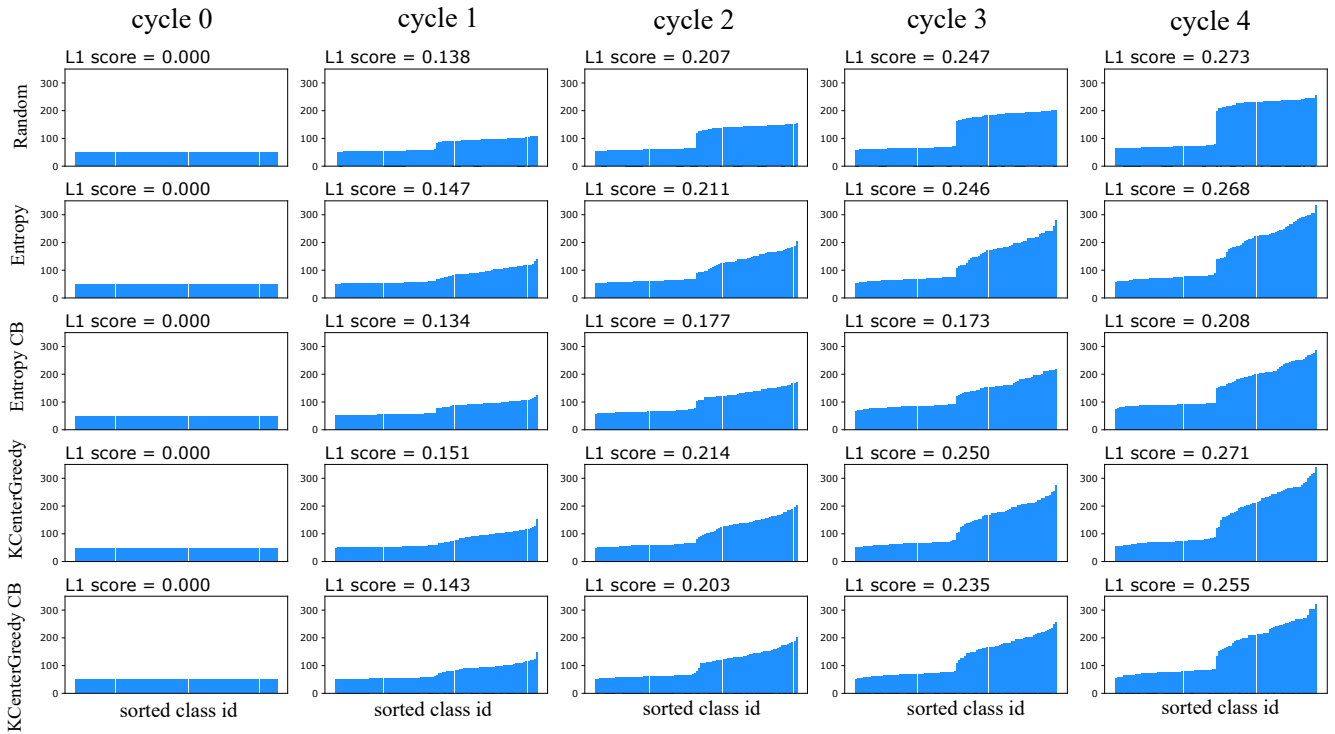


Figure 13. Distribution of samples selected by our proposed method (CB) compared to baselines on CIFAR100 with IF=0.1.