A Experiments on Fine-Grained Classification Datasets

In this section we present experiments on four fine-grained image classification datasets, where the goal is to estimate the number of categories — **Caltech-UCSD Birds** (CUB) [57] consists of 11,788 images of 200 bird species, **Stanford Cars** [29] contains 16,185 images of 196 car models, **FGVC Aircraft** [36] comprises 10,000 images of 100 aircraft models, and **Oxford Flowers** [39] includes 8,189 images of 102 flower categories. Although the names of categories are known, we use the same pairwise setting as the Re-ID tasks.

A.1 Performance of Estimating k

Here we show the estimated k as a function of the number of sampled pairs, as with the Re-ID datasets in Fig. 4. Similarly to Re-ID tasks, our method outperforms all the baselines when using feature embeddings from an ImageNet-pretrained ResNet50, DINO ViT-B/8, and CLIP ViT-L/14. We calculate the similarity as described in § 4.

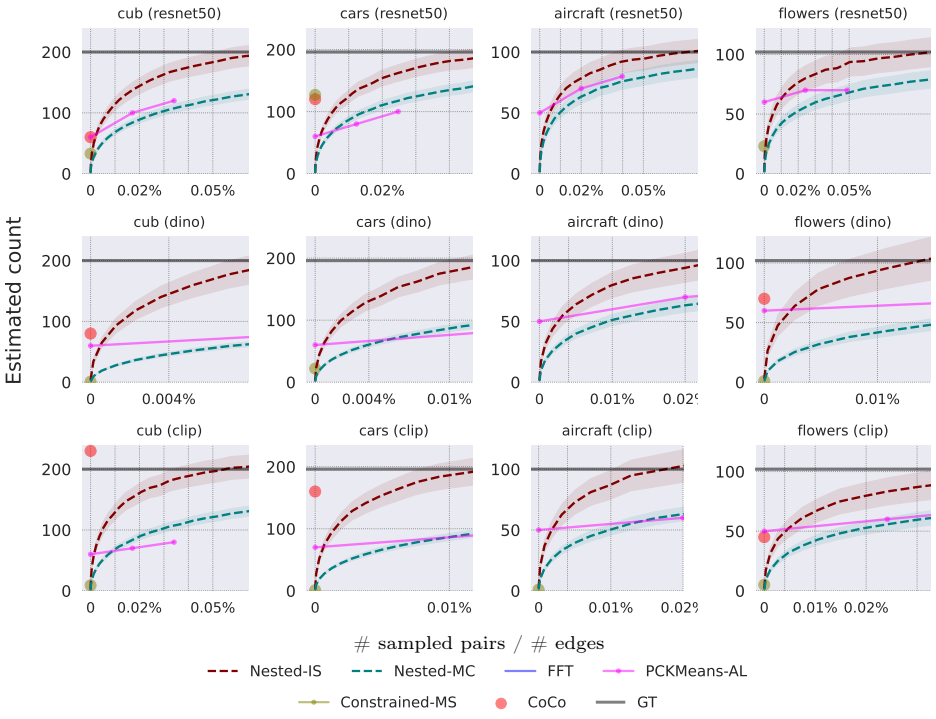

Fig. A1: Performance of Estimating k per Human Effort across fine-grained classification datasets. We use the cosine similarity built from ImageNet pretrained ResNet50, DINO ViT-B/8, and CLIP ViT-L/14 image embeddings. The human effort is measured as the fraction of the sampled pairs and total pairs $|E|$ in the dataset G. Our method estimates the true count with less human effort compared to the other baselines. Dashed lines indicate the mean estimates and shaded regions indicate the mean 95% confidence interval across 100 trials.

A.2 Measuring Clustering Accuracy

Similarly to the Re-ID datasets (See Fig. 8), we find that accurately estimating k has a bigger impact on the quality of clustering than active clustering, which suggests that human effort is better spent to estimate k initially.

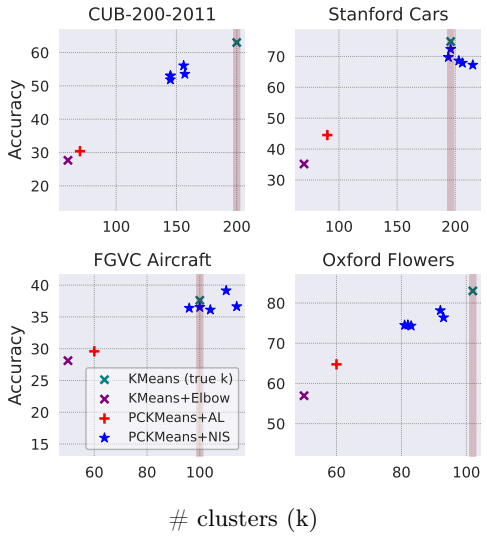

Fig. A2: Clustering accuracy on fine-grained classification datasets using the right number of clusters. Using the improved k and pairwise constraints (pc$k$-means+AL) improves the clustering accuracy over k-means+elbow, while clustering accuracy with our estimated k improves accuracy further (pc$k$-means+NIS). We plot results with five estimated ks and constraints from our proposed approach (blue stars). The red shaded line indicates the true number of clusters in the dataset.