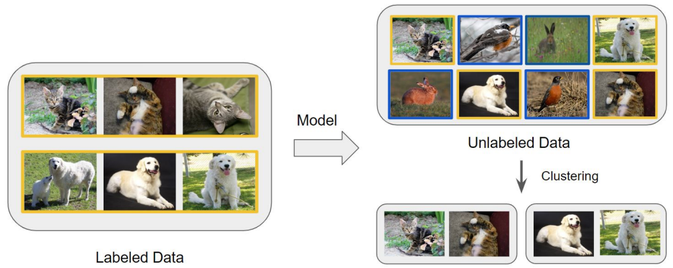
XCon: Learning with Experts for Fine-grained Category Discovery

Yixin Fei Zhongkai Zhao Siwei Yang Bingchen Zhao

BMVC 2022

Problem Definition & Motivation



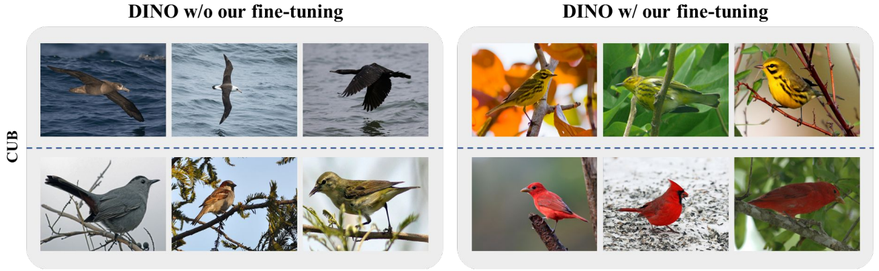
- Objective: cluster unlabeled images leveraging the information from seen classes
- Challenge: large inter-class similarity & intra-class variance
- The model aims to learn subtle discriminative cues between categories to be able to distinguish.

Contrastive Learning

unsupervised contrastive learning

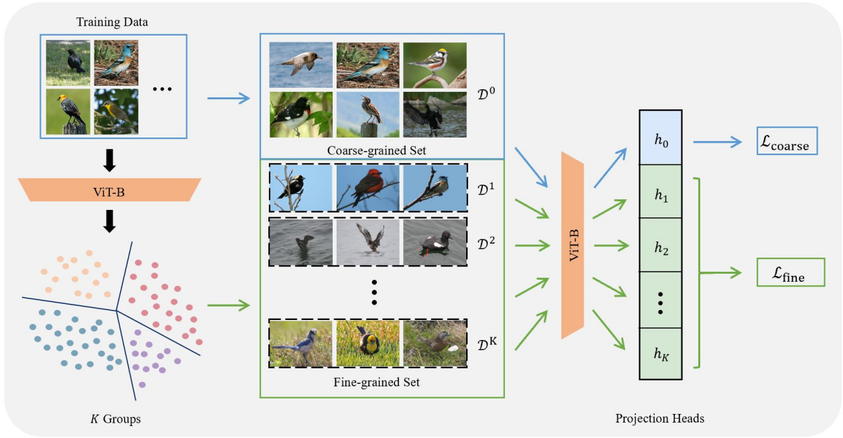
supervised contrastive learning

Self-supervised Representations



Cluster the data based on class irrelevant cues such as the object pose or the background

Method



- **Dataset partitioning**
  features clustering into K groups by k-means
- **Learning discriminative representations**
  apply contrastive loss on each sub-dataset

Experimental Results

- the best performance on the ‘All’ and ‘New’
- cluster images based on correct cues more tightly

Table 2: Results on generic datasets.

	CIFAR10			CIFAR100			ImageNet-100		
Method	All	Old	New	All	Old	New	All	Old	New
k-means [18]	83.6	85.7	82.5	52.0	52.2	50.8	72.7	75.5	**71.3**
RankStats+	46.8	19.2	60.5	58.2	77.6	19.3	37.1	61.6	24.8
UNO+	68.6	**98.3**	53.8	69.5	80.6	47.2	70.3	**95.0**	57.9
GCD [24]	91.5	97.9	88.2	73.0	76.2	**66.5**	74.1	89.8	66.3
XCon	**96.0**	97.3	**95.4**	**74.2**	**81.2**	60.3	**77.6**	93.5	69.7

Table 3: Results on fine-grained datasets.

	CUB-200			Stanford-Cars			FGVC-Aircraft			Oxford-Pet		
Method	All	Old	New	All	Old	New	All	Old	New	All	Old	New
k-means [18]	34.3	38.9	32.1	12.8	10.6	13.8	16.0	14.4	16.8	77.1	70.1	80.7
RankStats+	33.3	51.6	24.2	28.3	61.8	12.1	26.9	36.4	22.2	-	-	-
UNO+	35.1	49.0	28.1	35.5	**70.5**	18.6	40.3	**56.4**	32.2	-	-	-
GCD [24]	51.3	**56.6**	48.7	39.0	57.6	29.9	45.0	41.1	46.9	80.2	85.1	77.6
XCon	**52.1**	54.3	**51.0**	**40.5**	58.8	**31.7**	**47.7**	44.4	**49.4**	**86.7**	**91.5**	**84.1**