

Beyond Supervised vs. Unsupervised: Representative Benchmarking and Analysis of Image Representation Learning

Supplementary Material

A. Code and Assets

To reproduce our results, please visit our repository at <https://github.com/mgwillia/unsupervised-analysis>. Where specified (see our repository for details), we use code from VISSL (<https://github.com/facebookresearch/vissl/>) and SCAN (<https://github.com/wvangansbeke/Unsupervised-Classification>). These repositories have an MIT License and Creative Commons License, respectively.

B. k-NN Details

For k-NN classification, we use the VISSL defaults for ImageNet: 200 neighbors. For the FGVC datasets, there are too few images per class to use this approach. Instead, we try values in the set $\{0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$ and choose whichever value maximizes accuracy.

C. More Benchmark Results

Here, we give an expanded look at our benchmarks. Table 1 complements Figure 3 by providing the same data, in tabular form. Figures 1 and 2 along with Tables 2 and 3 do the same for k-NN and k-means, offering an expansion of the results shown in Tables 2 and 3.

We verify claims we make about the SimCLR models from the main paper. Specifically, we say that training time has a significant impact on results, while not changing the representations substantially (see Figure 6). Table 4 offers evidence supporting our claim.

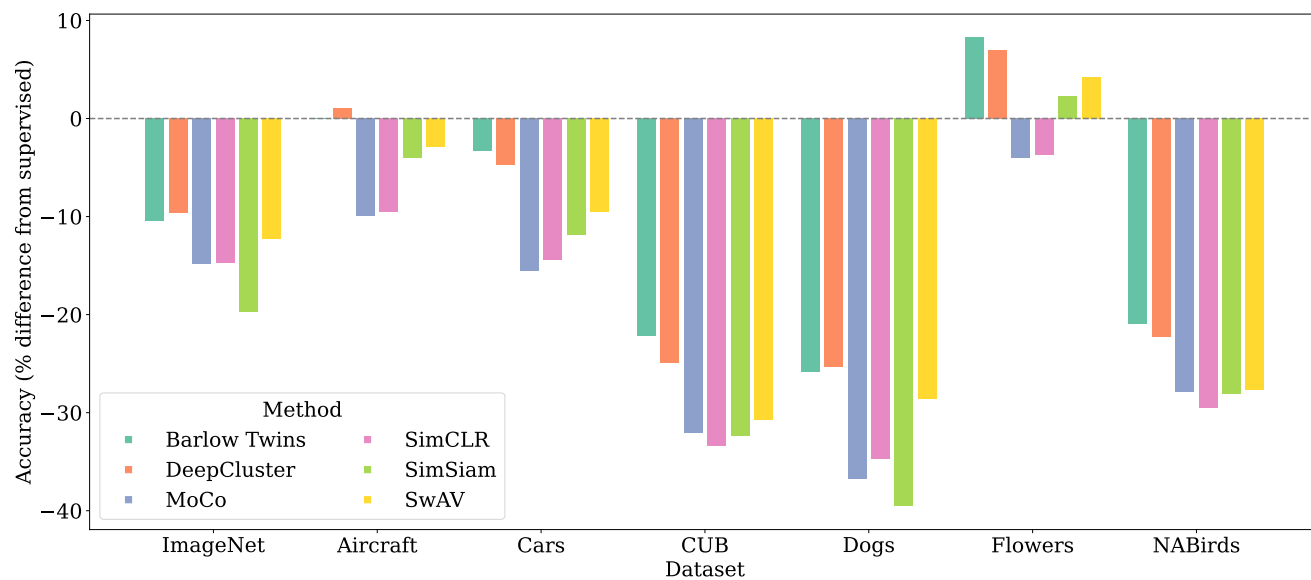


Figure 1. **k-NN results.**

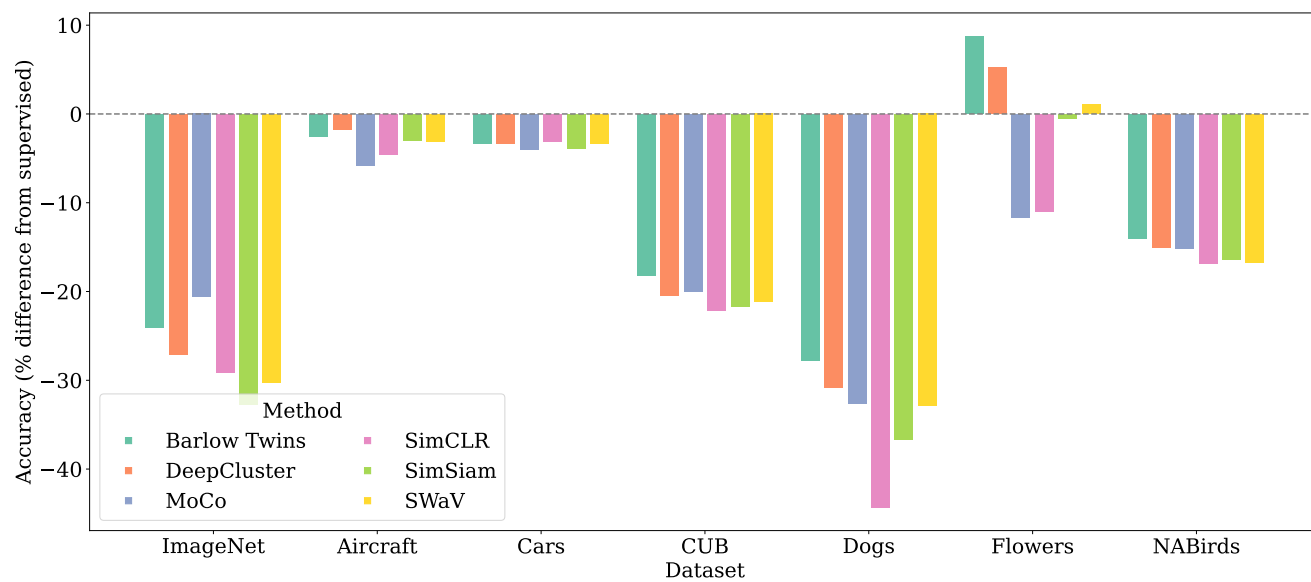


Figure 2. **K-Means results.**

Table 1. **Linear evaluation results.**

Method	Dataset						
	ImageNet	Aircraft	Cars	CUB	Dogs	Flowers	NABirds
Supervised	76.04	48.05	57.72	70.57	88.92	91.30	61.10
Barlow Twins	71.78	58.51	65.30	63.98	74.35	94.21	54.59
DeepCluster	75.19	58.36	67.76	69.82	77.80	94.46	59.89
MoCo	69.95	49.88	49.68	51.95	67.93	91.14	43.51
SimCLR	68.95	44.39	45.00	47.36	65.84	88.90	38.43
SimSiam	67.89	52.22	53.34	52.55	66.50	91.86	44.18
SwAV	74.87	55.73	61.95	65.10	75.99	93.97	56.52

Table 2. **k-NN results.**

Method	Dataset						
	ImageNet	Aircraft	Cars	CUB	Dogs	Flowers	NABirds
Supervised	73.41	31.59	30.16	56.63	88.38	77.96	43.25
Barlow Twins	62.90	31.83	26.94	34.41	62.53	86.18	22.29
DeepCluster	63.70	32.70	25.48	31.74	62.97	84.76	21.05
MoCo	58.59	21.39	14.64	24.35	51.60	74.53	15.40
SimCLR	54.57	21.21	14.74	23.21	49.63	74.78	14.03
SimSiam	53.66	27.39	18.41	24.20	48.97	80.01	15.18
SwAV	61.14	28.77	20.84	25.75	59.87	82.24	15.72

Table 3. **K-Means results.**

Method	Dataset						
	ImageNet	Aircraft	Cars	CUB	Dogs	Flowers	NABirds
Supervised	58.92	15.69	11.95	35.23	53.69	54.97	25.95
DeepCluster	31.79	13.92	8.66	14.81	22.84	60.20	10.86
MoCo	38.30	9.84	7.98	15.21	21.10	43.34	10.75
Barlow Twins	34.88	13.20	8.63	17.07	25.94	63.70	11.87
SimCLR	29.78	11.16	8.80	13.07	9.41	43.99	9.08
SimSiam	26.20	12.66	8.03	13.57	17.07	54.51	9.53
SwAV	28.69	12.60	8.66	14.05	20.79	56.04	9.26

Table 4. **Linear Evaluation for SimCLR** with varying training time.

Method	Dataset						
	ImageNet	Aircraft	Cars	CUB	Dogs	Flowers	NABirds
100 Epochs	64.76	44.81	44.67	43.17	60.44	88.72	34.34
200 Epochs	66.92	45.56	46.31	46.05	62.48	89.39	36.90
400 Epochs	67.93	44.84	46.36	46.00	64.35	89.08	37.23
800 Epochs	68.95	44.39	45.00	47.36	65.84	88.90	38.43
1000 Epochs	64.57	45.26	44.55	46.93	66.25	88.57	37.93