On Partial Prototype Collapse in the DINO Family of Self-Supervised Methods

Hariprasath Govindarajan^{1,2}, Per Sidén¹, Jacob Roll¹, Fredrik Lindsten²

¹ Qualcomm Technologies, Inc. ² Linköping University, Sweden

Qualcom







All self-supervised methods in the DINO family suffer from a *partial prototype collapse*.

- Hardly 1-10% of the prototypes are unique/used!

What happens if we effectively utilize the prototypes?

■ IF pre-training data == uniform distribution :

- Classification performance on in-domain data
 Few-shot learning performance on in-domain data
- ELIF pre-training data == long-tailed distribution :
 - Long-tailed classification performance
 Transfer learning performance

Transfer learning performance

DINO - Preliminaries

DINO learns representations by clustering the images into a set of clusters such that the clusters are consistent across multiple different views of the images.



Mitigating partial prototype collapse

We use a KoLeo-Proto regularization to encourage diverse prototypes – by maximizing the differential entropy of the prototypes.





- KoLeo-Proto regularization enables effective utilization of prototypes.
- With effective utilization of prototypes, more prototypes improve kNN performance.

DINO Family – Partial Prototype Collapse



DINO family of methods are regularized by encouraging the distribution of data over the

Uniformly distributed data (ImageNet-1K)

	Method	kNN	Linear	Finetuning	Avg. Transfer
In-domain 🕨	ViT-Base/16				
classification	DINO-vMF	77.4	78.8	83.6	86.8
	MSN	73.3	74.8	_	85.3
	WE-SSL	77.2	78.9	_	_
	iBOT-vMF	<u>78.7</u>	<u>80.3</u>	84.1	87.3
	iBOT-vMF (kp)	78.8	80.5	84.1	<u>86.6</u>
	ViT-Small/16				
	DINO-vMF	74.7	77.0	81.8	85.4
	MSN	74.9	76.6	_	84.1
	WE-SSL	75.2	77.4	_	84.6
	iBOT-vMF	<u>75.3</u>	77.9	82.3	<u>85.5</u>
	iBOT-vMF (kp)	75.5	77.9	82.3	85.6
	Baseline KoLeo-Proto				

clusters to match a uniform prior.

The *partial prototype collapse* serves as a shortcut to achieve such a uniform distribution.

Prototype Utilization

	DINO (ViT-S/16) Imagenet-1K	iBOT-vMF (ViT-B/16) Imagenet-1K	iBOT (ViT-L/16) Imagenet-22K	EsViT (Swin-T/7) Imagenet-1K	DINO (Resnet50) Imagenet-1K
Number of prototypes	65536	8192	8192	65536	60000
Number of "unique" prototypes	1078 (1.6%)	1170 (14.3%)	969 (11.8%)	1157 (1.8%)	984 (1.6%)





Long-tail distributed data (iNaturalist-2018)



75 –

Transfer Learning

ViT-S/16

ViT-B/16

