ICLR
International Conference On Learning Representations

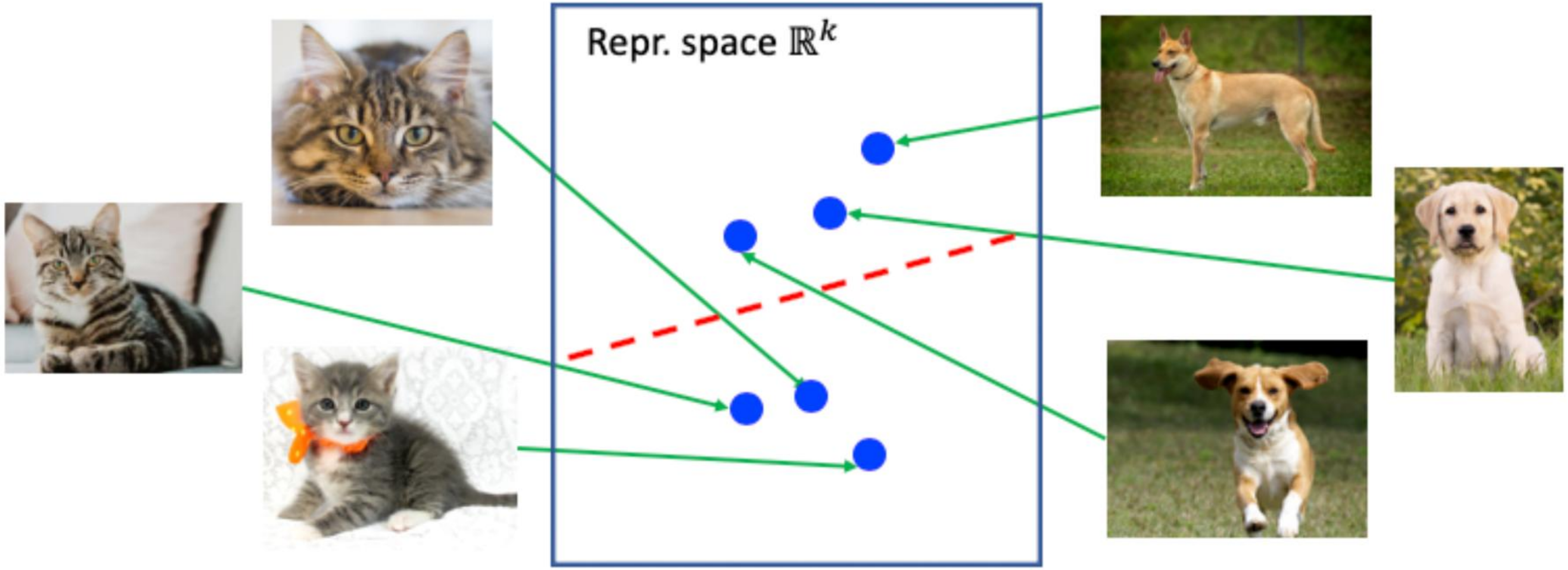
Theoretical Analysis of Contrastive Learning under Imbalanced Data: From Training Dynamics to a Pruning Solution

Haixu Liao[1] Yating Zhou[2] Songyang Zhang[3] Meng Wang[4] Shuai Zhang[1]

[1] New Jersey Institute of Technology [2] Cornell University [3] University of Louisiana at Lafayette [4] Rensselaer Polytechnic Institute

Background: Contrastive Learning

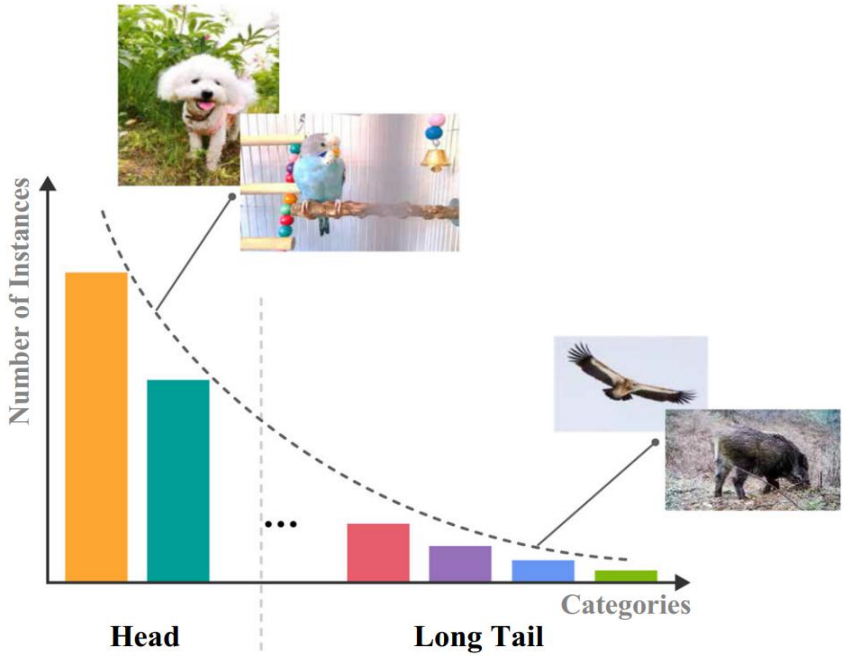
- Contrastive learning makes similar images closer together in the representation space



Background: Imbalanced Data

Long-tailed datasets exist in many real-world scenarios

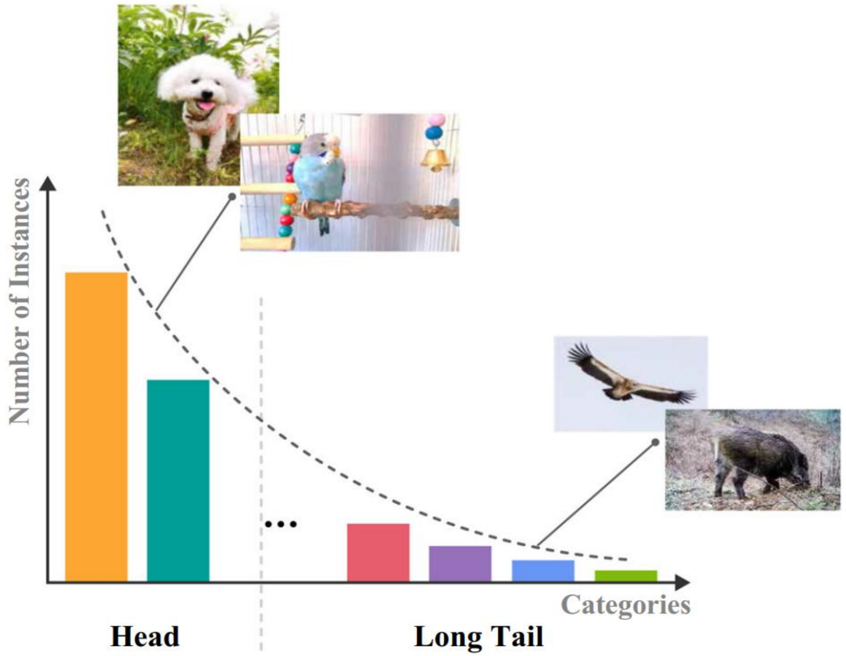
- Common classes: dog, budgie, and others
- Uncommon classes: eagle and many more



Background: Imbalanced Data

Long-tailed datasets exist in many real-world scenarios

- Common classes: dog, budgie, and others
- Uncommon classes: eagle and many more



However, contrastive learning still struggles with class imbalance in real-world datasets

Main contribution: Training Dynamics

Contribution 1: we build a theoretical framework to understand how contrastive learning behaves during training

- Stage 1: All neurons grow in feature directions but non-feature directions are ignored

- Stage 2: Lucky neurons then specialize in single features, and ordinary neurons learn a mix of features

- Stage 3: Each neuron becomes strongly aligned with one or more features, weakly aligned with other features, and remaining small in non-feature directions

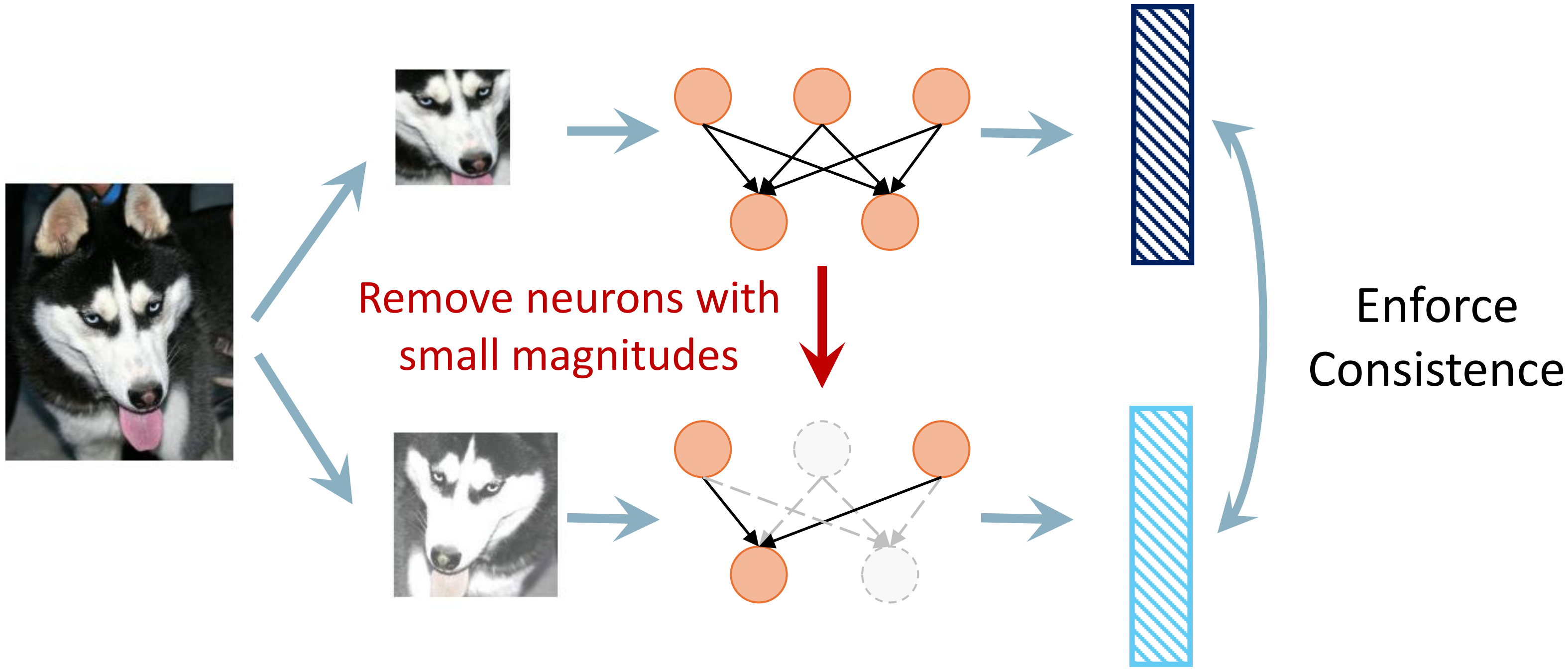
Main contribution: Data Imbalance

Contribution 2: our analysis shows that data imbalance hurts the quality of representations in several ways

- Slower learning of minority features requires a longer time to converge
- Fewer neurons specializing in single features
- Higher model complexity needed to capture all features, which implies longer training time and more neurons

Pruning Algorithm

- Dynamically removes small-magnitude neuron weights during the forward pass
- Keeps all parameters trainable during the backward pass



Credit:
Jiang et al., ICML 2021;
Qian et al., NeurIPS 2022.

Main contribution: Pruning

Contribution 3: magnitude-based pruning can enhance the learning of minority features

- Enhanced neuron updates along minority-feature directions
- More neurons focusing on pure minority features
- More balanced and robust representations

Numerical Result: Pruning

- Consistently improves accuracy across all datasets
- Narrows performance gap between head and tail classes

Dataset	ρ	**Accuracy**		$\boldsymbol{\Delta_{20}}$	
		w/o pruning	**w/ pruning**	**w/o pruning**	**w/ pruning**
CIFAR10-LT	10	79.25 ± 1.03	84.92 ± 0.67	3.42 ± 1.02	2.99 ± 0.92
	50	75.58 ± 0.84	83.60 ± 1.02	3.92 ± 1.21	3.35 ± 0.76
	100	74.24 ± 0.82	81.31 ± 0.94	5.69 ± 1.35	5.62 ± 0.99
CIFAR100-LT	10	51.21 ± 1.21	56.33 ± 1.51	2.45 ± 0.57	1.37 ± 0.46
	50	49.32 ± 0.45	56.12 ± 0.32	4.95 ± 1.02	2.57 ± 0.92
	100	47.12 ± 0.51	54.93 ± 0.50	7.11 ± 0.45	4.38 ± 0.22
ImageNet-LT	256	63.21	65.12	8.47	7.21