HEX: Hierarchical Emergence Exploitation in Self-Supervised Algorithms

Kiran Kokilepersaud, Seulgi Kim, Mohit Prabhushankar, Ghassan AlRegib

Omni Lab for Intelligent Visual Engineering and Science (OLIVES)

{ kpk6, seulgi.kim, mohit.p, alregib } @ gatech.edu

WACV 2025 — TUCSON, ARIZONA • FEB 28 - MAR 4

ArXiv   Github

OLIVES @GeorgiaTech

https://alregib.ece.gatech.edu/

Motivation

50 Epochs of Training

450 Epochs of Training

a) Traditional View of Dimensional Collapse
b) HEX view of Dimensional Collapse

Legend
- Global Collapse Point
- Bird Collapse Point
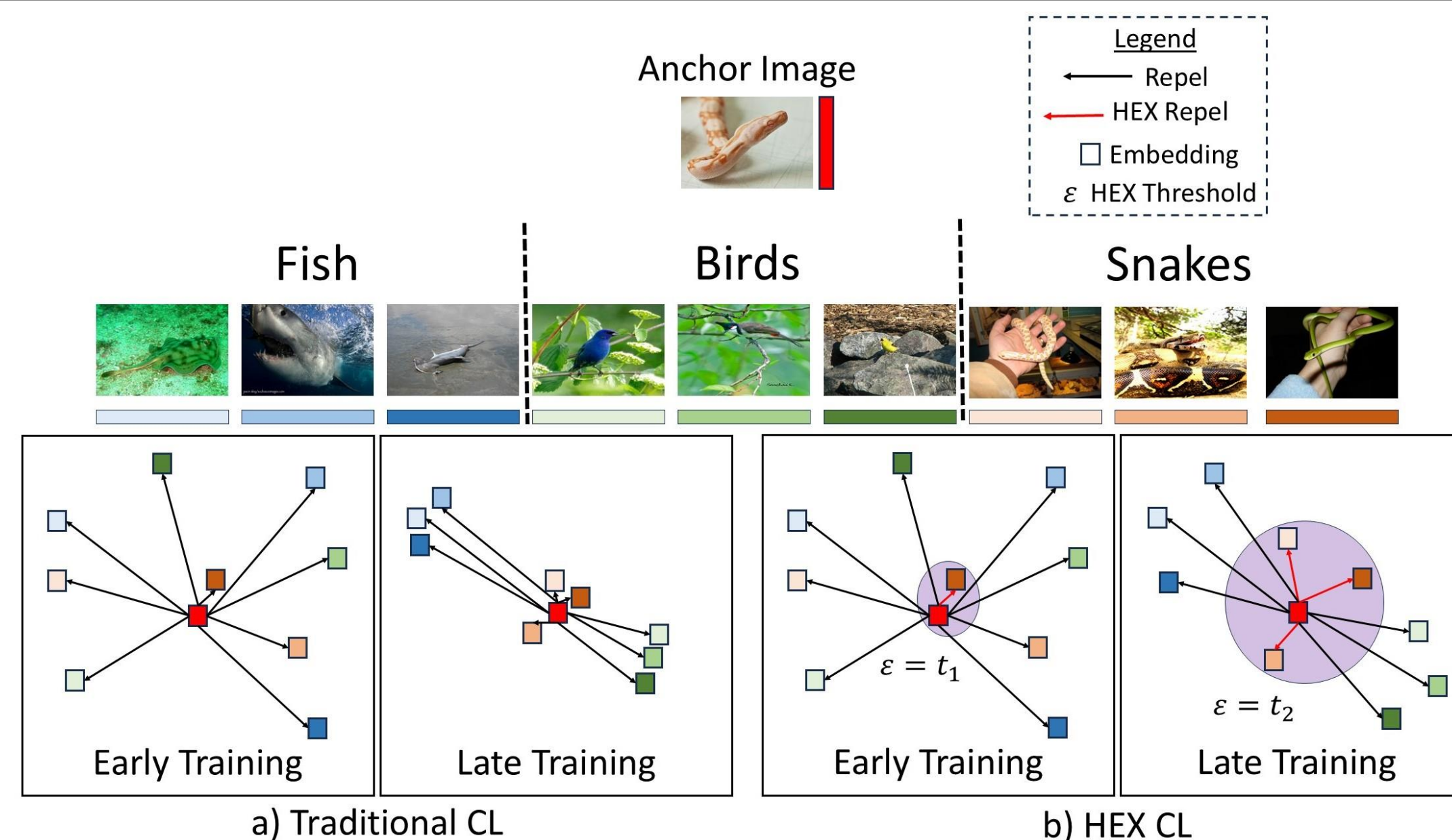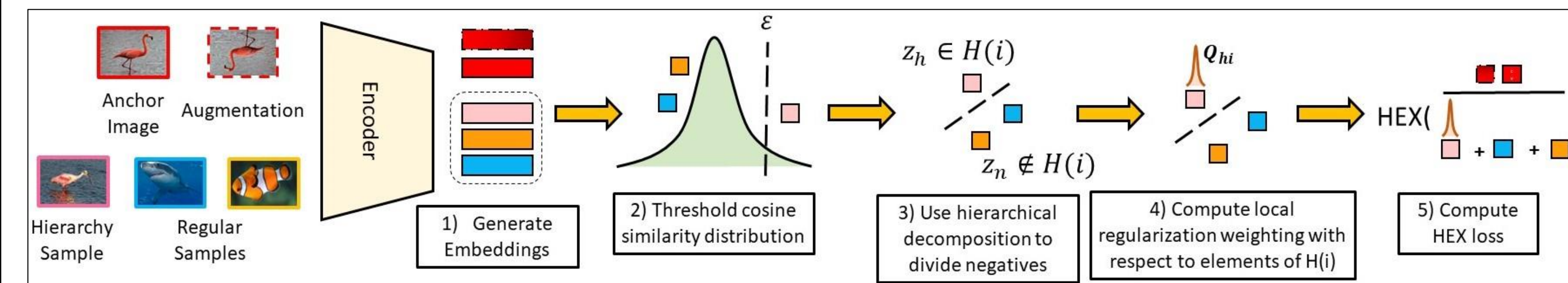- Fish Collapse Point
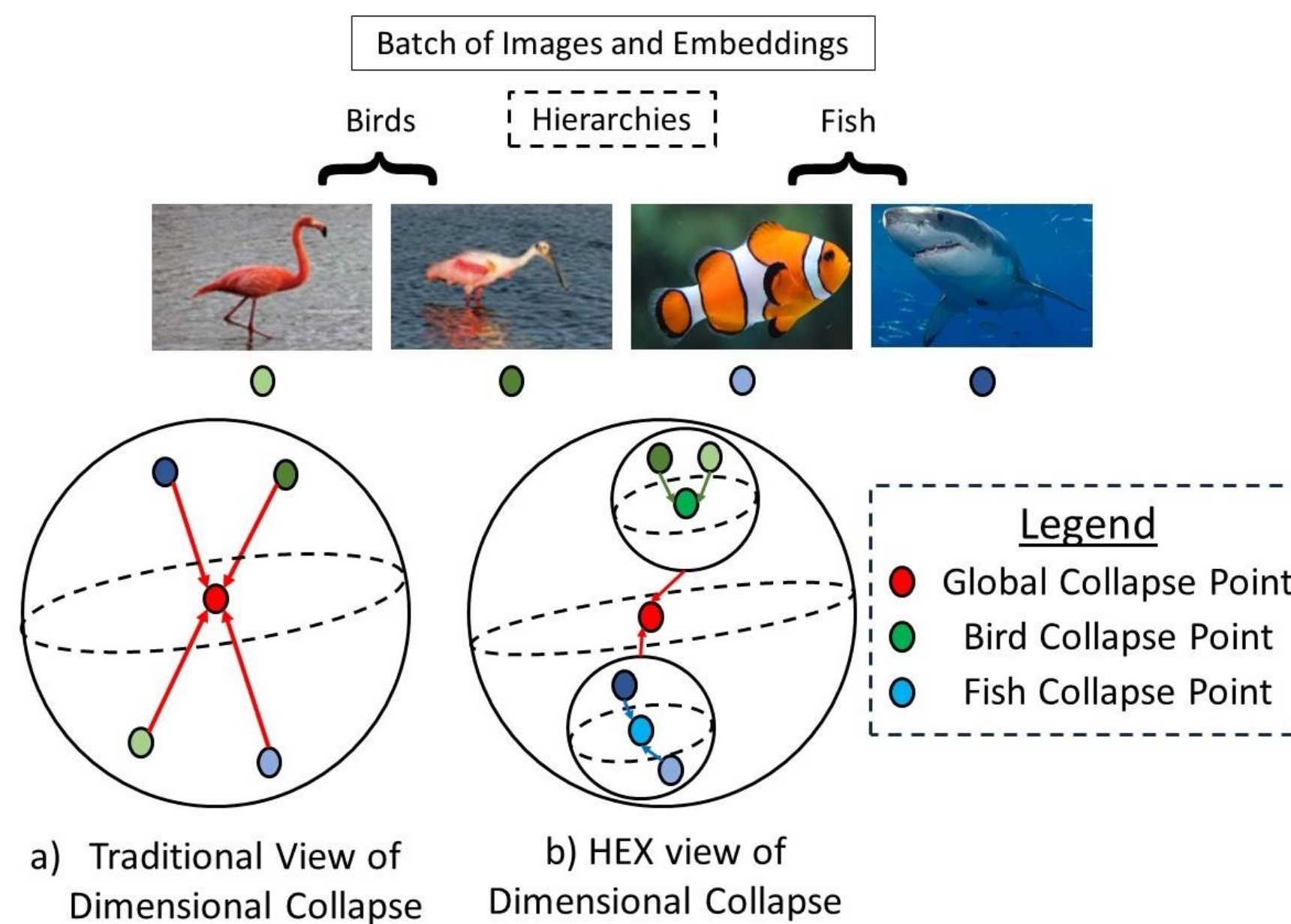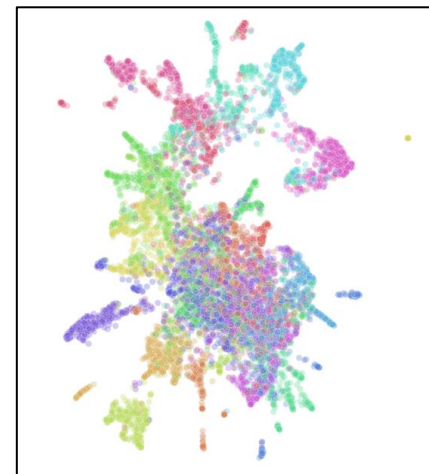
Observations

1. SSL algorithms are optimized to prevent **global dimensional** collapse.
2. Classes of **same hierarchy can share features** which collapses locally.
3. Hierarchies **emerge naturally** during training of SSL algorithm.

HEX Intuition

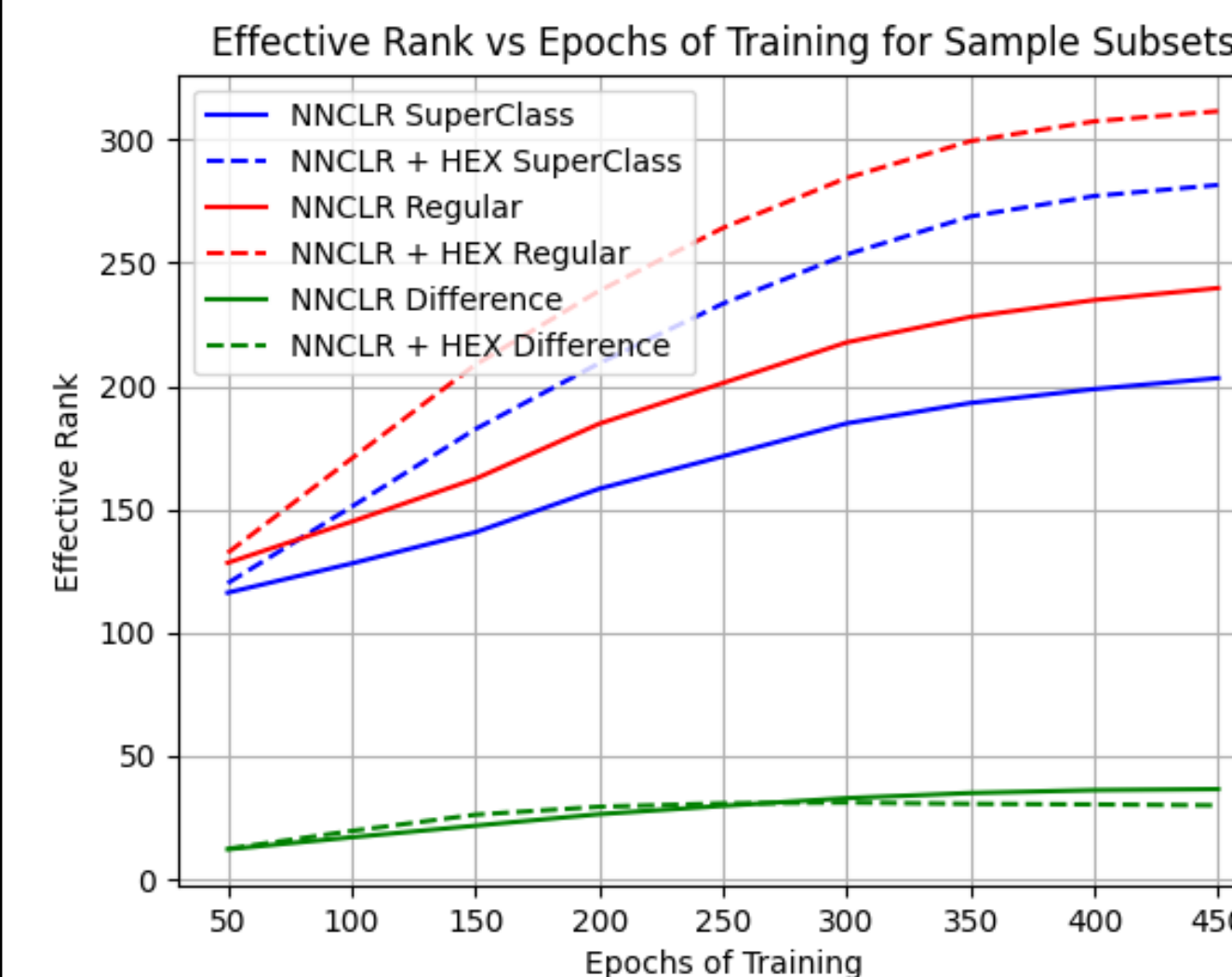Local hierarchies emerge. Can we exploit this behavior?

Anchor Image

Legend
- Repel
- HEX Repel
- Embedding
- ε HEX Threshold

Fish   Birds   Snakes

Early Training   Late Training $\varepsilon = t_1$ Early Training $\varepsilon = t_2$ Late Training

a) Traditional CL   b) HEX CL

Methodology

1) Generate Embeddings
2) Threshold cosine similarity distribution
3) Use hierarchical decomposition to divide negatives
4) Compute local regularization weighting with respect to elements of H(i)
5) Compute HEX loss

Anchor Image, Augmentation, Hierarchy Sample, Regular Samples

$$L_{HEX} = -\sum_{i \in I} \log\left(\frac{\exp\left(z_i \cdot \frac{z_{j(i)}}{\tau}\right)}{Q_{hi}(\sum_{h \in H(i)} \exp(z_i \cdot \frac{z_h}{\tau})) + \sum_{n \notin T(i)} \exp(z_i \cdot \frac{z_n}{\tau})} \right)$$

z_i = embedding of anchor image
$i \in I$ = anchor image
$z_{j(i)}$ = embedding of augmentation
z_h = embedding of hierarchy negatives
Q_{hi} = weighting function
z_a = embedding of negatives
z_n = embedding of regular negatives
$H(i)$ = Set of Hierarchy Negatives
τ = Temperature

Analysis of Hierarchical Representations

Effective Rank vs Epochs of Training for Sample Subsets

- NNCLR SuperClass
- NNCLR + HEX SuperClass
- NNCLR Regular
- NNCLR + HEX Regular
- NNCLR Difference
- NNCLR + HEX Difference
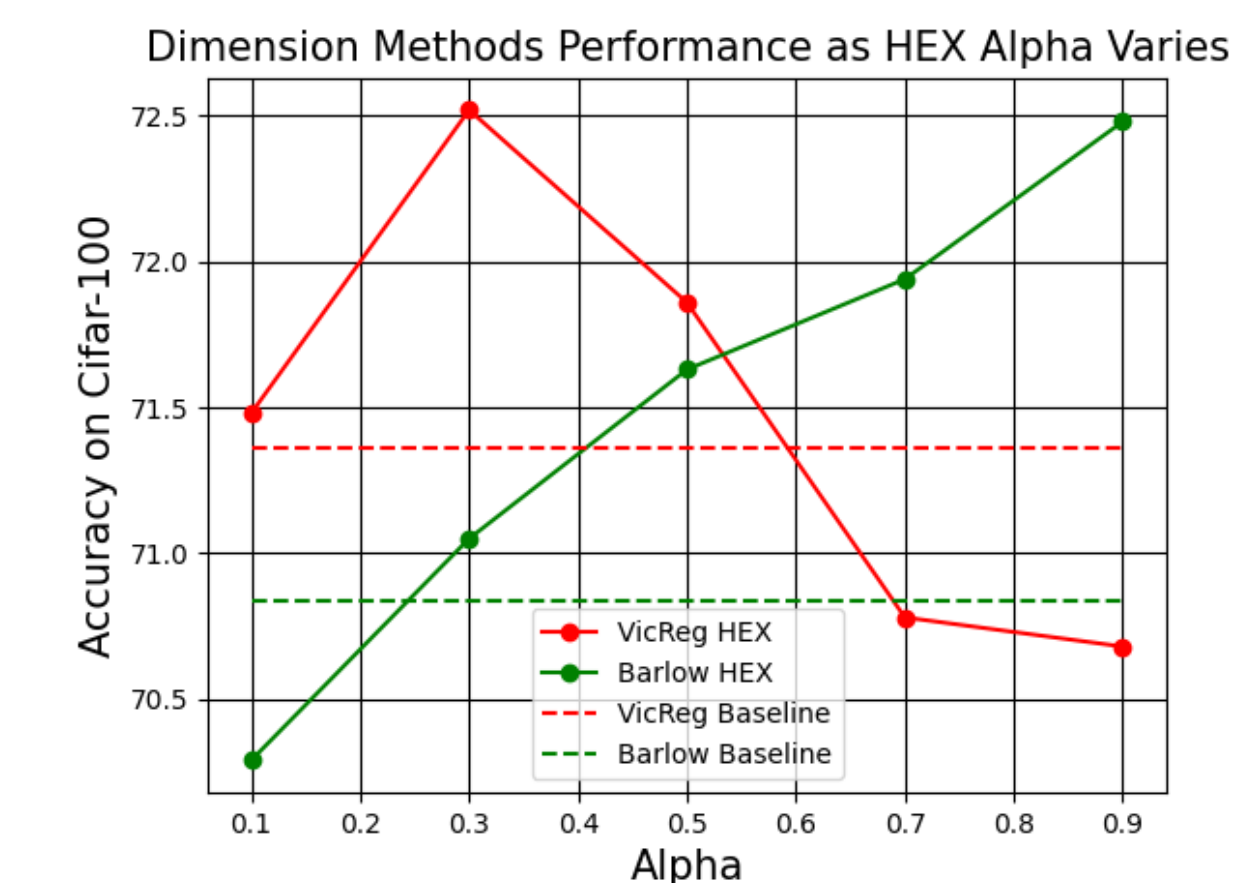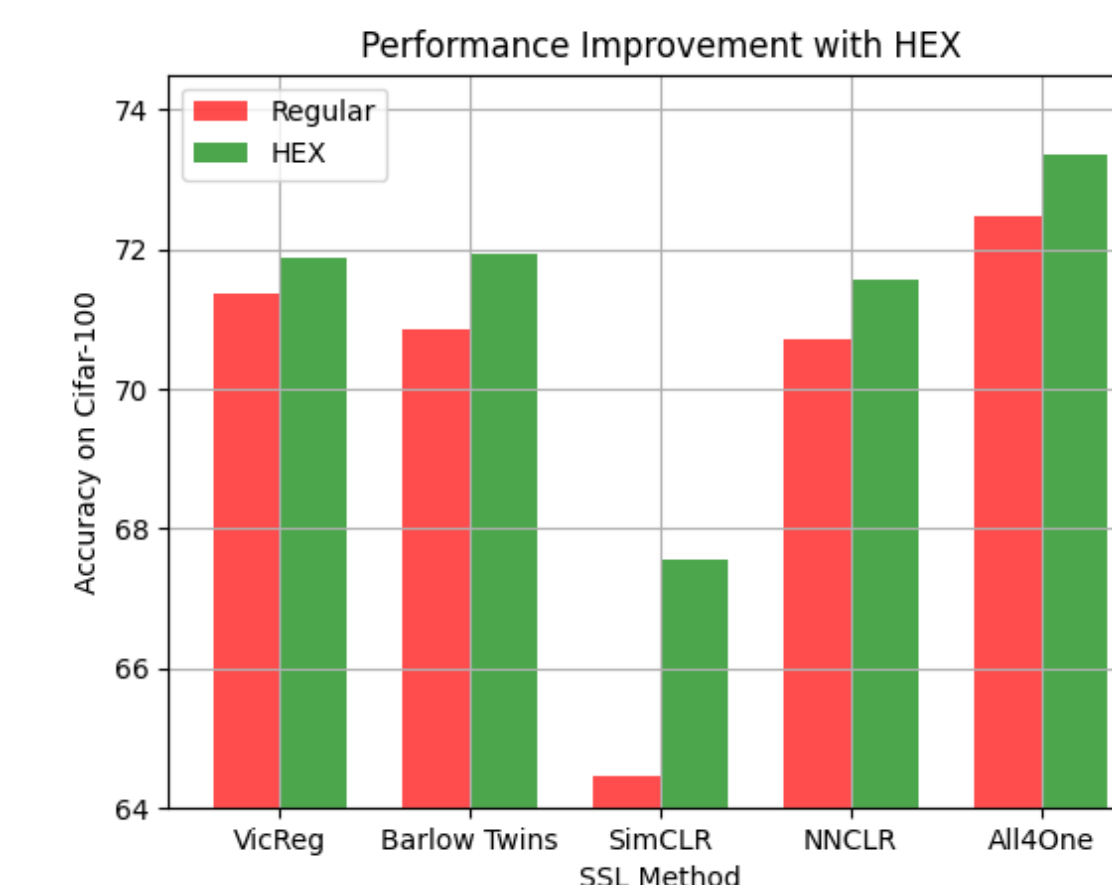
Members of same hierarchy have **greater local collapse**.

Cosine similarity distribution varies for superclass and regular samples at different epochs.
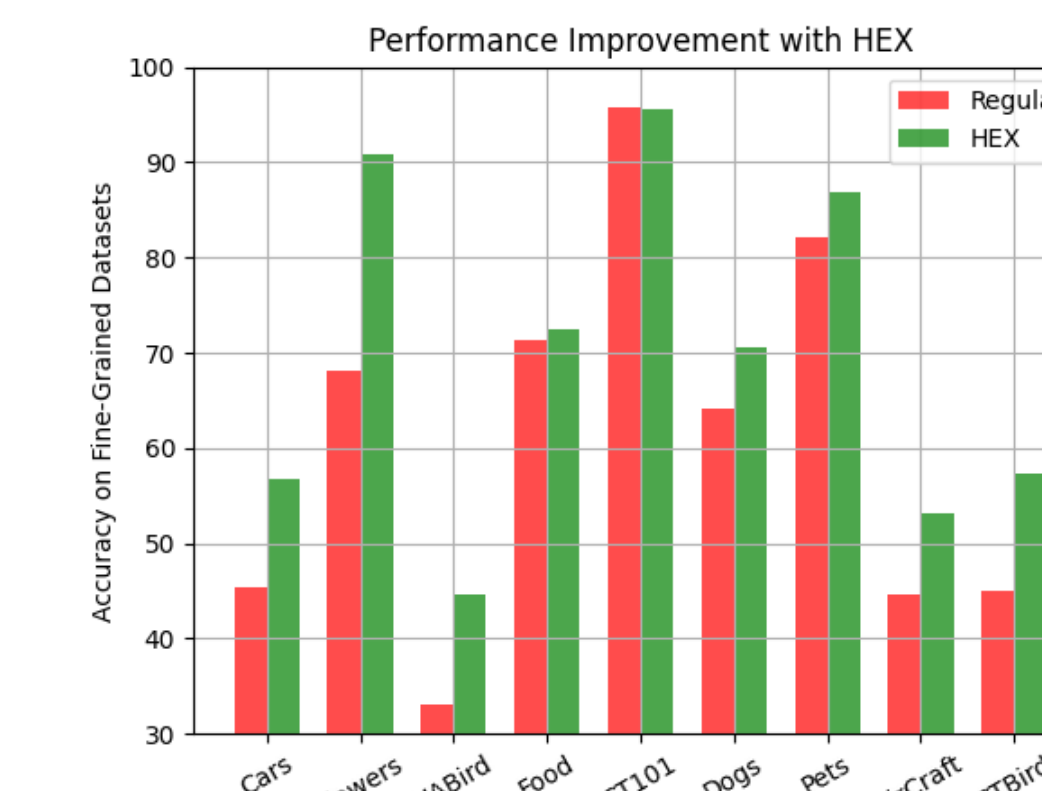
Can separate hierarchical and regular samples.

a) **SuperClass Cosine Similarity Distribution Across Training**

b) **Regular Cosine Similarity Distribution Across Training**
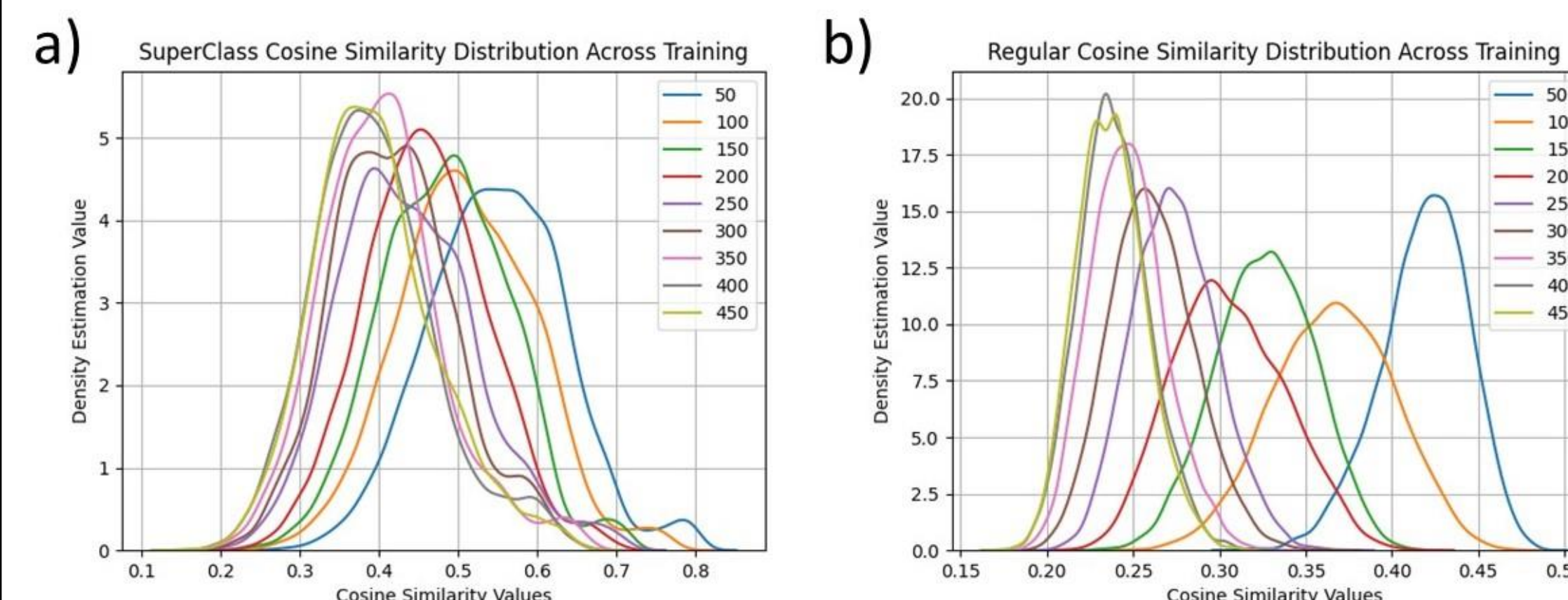
Results

HEX can be applied to both Sample Contrastive and Dimension Contrastive Strategies

Performance Improvement with HEX

SSL Methods: VicReg, Barlow Twins, SimCLR, NNCLR, All4One

Dimension Methods Performance as HEX Alpha Varies

- VicReg HEX
- Barlow HEX
- VicReg Baseline
- Barlow Baseline

Performance improvement consistent across datasets

Performance Improvement with HEX

Datasets: Cars, Flowers, NABird, Food, CT101, Dogs, Pets, AirCraft, CTBird

Method	ImageNet	ImageNet Semi-Supervised (1%)
SimCLR	57.74%	29.61%
SimCLR + HEX	**59.03%**	**33.72%**
NNCLR	62.47%	39.94%
NNCLR + HEX	**65.96%**	**46.00%**

Conclusion

- We propose a self-supervised methodology based on cosine similarity distributions to *identify the existence of local hierarchical groupings of images* within the representation space.
- We *introduce a local hierarchical regularization into the training process* for both sample and dimension contrastive strategies.
- We introduce an ε threshold term during the training process that can be tuned adaptively or manually to reflect the **growing hierarchical organization** of the representation space.