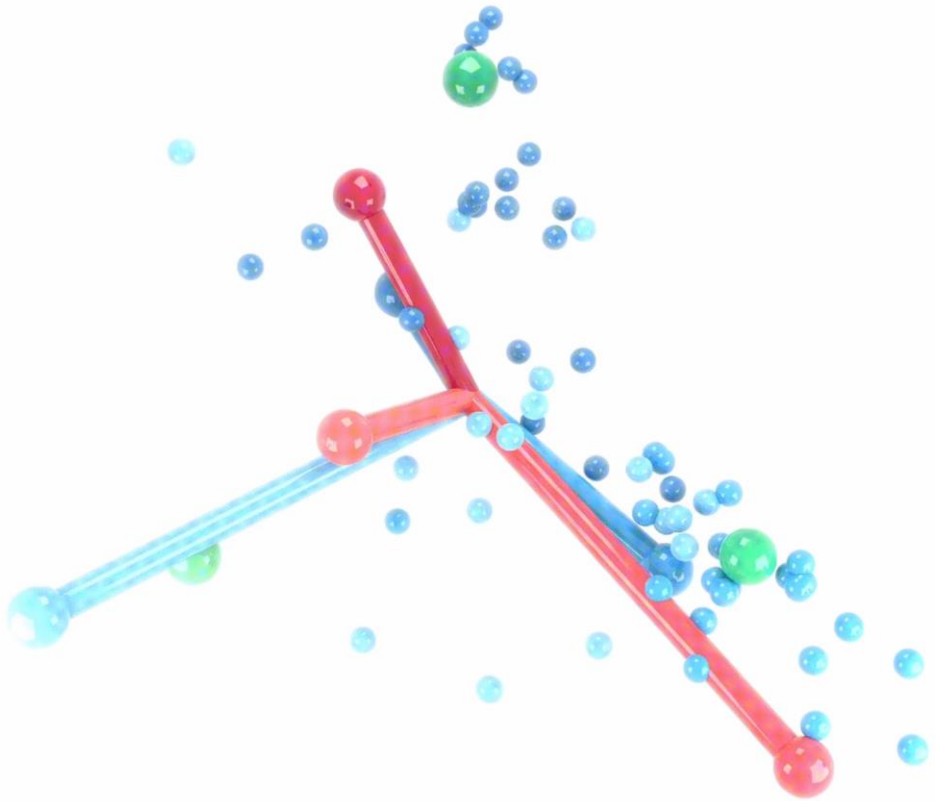
Understanding Imbalanced Semantic Segmentation Through Neural Collapse

Zhisheng Zhong [1*] **Jiequan Cui** [1*] **Yibo Yang** [2*] **Xiaoyang Wu** [3] **Xiaojuan Qi** [3] **Xiangyu Zhang** [4] **Jiaya Jia** [1]

[1] CUHK [2] JD Explore Academy [3] HKU [4] MEGVII Technology



Red ball-and-sticks: *Linear classifiers*

Blue ball-and-sticks: *Class-means (centers)*

Small blue spheres: *Last-layer features*

Green spheres: *Simplex equiangular tight frame (ETF)*

Neural Collapse*: As training progresses ……*

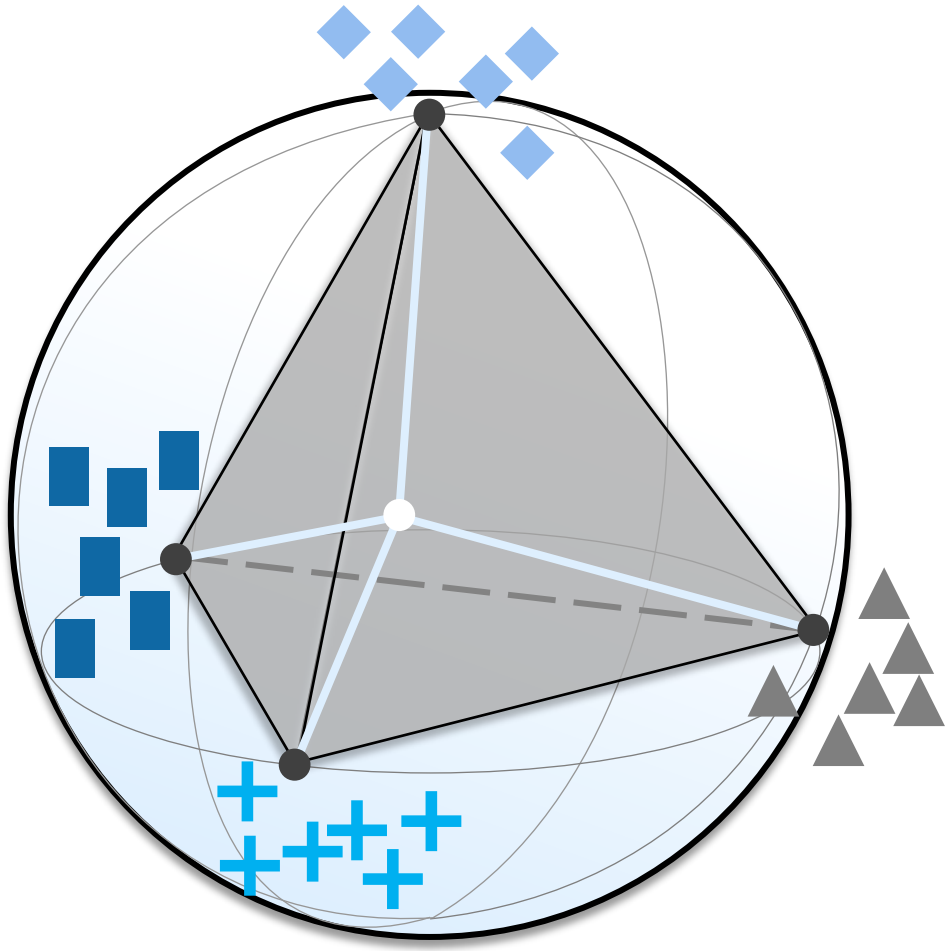
Linear classifiers

Class-means ***All*** *collapse to* ***Simplex ETF***

Last-layer features

V. Papyan, X. Han, and D. L. Donoho. Prevalence of Neural Collapse During the Terminal Phase of Deep Learning Training. Proceedings of the National Academy of Sciences (PNAS) 2020.

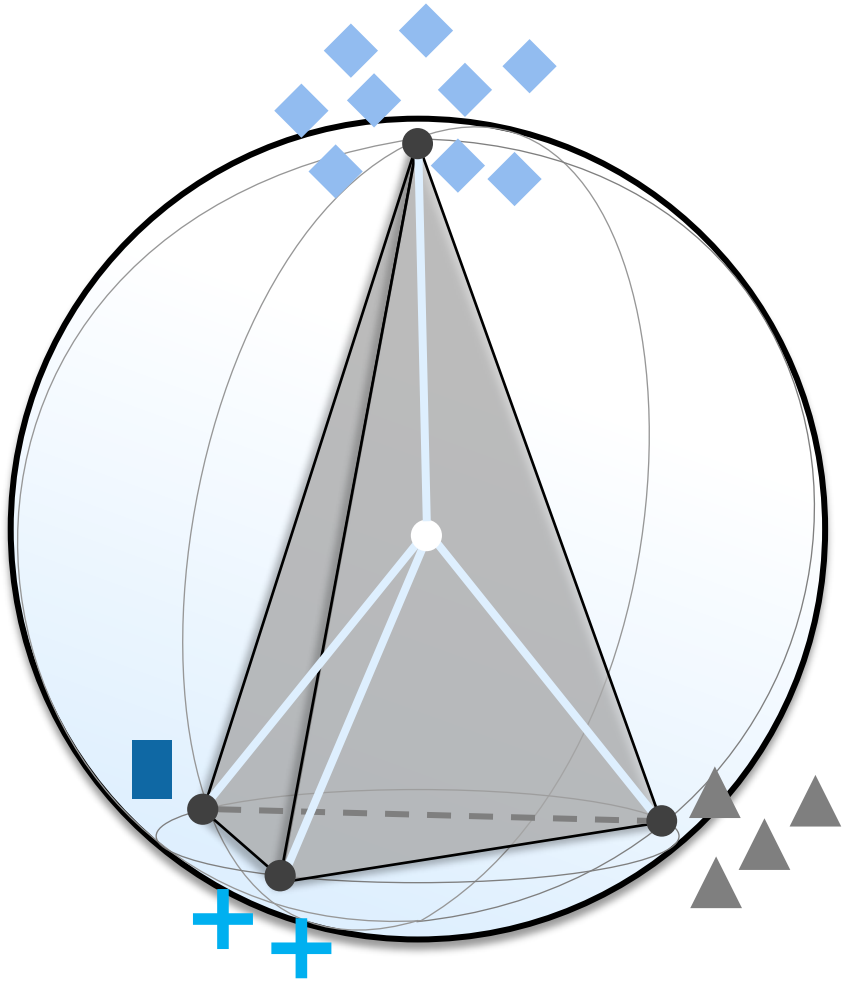
Classification



*normally **balanced distribution** &*
relatively unique among instances

→ *easier to reach the **equiangular** and **maximal separated** structure (ETF) of neural collapse*

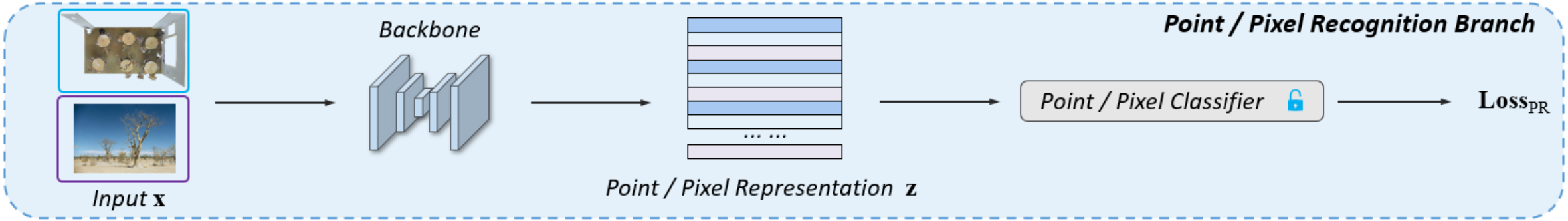
Semantic Segmentation



*naturally brings **contextual correlation** &*
imbalanced distribution (point/pixel)

⟹ *break the **equiangular** and **maximal separated** structure (ETF) of neural collapse*

Conventional point cloud & image semantic segmentation framework

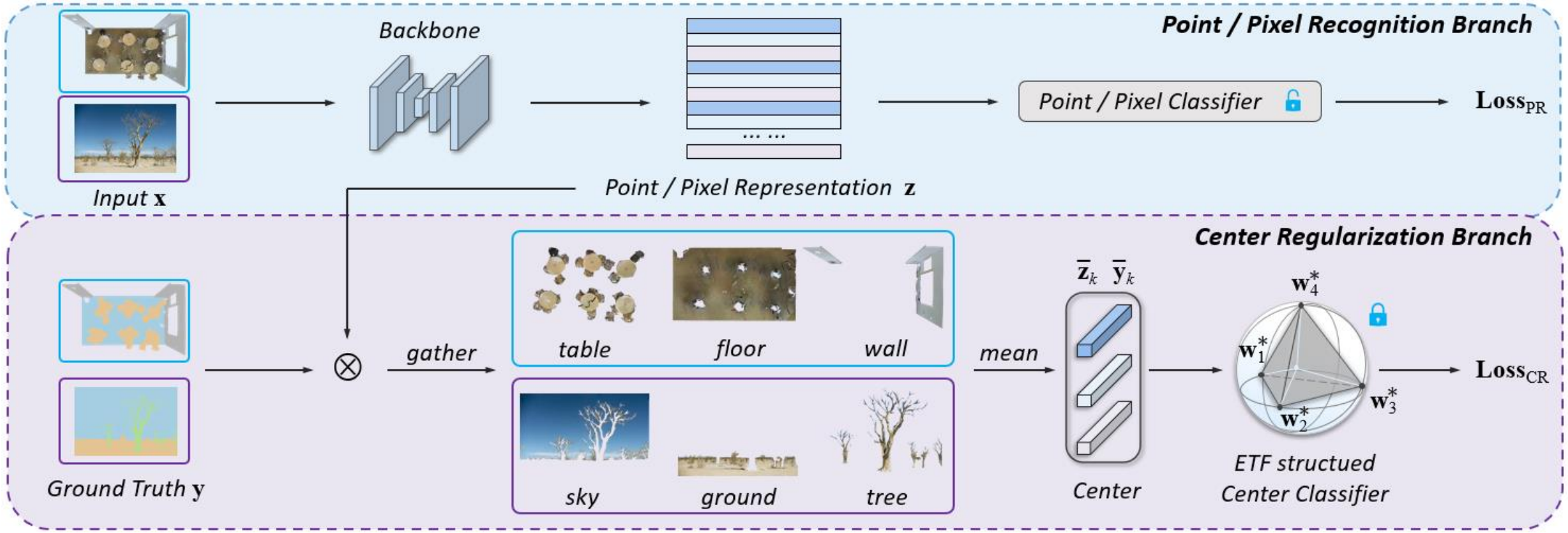


Feature centers and classifiers *converge to a* ***non-equiangular*** *and* ***non-maximal separated*** *structure*

However, for imbalanced semantic segmentation:

This asymmetric structure is detrimental to discrimination among the tailed/minor classes.

CeCo (Ours)



*We introduce a **center collapse (CeCo) loss** to regularize the feature centers to close the **equiangular and maximal separated ETF structure**.*

ScanNet200 Benchmark

https://kaldir.vc.in.tum.de/scannet_benchmark

ScanNet200 3D Semantic Label Benchmark

This table lists the benchmark results for the ScanNet200 3D semantic label scenario.

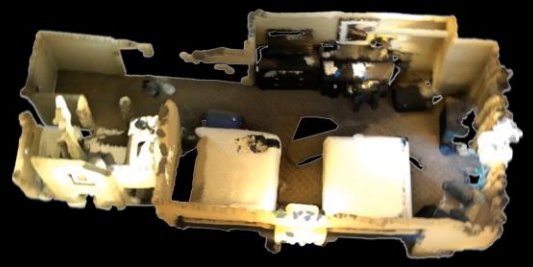
Column Sorting: Column Sort Alphabetically

Method	Info	avg iou	head iou	common iou	tail iou	alarm clock	armchair	backpack	bag	ball	bar	basket	bathroom cabinet	bathroo
CeCo		**0.340** 1	**0.551** 1	**0.247** 1	**0.181** 1			0.475 2	0.057 4	0.142 3	**0.000** 1	**0.000** 1	**0.000** 1	
: Understanding Imbalanced Semantic Segmentation Through Neural Collapse.														
LGround	P	0.272 2	0.485 2	0.184 2	0.106 2			**0.476** 1	0.077 2	**0.218** 1	**0.000** 1	**0.000** 1	**0.000** 1	
David Rozenberszki, Or Litany, Angela Dai: Language-Grounded Indoor 3D Semantic Segmentation in the Wild. arXiv														
Minkowski 34D	P	0.253 3	0.463 3	0.154 4	0.102 3			0.381 4	**0.084** 1	0.134 4	**0.000** 1	**0.000** 1	**0.000** 1	
C. Choy, J. Gwak, S. Savarese: 4D Spatio-Temporal ConvNets: Minkowski Convolutional Neural Networks. CVPR 2019														
CSC-Pretrain	P	0.249 4	0.455 4	0.171 3	0.079 4			0.418 3	0.059 3	0.186 2	**0.000** 1	**0.000** 1	**0.000** 1	
Ji Hou, Benjamin Graham, Matthias Nießner, Saining Xie: Exploring Data-Efficient 3D Scene Understanding with Contrastive Scene Contexts. CVPR 2021														

Rank 1st

Improve a lot (+6.8% mIoU)

Org.



GT

CE

CeCo

Org.

GT

CE

CeCo

2D Visualization Example:

2D Visualization Example:

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Thanks for your watching!

For more details, please visit:

Our paper

Our code