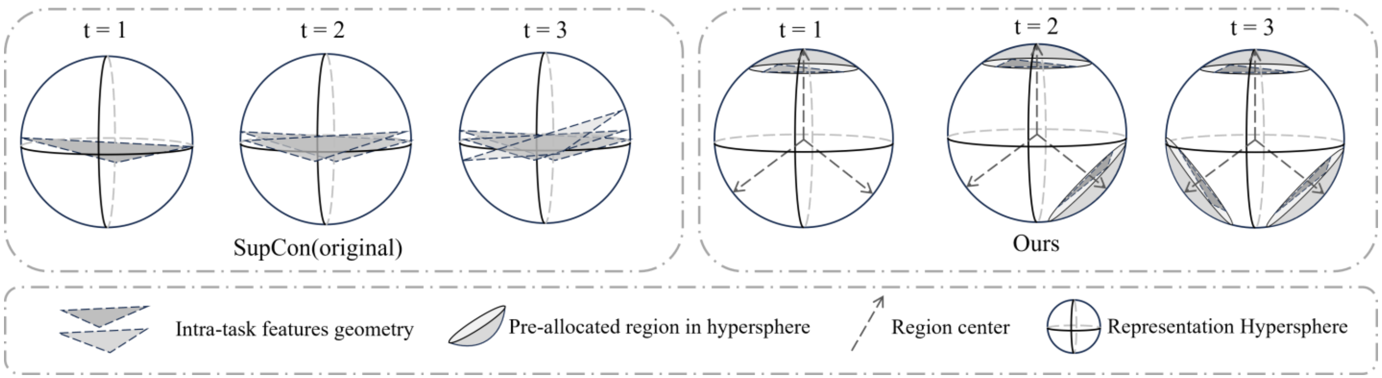
Global Pre-fixing, Local Adjusting: A Simple yet Effective Contrastive Strategy for Continual Learning

Jia Tang, Xinrui Wang, Songcan CHEN

Frontiers of Computer Science, DOI: 10.1007/s11704-025-50623-6

Problems & Ideas

- Problems of continual learning:
  - Performance is still limited due to confusion arising from both inter-task and intra-task features.
  - Existing methods leverages SupCon to construct intra-task feature geometry with better transferability and discriminability, but they cannot form a task-level well-separated feature geometry spontaneously.
- Ideas:
  - **For inter-task features: global pre-fixing.** Pre-allocate non-overlapping feature regions for different tasks from the beginning.
  - **For intra-task features: local adjusting.** Optimize each task's class features exclusively within its dedicated region to achieve locally optimal discrimination.



The difference between our method and SupCon in CL. (left) Due to the sequential arrival of tasks and the (global) invisibility of inter-task data, the SupCon cannot spontaneously form a task-level well-separated feature structure in CL. (right) We pre-divide the unit representation hypersphere into non-overlapping regions according to the pre-estimated maximum task counts, with the center of the regions forming a task-level pre-fixed ETF. The proposed loss helps adjust feature structure and form an intra-task ETF in the allocated region.

Main Contributions

- Contributions:
  - We address feature confusion in contrastive continual learning by introducing a global prefixing and local adjustment strategy that separates inter-task and intra-task feature learning.
  - Building on SupCon and the ETF principle, we design a new contrastive loss that shapes intra-task ETF structures within pre-allocated regions, with theoretical justification.
  - Extensive experiments with multiple baselines show our method consistently improves CIL performance across datasets and buffer sizes.

Buffer	Dataset	Seq-Cifar-10		Seq-Cifar-100		Seq-Tiny-ImageNet	
	Scenario	Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL
200	ER	49.16±2.08	91.92±1.01	21.78±0.48	60.19±1.01	8.65±0.16	38.83±1.15
	iCaRL	32.44±0.93	74.59±1.24	28.0±0.91	51.43±1.47	5.5±0.52	22.89±1.83
	GEM	29.99±3.92	88.67±1.76	20.75±0.66	58.84±1.00	-	-
	GSS	38.62±3.59	90.0±1.58	19.42±0.29	55.38±1.34	8.57±0.13	31.77±1.34
	DER	63.69±2.35	91.91±0.51	31.23±1.38	63.09±1.09	13.22±0.92	42.27±0.90
	GCR	64.84±1.63	90.8±1.05	33.69±1.40	64.24±0.83	13.05±0.91	42.11±1.01
	[ICCV, 2021] Co2L	65.57±1.37	93.43±0.78	27.73±0.54	54.33±0.36	13.88±0.40	42.37±0.74
	Co2L + Ours	**70.59±1.26**	**95.55±0.37**	**32.48±1.08**	**62.01±1.24**	**14.30±0.51**	**44.53±0.83**
	[ICML, 2024] CILA	67.06±1.59	94.29±0.24	30.18±0.39	58.19±0.28	14.55±0.39	**44.15±0.70**
	CILA + Ours	**71.8±1.35**	**96.41±0.86**	**33.13±1.34**	**62.13±1.88**	**15.51±0.47**	44.12±0.65
	[AAAI, 2024] CCLIS	74.95±0.61	96.20±0.26	42.39±0.37	**72.93±0.46**	16.13±0.19	48.29±0.78
	CCLIS + Ours	**76.33±1.14**	**96.73 ±0.48**	**44.48±1.27**	72.91±1.45	**17.16±0.34**	**48.80±0.68**

Class-IL and Task-IL Continual Learning. We report our performance and the results of rehearsal-based baselines on Seq-Cifar-10, Seq-Cifar-100 and Seq-Tiny-ImageNet with memory sizes 200, all of which are averaged across five independent trails.