Class-incremental Novel Class Discovery

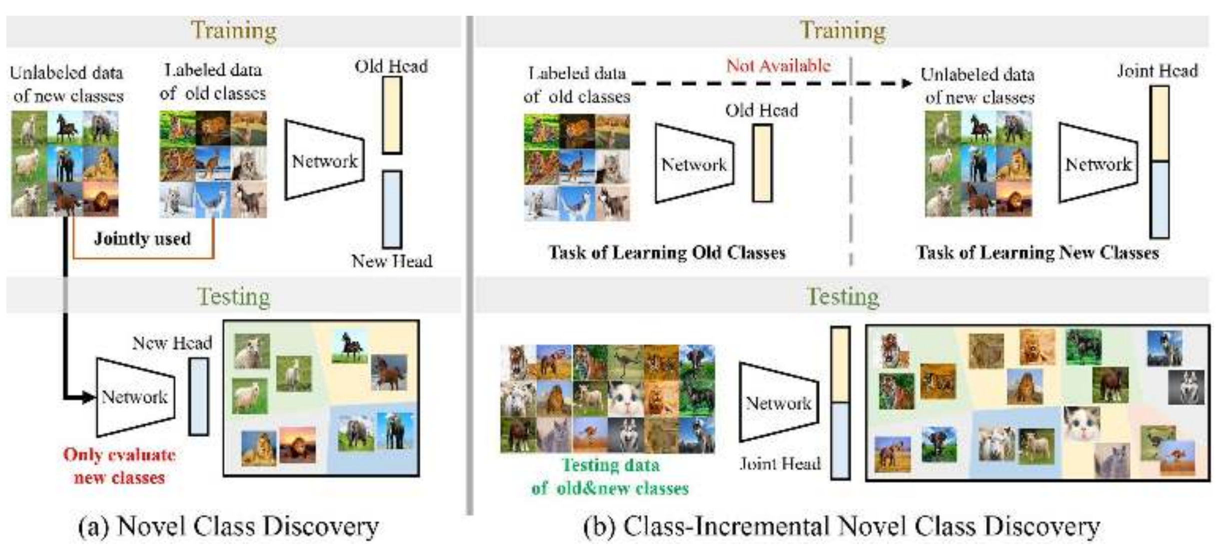
Subhankar Roy† Mingxuan Liu† Zhun Zhong Nicu Sebe Elisa Ricci

†equal contribution

University of Trento Fondazione Bruno Kessler

Problem

Class-incremental Novel Class Discovery (class-iNCD): Discovering novel categories in an unlabelled data set by leveraging a pre-trained model that has been trained on a labelled data set containing disjoint yet related categories, while preserving the ability of the model to recognize previously seen categories, without access to the previously seen data and task-id of an input sample during inference.



(a) Novel Class Discovery (b) Class-Incremental Novel Class Discovery

Motivation

- To facilitate learning of novel classes, we dedicate a task specific classifier that is optimized with robust rank statistics:

$$\mathcal{L}_{\text{bce}} = -\mathbb{E}_{p(\mathbf{z}^{[\text{U}]})}\tilde{y}_{ij}^{[\text{U}]}\log(p_{ij}) + (1-\tilde{y}_{ij}^{[\text{U}]})\log(1-p_{ij})$$

- To overcome reliance on task-id, we propose to maintain a joint classifier for both the base and novel classes, which is trained with the pseudo-labels generated by the task specific one:

$$\mathcal{L}_{\text{self}} = -\mathbb{E}_{(\mathbf{x}^{[\text{U}]},\hat{\mathbf{y}}^{[\text{U}]})}\frac{1}{|C^{[\text{A}]}|}\sum_{k=1}^{|C^{[\text{A}]}|}\hat{y}_k^{[\text{U}]}\log\sigma_k(h^{[\text{A}]}(g(\mathbf{x}^{[\text{U}]})))$$

$$\hat{y}^{[\text{U}]} = C^{[\text{L}]} + \arg\max_{k\in C^{[\text{U}]}} h^{[\text{U}]}(g(\mathbf{x}^{[\text{U}]})).$$

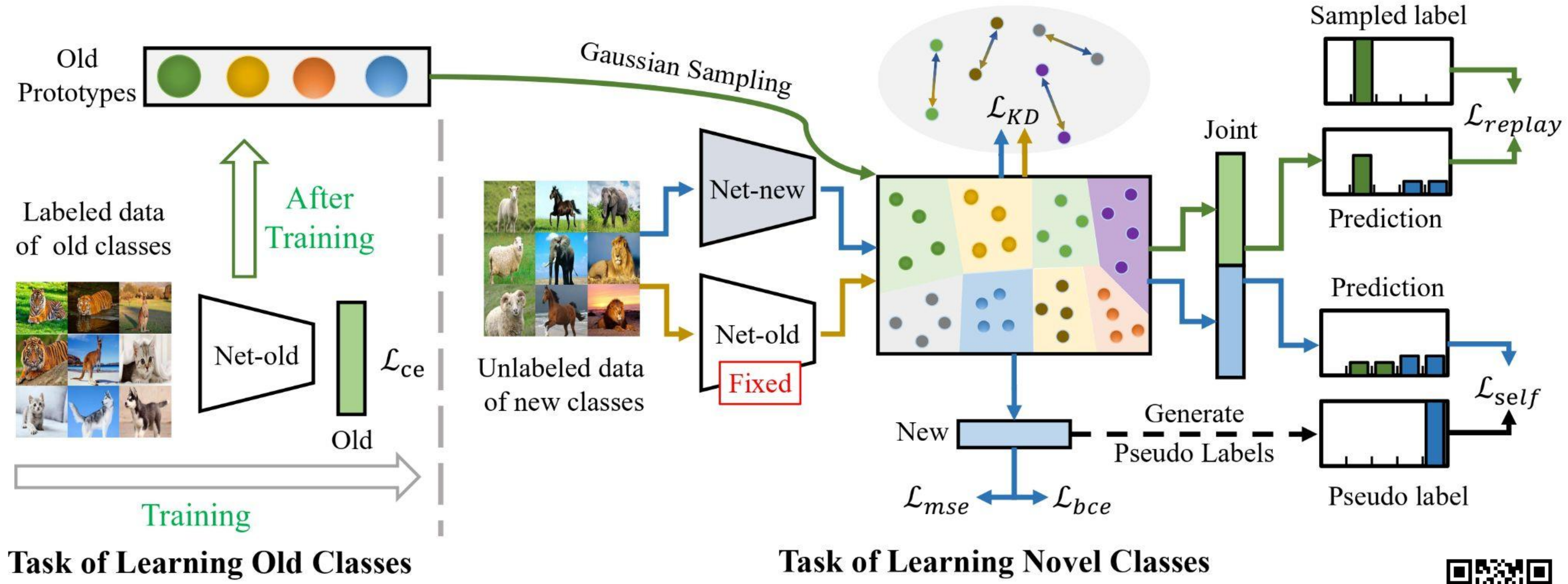
- We propose to store the base class feature prototypes from the previous task as exemplars. Features derived from the stored prototypes are then replayed to prevent forgetting old information on the base classes:

$$\mathcal{L}_{\text{replay}} = -\mathbb{E}_{c\sim C^{[\text{L}]}}\mathbb{E}_{(\mathbf{z}^{[\text{L}]},\mathbf{y}_c^{[\text{L}]})\sim\mathcal{N}(\mu_c,v_c^2)}\sum_{k=1}^{|C^{[\text{A}]}|}y_{kc}^{[\text{L}]}\log\sigma_k(h^{[\text{A}]}(\mathbf{z}^{[\text{L}]}))$$

- To keep the feature replay useful, we add an regularization on the current feature extractor:

$$\mathcal{L}_{\text{KD}}^{\text{feat}} = -\mathbb{E}_{p(\mathbf{x}^{[\text{U}]})}\left|\left|g^{[\text{L}]}(\mathbf{x}^{[\text{U}]}) - g(\mathbf{x}^{[\text{U}]})\right|\right|_2$$

Overall Framework

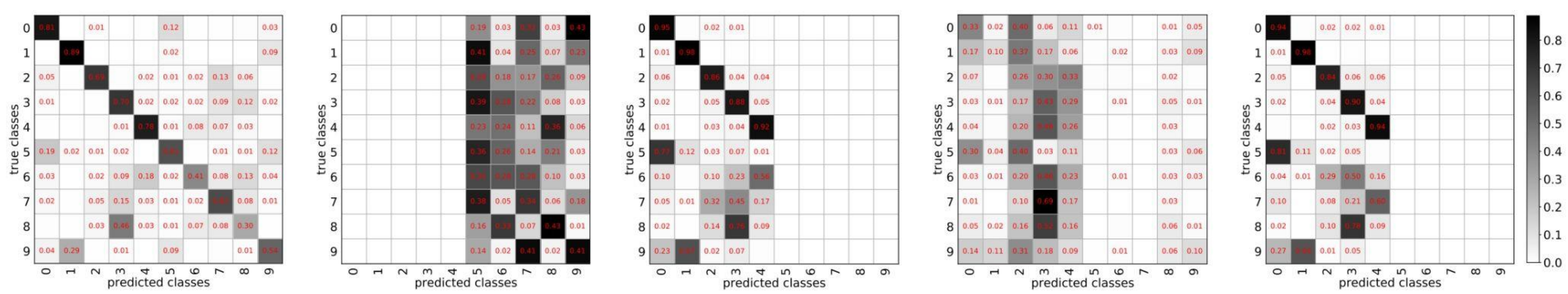


Feature **R**eplay and Distillation with **S**elf-Training (**FRoST**)

Code:

Single-step State-of-The-Arts & Visualization

Methods	CIFAR-10			CIFAR-100			Tiny-ImageNet			Average		
	Old	New	All	Old	New	All	Old	New	All	Old	New	All
AutoNovel[15]	27.5	3.5	15.5	2.6	15.2	5.1	2.0	26.4	4.5	10.7	15.0	8.4
ResTune[29]	91.7	0.0	45.9	**73.8**	0.0	59.0	44.3	0.0	39.9	**69.9**	0.0	48.3
NCL[35]	**92.0**	1.1	46.5	73.6	10.1	**60.9**	0.8	6.5	1.4	55.5	5.9	36.3
DTC[14]	64.0	0.0	32.0	55.9	0.0	44.7	35.5	0.0	32.0	51.8	0.0	36.2
FRoST	77.5	**49.5**	**63.4**	64.6	**45.8**	59.2	**54.5**	**33.7**	**52.3**	65.5	**39.8**	**54.9**



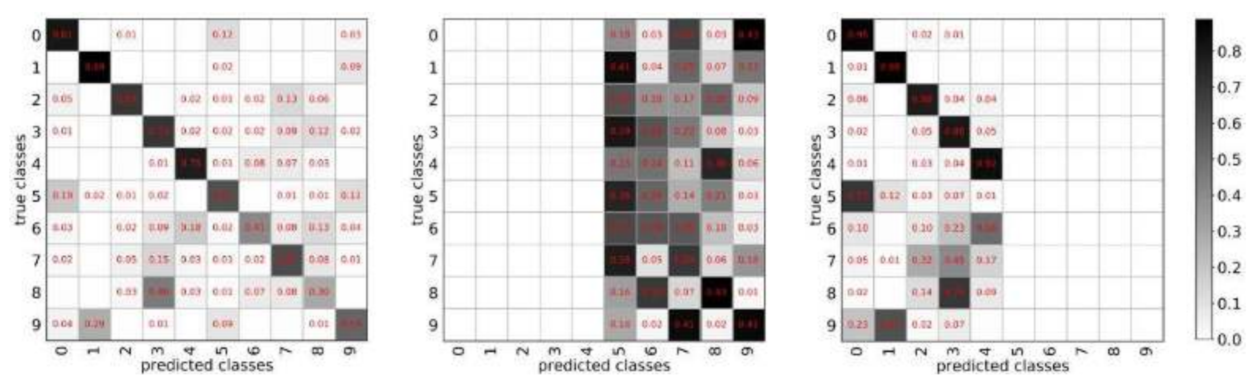
(a) Ours (b) Ours w/o FD&FR (c) Ours w/o ST (d) AutoNovel (e) ResTune

Two-step State-of-The-Arts

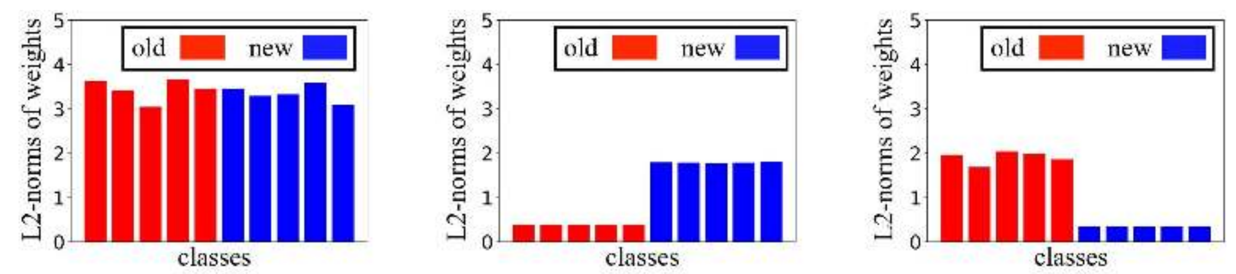
Methods	Tiny-ImageNet									
	First Step (180-10)				Second Step (180-10-10)					
	Old	New-1-J	New-1-N	All	Old	New-1-J	New-2-J	New-1-N	New-2-N	All
ResTune[29]	39.7	0.0	38.0	37.6	34.9	0.0	0.0	25.4	42.8	31.4
DTC[14]	38.9	0.0	**43.8**	36.9	33.4	0.0	0.0	28.0	**59.4**	30.1
NCL[35]	5.6	0.0	34.2	5.3	1.4	0.0	2.6	21.6	41.6	1.4
FroST	**55.2**	**27.6**	32.0	**53.8**	**42.5**	**34.8**	**31.2**	**31.2**	46.8	**41.6**

Ablation Study & Visualization

Methods	CIFAR-10			CIFAR-100			Tiny-ImageNet			Average		
	Old	New	All	Old	New	All	Old	New	All	Old	New	All
FRoST (Ours)	77.4	49.5	**63.5**	62.5	45.8	**59.2**	54.4	33.9	**52.4**	64.8	43.1	**58.3**
w/o FD & FR	0.0	36.4	18.2	0.0	33.1	6.6	0.0	37.2	3.7	0.0	35.6	9.5
w/o FD	0.0	39.4	19.7	0.0	33.1	6.6	0.0	34.3	3.4	0.0	35.6	9.9
w/o FR	0.0	**73.3**	36.6	0.0	**57.8**	11.6	0.0	**40.9**	4.1	0.0	**57.3**	17.4
w/o ST	**91.7**	0.0	45.8	**69.2**	0.0	55.4	**57.5**	0.0	51.7	**72.8**	0.0	51.0
w/o FD & FR & ST	16.6	0.0	8.3	2.7	0.0	2.1	2.0	0.0	1.8	7.1	0.0	4.1



(a) Ours (b) Ours w/o FD&FR (c) Ours w/o ST



(a) Ours (b) Ours w/o FD&FR (c) Ours w/o ST

Contribution

- We propose a novel framework, FRoST, that can tackle the newly introduced and relevant task of class-incremental novel class discovery (class-iNCD).
- Our FRoST is equipped with prototypes for feature-replay and employs feature-level knowledge distillation to prevent forgetting. Moreover, it uses pseudo-labels from the task specific head to efficiently learn novel classes without interference, enabling us to achieve a task-agnostic classifier.
- We run extensive experiments on three common benchmarks to prove the effectiveness of our method. FRoST also obtains state-of-the-art performance when compared with the existing baselines. Additionally, we run experiments on a sequence of tasks of unlabelled sets and verify its generality.