JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

南京大學 NANJING UNIVERSITY

R&L Group

Modeling Inter-Class and Intra-Class Constraints in Novel Class Discovery

Wenbin Li, Zhichen Fan, Jing Huo, Yang Gao

State Key Laboratory for Novel Software Technology, Nanjing University, China

Poster: TUE-AM-328

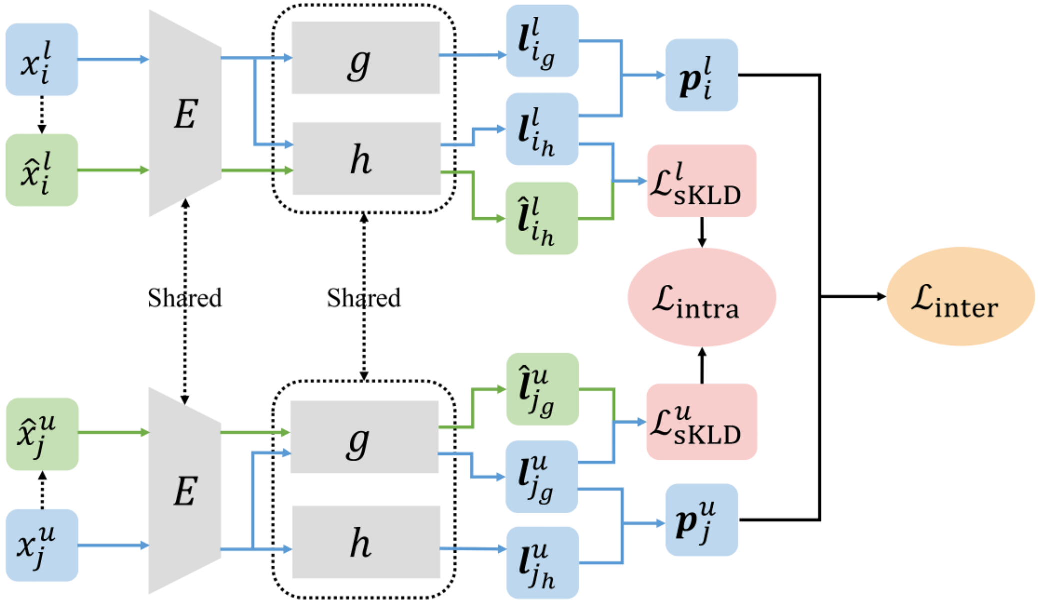
Overview

Task

- Novel Class Discovery (NCD) aims to discover new classes in an unlablled dataset with the latent common knowledge transferred from another class-disjoint labelled dataset.

Contributions

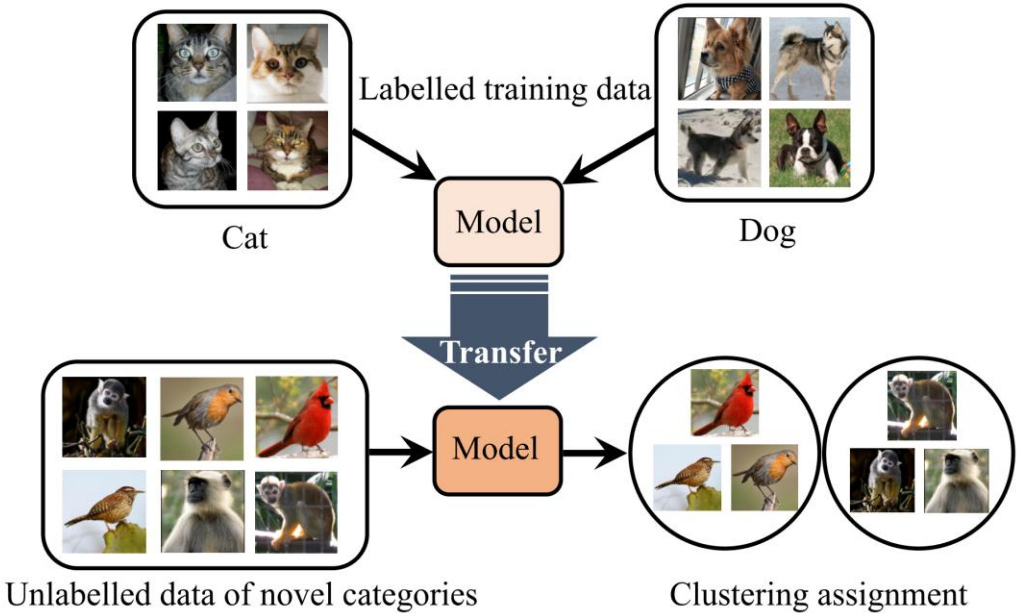
- Propose two symmetric Kullback-Leibler divergence (sKLD) based constraints from both inter-class and intra-class perspectives to learn more discriminative features for NCD.
- The proposed constraints achieve state-of-the-art results on three benchmark datasets.



Problem Definition

Novel Class Discovery (NCD)

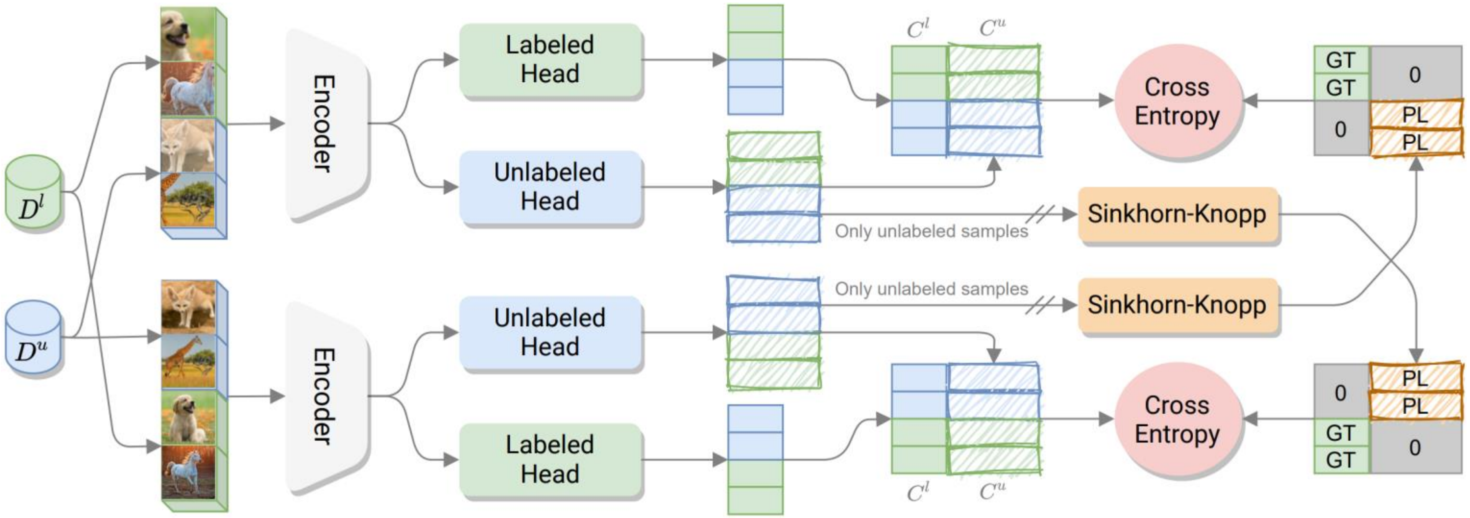
- Goal: Discover new classes in an unlablled dataset with the latent common knowledge transferred from another class-disjoint labelled dataset.
- Essence of NCD: The set of labelled categories is related but non-overlapped with the set of unlabelled categories.



The NCD setting [1].

Motivations

- Current single-stage based methods [2, 3] overlook the disjoint characteristic between labelled and unlabelled classes.
- Some previous methods [4, 5] employ the MSE as consistency regularization, which cannot perform well in practice.



Overview of UNO [2].

Method

Inter-class Symmetric KLD Constraint

- Enlarge the distance between each labelled sample and each unlabelled sample in a mini-batch.

$$\mathcal{L}_{\text{sKLD}} = \frac{1}{2}\left(D_{\text{KL}}(\boldsymbol{p}_i^l || \boldsymbol{p}_j^u) + D_{\text{KL}}(\boldsymbol{p}_j^u || \boldsymbol{p}_i^l)\right)$$

$$\mathcal{L}_{\text{inter-class}} = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\mathcal{L}_{\text{sKLD}}$$

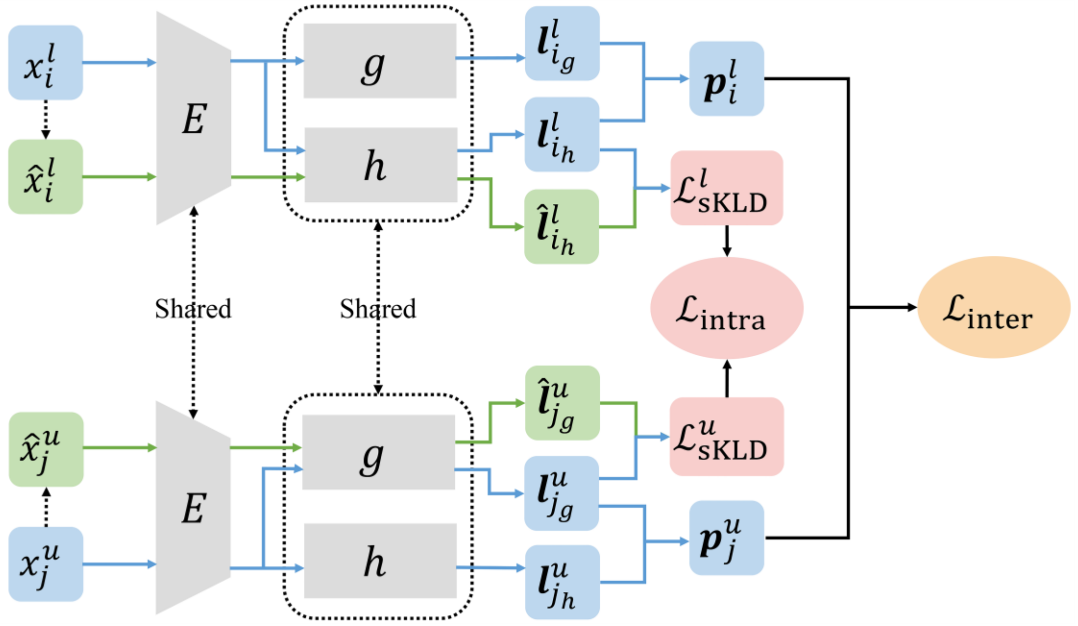
Intra-class Symmetric KLD Constraint

- The distance between any two probability distributions of different augmentations of the same image should be small.

$$\mathcal{L}_{\text{sKLD}}^l = \frac{1}{2}\left(D_{\text{KL}}(\boldsymbol{p}_{i_h}^l || \hat{\boldsymbol{p}}_{i_h}^l) + D_{\text{KL}}(\hat{\boldsymbol{p}}_{i_h}^l || \boldsymbol{p}_{i_h}^l)\right)$$

$$\mathcal{L}_{\text{sKLD}}^u = \frac{1}{2}\left(D_{\text{KL}}(\boldsymbol{p}_{j_g}^u || \hat{\boldsymbol{p}}_{j_g}^u) + D_{\text{KL}}(\hat{\boldsymbol{p}}_{j_g}^u || \boldsymbol{p}_{j_g}^u)\right)$$

$$\mathcal{L}_{\text{intra-class}} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{\text{sKLD}}^l + \frac{1}{M}\sum_{j=1}^{M}\mathcal{L}_{\text{sKLD}}^u$$



Overall Objective

- Maximize the inter-class sKLD and minimize the intra-class sKLD simultaneously.

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N+M}\sum_{i=1}^{N+M}\sum_{k=1}^{C^l+C^u}\boldsymbol{y}_i(k)\log\boldsymbol{p}_i(k)$$

$$\mathcal{L} = \mathcal{L}_{\text{CE}} - \alpha\mathcal{L}_{\text{inter-class}} + \beta\mathcal{L}_{\text{intra-class}}$$

Experimental Setup

Datasets

- Three datasets: CIFAR10, CIFAR100 and ImageNet.
- Assume the number of classes in the unlabelled subset is known a priori.

Dataset split	Labelled		Unlabelled	
	#Images	#Classes	#Images	#Classes
CIFAR10	25K	5	25K	5
CIFAR100-20	40K	80	10K	20
CIFAR100-50	25K	50	25K	50
ImageNet	1.25M	882	≈30K	30

Details of dataset splits used in the experiments.

Evaluation Metric

- Evaluation protocols: Task-aware and task-agnostic
- Evaluation criterion: Average clustering accuracy (ACC)

$$\mathcal{ACC} = \max_{\mathrm{perm} \in P} \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{y_i = \mathrm{perm}(\hat{y}_i)\}$$

Ablation Study

Symmetric KLD vs. MSE for Intra-class Constraint

Method	CIFAR10	CIFAR100-20	CIFAR100-50
Baseline	93.25	90.54	62.27
+ MSE	93.11	90.96	62.47
+ sKLD	**93.60**	**91.13**	**63.35**

Symmetric KLD Constraints

- Extra evaluation criteria: Normalized mutual information (NMI) and adjusted rand index (ARI)

#	Inter-class	Intra-class	CIFAR10			CIFAR100-20			CIFAR100-50		
			ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
1	✗	✗	93.25	0.8717	0.8565	90.54	0.8475	0.8123	62.27	0.6781	0.4688
2	✓	✗	99.07	0.9644	0.9770	92.17	0.8695	0.8448	64.31	0.7046	0.5102
3	✗	✓	93.60	0.8753	0.8624	91.13	0.8506	0.8193	63.35	0.6800	0.4729
4	✓	✓	**99.11**	**0.9657**	**0.9780**	**92.48**	**0.8727**	**0.8508**	**65.85**	**0.7106**	**0.5238**

Comparison with State of the Arts

Results on Task-aware Protocol

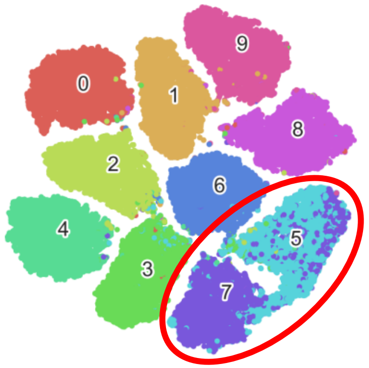
Method	Venue	Type	CIFAR10	CIFAR100-20	CIFAR100-50	ImageNet
k-means	Classic	-	72.5±0.0	56.3±1.7	28.3±0.7	71.9
KCL	ICLR'18	Two-stage	72.3±0.2	42.1±1.8	–	73.8
MCL	ICLR'19	Two-stage	70.9±0.1	21.5±2.3	–	74.4
DTC	ICCV'19	Two-stage	88.7±0.3	67.3±1.2	35.9±1.0	78.3
RS	ICLR'20	Single-stage	90.4±0.5	73.2±2.1	39.2±2.3	82.5
RS+	ICLR'20	Single-stage	91.7±0.9	75.2±4.2	44.1±3.7	82.5
OpenMix	CVPR'21	Single-stage	95.3	–	–	85.7
NCL	CVPR'21	Single-stage	93.4±0.5	86.6±0.4	–	90.7
Joint	ICCV'21	Single-stage	93.4±0.6	76.4±2.8	–	86.7
UNOv1	ICCV'21	Single-stage	96.1±0.5	85.0±0.6	52.9±1.4	90.6
UNOv2†	ICCV'21	Single-stage	93.3±0.4	90.5±0.7	62.3±1.4	90.7
DualRank	NeurIPS'21	Single-stage	91.6±0.6	75.3±2.3	–	88.9
ComEx	CVPR'22	Single-stage	93.6±0.3	85.7±0.7	53.4±1.3	90.9
IIC (Ours)	CVPR'23	Single-stage	**99.1±0.0**	**92.4±0.2**	**65.8±0.9**	**91.9**

Results on Task-agnostic Protocol

Method	CIFAR10			CIFAR100-20			CIFAR100-50		
	Label	Unlabel	All	Label	Unlabel	All	Label	Unlabel	All
KCL	79.4	60.1	69.8	23.4	29.4	24.6	–	–	-
MCL	81.4	64.8	73.1	18.2	18.0	18.2	–	–	-
DTC	58.7	78.6	68.7	47.6	49.1	47.9	30.2	34.7	32.5
RS+	90.6	88.8	89.7	71.2	56.8	68.3	69.7	40.9	55.3
UNOv1	93.5	93.3	93.4	73.2	73.1	73.2	71.5	50.7	61.1
ComEx	95.0	92.6	93.8	75.2	77.3	75.6	**75.3**	53.5	64.4
IIC (Ours)	**96.0**	**97.2**	**96.6**	**75.9**	**78.4**	**77.2**	75.1	**61.0**	**68.1**

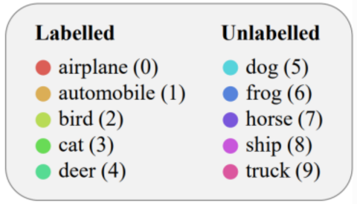
Visualization

t-SNE Visualization

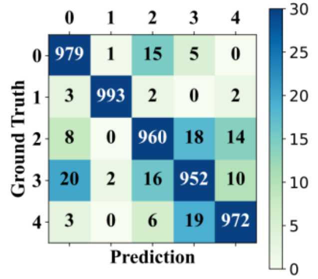


(a) UNOv2

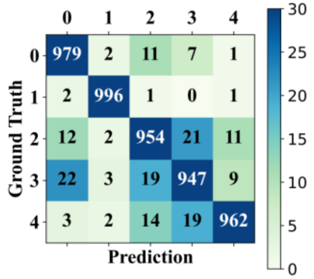
(b) Ours



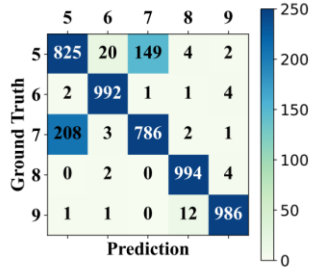
Confusion Matrix on Task-aware/-agnostic Protocol



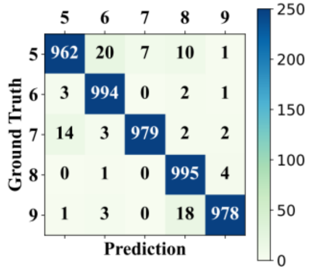
(a) UNOv2 on labelled subset



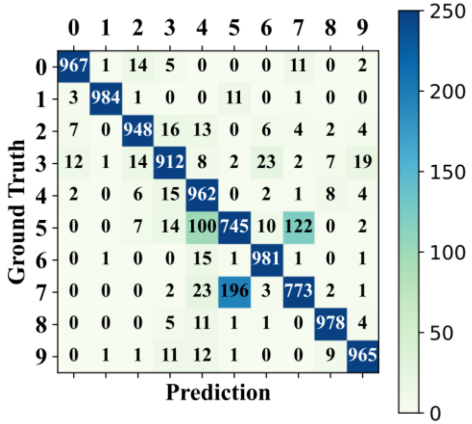
(b) UNOv2+ on labelled subset



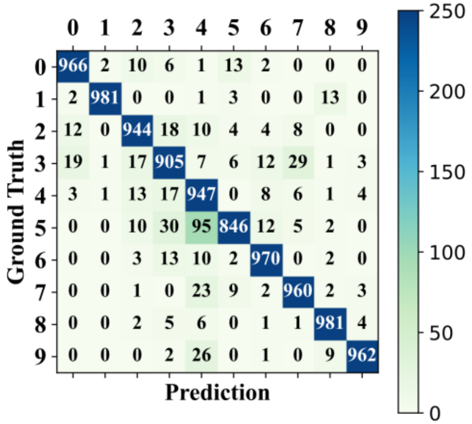
(c) UNOv2 on unlabelled subset



(d) UNOv2+ on unlabelled subset



(a) UNOv2



(b) UNOv2+

Conclusion

- We propose to model both inter-class and intra-class constraints built on the symmetric Kullback-Leibler divergence (sKLD) for novel class discovery (NCD).
- We conduct extensive experiments on four popular benchmarks and show that our method could outperform the existing state-of-the-art methods by a large margin.
- From the experimental results, we have the following findings: (1) making use of the disjoint characteristic between the labelled and unlabelled classes, i.e., constraining an inter-class constraint, is important and effective for NCD; (2) using a slacker sKLD measure instead of the MSE for constraining the intra-class constraint is reasonable and beneficial to NCD.

References

[1] Kai Han, Andrea Vedaldi, and Andrew Zisserman. Learning to discover novel visual categories via deep transfer clustering. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 8401–8409, 2019.
[2] Enrico Fini, Enver Sangineto, Stephane Lathuiliere, Zhun Zhong, Moin Nabi, and Elisa Ricci. A unified objective for novel class discovery. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 9284–9292, 2021.
[3] Zhun Zhong, Enrico Fini, Subhankar Roy, Zhiming Luo, Elisa Ricci, and Nicu Sebe. Neighborhood contrastive learning for novel class discovery. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 10867–10875, 2021.
[4] Kai Han, Sylvestre-Alvise Rebuffi, Sebastien Ehrhardt, Andrea Vedaldi, and Andrew Zisserman. Automatically discovering and learning new visual categories with ranking statistics. In Proceedings of the International Conference on Learning Representations (ICLR), 2020.
[5] Bingchen Zhao and Kai Han. Novel visual category discovery with dual ranking statistics and mutual knowledge distillation. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), volume 34, pages 22982–22994, 2021.

Thanks

Paper: https://arxiv.org/abs/2210.03591

Code: https://github.com/FanZhichen/NCD-IIC