Label Hallucination for Few-Shot Classification

Yiren Jian, Lorenzo Torresani

DARTMOUTH
Department of Computer Science

AAAI-22

Few-Shot Learning

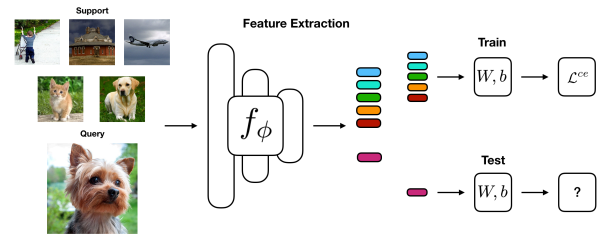
Few-shot learning aims at adapting knowledge extracted from data-rich base categories to novel categories where examples are limited [1,2].

Visualization and Intuition

Our method trains the model on the large-scale base dataset automatically re-labeled according to the novel classes.

- ➡ First row shows one-shot examples of 5 novel classes. Underneath each one-shot image, we show the 3 base images receiving the largest pseudo-label scores.
- ➡ Our method assigns novel-class labels to base images that match the few-shot examples in terms of background, shape, spatial layout, color, or texture.

Prior Work

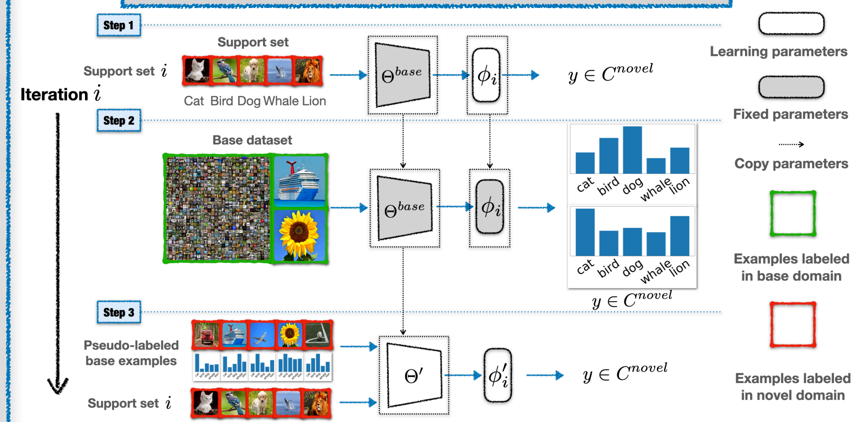


- ➡ Train a large-capacity model using a multi-way classification loss on the base dataset to learn a discriminative representation.

Then:

- ➡ Method 1: Train a linear classifier on top of the frozen representation for each set of novel classes [3]. Weakness: limited learning capacity
- ➡ Method 2: Finetune entire model on the novel set [4]. Weakness: high risk of overfitting

Label Hallucination



1. Learn linear classifier for novel classes on top of pretrained backbone.
2. Pseudo-label the base dataset using the label space of the novel classes.
3. Finetune the entire model on the novel examples and the pseudo-labeled base dataset.

Experiments

miniImageNet and tieredImageNet

model	backbone	miniImageNet 5-way 1-shot	miniImageNet 5-way 5-shot	tieredImageNet 5-way 1-shot	tieredImageNet 5-way 5-shot
DeepEMD [56] (CVPR'20)	ResNet-12	65.91 ± 0.82	82.41 ± 0.56	71.16 ± 0.87	86.03 ± 0.58
RFS-simple [48] (ECCV'20)	ResNet-12	62.02 ± 0.63	79.64 ± 0.44	69.74 ± 0.72	84.41 ± 0.55
RFS-distill [48] (ECCV'20)	ResNet-12	64.82 ± 0.82	82.41 ± 0.43	71.52 ± 0.69	86.03 ± 0.49
AssoAlign [1] (ECCV'20)	ResNet-18†	59.88 ± 0.67	80.35 ± 0.73	69.29 ± 0.56	85.97 ± 0.49
AssoAlign [1] (ECCV'20)	WRN-28-10‡	65.92 ± 0.60	82.85 ± 0.55	74.40 ± 0.68	86.61 ± 0.59
SKD-GEN1 [35] (Arxiv'20)	ResNet-12	$66.54 \pm 0.97^\S$	$83.18 \pm 0.54^\S$	$72.35 \pm 1.23^\S$	$85.97 \pm 0.63^\S$
MELR [14] (ICLR'21)	ResNet-12	67.40 ± 0.43	83.40 ± 0.28	72.14 ± 0.51	87.01 ± 0.35
IEPT [57] (ICLR'21)	ResNet-12	67.05 ± 0.44	82.90 ± 0.30	72.24 ± 0.50	86.73 ± 0.34
IER-distill [39] (CVPR'21)	ResNet-12	$66.85 \pm 0.76^\S$	$84.50 \pm 0.53^\S$	$72.74 \pm 1.25^\S$	$86.57 \pm 0.81^\S$
Label-Halluc (pretrained w/ SKD-GEN1)	ResNet-12	67.50 ± 1.01	85.60 ± 0.52	72.80 ± 1.20	86.93 ± 0.60
Label-Halluc (pretrained w/ IER-distill)	ResNet-12	$\mathbf{68.28 \pm 0.77}$	$\mathbf{86.54 \pm 0.46}$	$\mathbf{73.34 \pm 1.25}$	$\mathbf{87.68 \pm 0.83}$

Comparison of our method (Label-Halluc) against the state-of-the-art on miniImageNet and tieredImageNet.

CIFAR-FS and FC100

model	backbone	CIFAR-FS 5-way 1-shot	CIFAR-FS 5-way 5-shot	FC-100 5-way 1-shot	FC-100 5-way 5-shot
DeepEMD [56] (CVPR'20)	ResNet-12	-	-	46.5 ± 0.8	63.2 ± 0.7
RFS-simple [48] (ECCV'20)	ResNet-12	71.5 ± 0.8	86.0 ± 0.5	42.6 ± 0.7	59.1 ± 0.6
RFS-distill [48] (ECCV'20)	ResNet-12	73.9 ± 0.8	86.9 ± 0.5	44.6 ± 0.7	60.9 ± 0.6
AssoAlign [1] (ECCV'20)	ResNet-18‡	-	-	45.8 ± 0.5	59.7 ± 0.6
SKD-GEN1 [35] (Arxiv'20)	ResNet-12	$76.6 \pm 0.9^\S$	$88.6 \pm 0.5^\S$	$46.5 \pm 0.8^\S$	$64.2 \pm 0.8^\S$
InfoPatch [18] (AAAI'21)	ResNet-12	-	-	43.8 ± 0.4	58.0 ± 0.4
IER-distill [39] (CVPR'21)	ResNet-12	$77.6 \pm 1.0^\S$	$89.7 \pm 0.6^\S$	$48.1 \pm 0.8^\S$	$65.0 \pm 0.7^\S$
Label-Halluc (pretrained w/ SKD-GEN1)	ResNet-12	77.3 ± 0.9	89.5 ± 0.5	47.3 ± 0.8	67.2 ± 0.8
Label-Halluc (pretrained w/ IER-distill)	ResNet-12	$\mathbf{78.0 \pm 1.0}$	$\mathbf{90.5 \pm 0.6}$	$\mathbf{49.1 \pm 0.8}$	$\mathbf{68.0 \pm 0.7}$

Comparison of our method (Label-Halluc) against the state-of-the-art on CIFAR-FS and FC-100

Ablation: Learning embedding or classifier with LabelHalluc

- **Support**: learning with support set.
- **Base**: learning with base set
- **Net**: updating network encoder
- **Clf**: updating the last classifier

Support Net	Support Clf	Base Net	Base Clf	miniImageNet 1-shot	miniImageNet 5-shot
✓	✓			61.43	80.03
✓	✓		✓	63.59	81.53
✓	✓	✓		66.18	84.36
✓	✓	✓	✓	**67.50**	**85.60**

The largest improvements come from learning the capacity embedding network, and fine-tuning both the embedding and classifier yields best results.

Ablation: Soft or Hard Pseudo-labels

	mini-IN 1-shot	mini-IN 5-shot	CIFAR-FS 1-shot	CIFAR-FS 5-shot	FC100 1-shot	FC100 5-shot
Transfer w/ frozen backbone (LR)	66.54	83.18	76.6	88.6	46.5	64.2
Transfer w/ finetuning	61.43	80.03	68.8	85.7	43.1	61.9
Hard LabelHalluc + finetuning	65.04	80.68	75.3	85.3	44.6	62.4
Soft LabelHalluc + finetuning	**67.50**	**85.60**	**77.3**	**89.5**	**47.3**	**67.2**

Soft Label Hallucination works the best, outperforming Hard Label Hallucination and the frozen backbone baseline.

Ablation: Different Pre-training

	miniImageNet LR	miniImageNet ours	CIFAR-FS LR	CIFAR-FS ours	FC100 LR	FC100 ours
RFS-simple [48]	79.33	**81.75**	86.6	**87.3**	58.1	**61.2**
RFS-distill [48]	81.15	**82.74**	86.5	**87.1**	61.0	**63.9**
SKD-gen0 [35]	82.31	**84.14**	87.8	**88.8**	62.8	**66.5**
SKD-gen1 [35]	83.18	**85.60**	88.6	**89.5**	64.2	**67.2**
IER-gen0 [39]	83.88	**85.86**	89.5	**90.2**	63.8	**67.2**
IER-distill [39]	84.50	**86.54**	89.7	**90.5**	65.0	**68.0**
Average improvement		**+2.05**		**+0.8**		**+3.2**

Our LabelHalluc can apply to different pretraining methods. Experiments with six different pretrainings show the consistent improvements enabled by our method.

References

(1) Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. C Finn, P Abbeel, S Levine. In ICML, 2017.
(2) Prototipical Networks for Few-shot Learning. J Snell, K Swersky, R Zemel. In NeurIPS, 2017.
(3) Rethinking Few-Shot Image Classification: a Good Embedding is All You Need? Y Tian, Y Wang, D Krishnan, J Tenenbaum, P Isola, In ECCV, 2020.
(4) A Baseline for Few-shot Image Classification. G Dhillon, P. Chaudhari, A Ravichandran, S Soatto, in ICLR, 2020.