An Embarrassingly Simple Approach to Semi-Supervised Few-Shot Learning

Xiu–Shen Wei[1,2], He–Yang Xu[1], Faen Zhang[3], Yuxin Peng[4], Wei Zhou[5]

1 School of Computer Science and Engineering, Nanjing University of Science and Technology, China

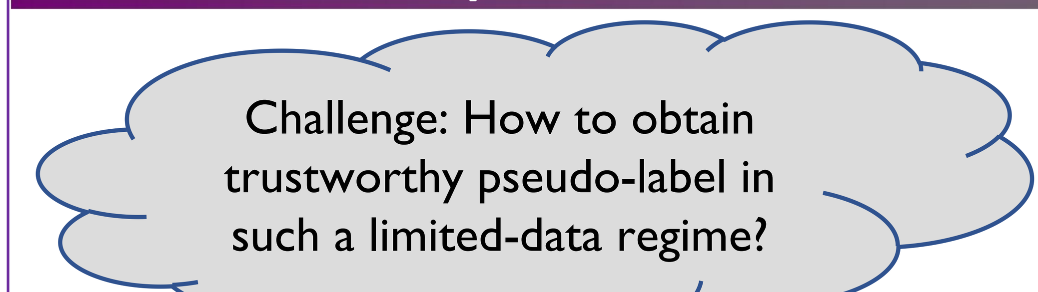
2 State Key Laboratory of Integrated Services Networks, Xidian University, China

3 Qingdao AInnovation Technology Group Co., Ltd

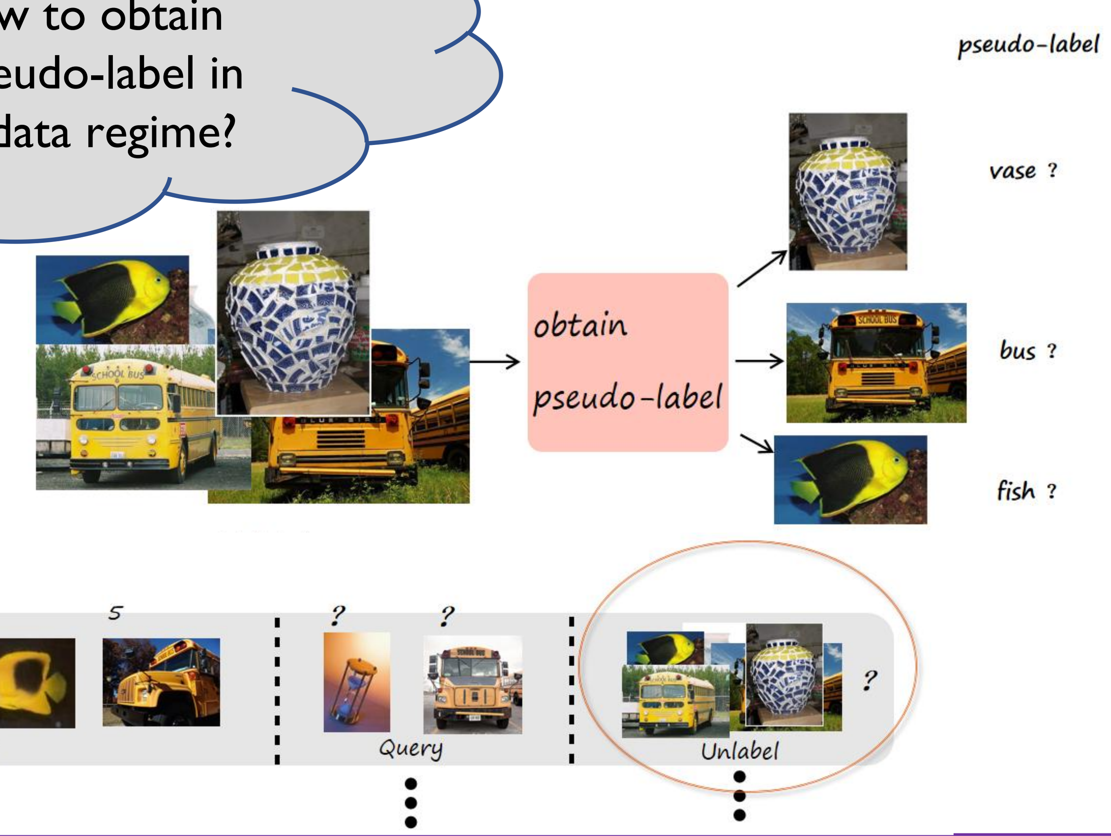
4 Wangxun Institute of Computer Technology, Peking University

5 CICC Alpha (Beijing) Private Equity

1 – Semi–Supervised Few–Shot Learning(SSFL)

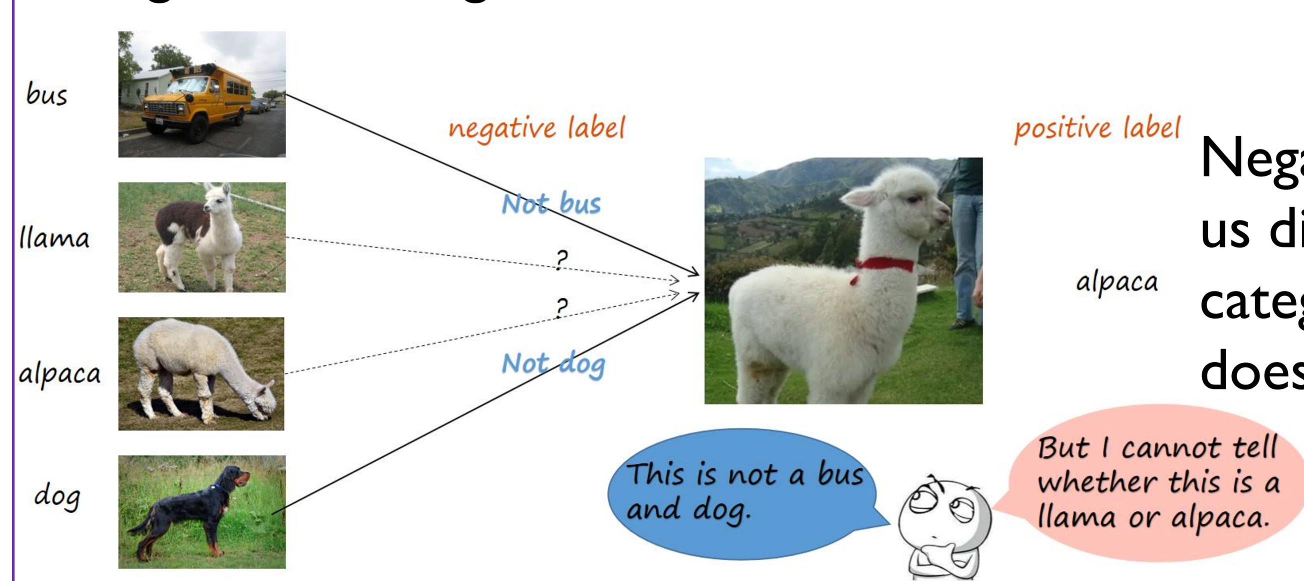


SSFSL supposes that we can utilize unlabeled data to help to learn the model.



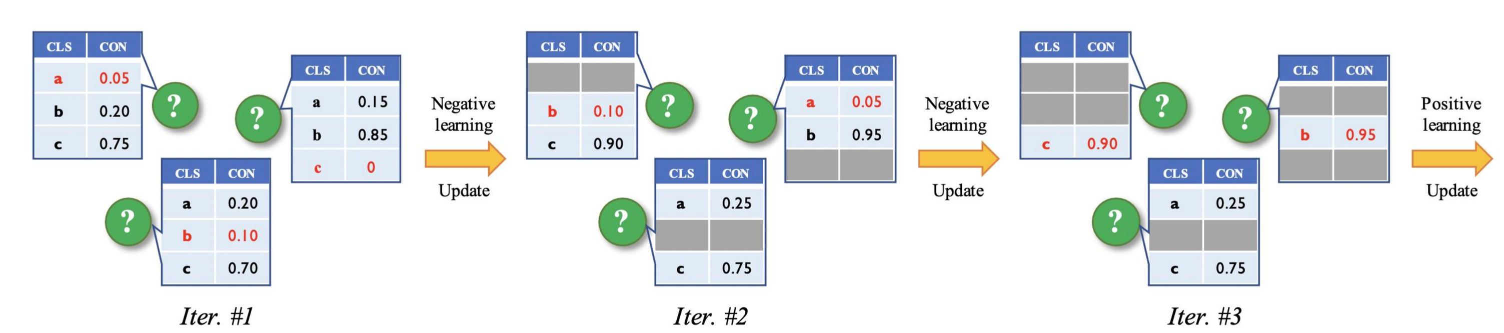
2 — The proposed MUSIC

Negative learning



Negative label helps us distinguish what category this object does not belong to.

Our Method of sUccesSIve exClusions (MUSIC)



Cross-entropy loss

$$\mathcal{L}(f,\mathbf{y}) = -\sum_k y_k \log p_k$$

Minimum-entropy loss

$$\mathcal{L}(f,\overline{\mathbf{y}}^u) = -\sum_k \overline{y}_k^u \log(1-p_k^u)$$

Negative cross-entropy loss

$$\mathcal{L}(f,\overline{\mathbf{y}}^u) = -\sum_k \overline{y}_k^u \log(1-p_k^u)$$

Algorithm 1 Pseudo-code of the proposed MUSIC

```
# f: a classifier, cf. Eqn. (2) of the paper
# δ: a reject option to select the negative label, cf. Eqn. (4) of the paper
# c: the number of classes
# Position: a list to record the label which has been selected as the negative label in each iteration
# S, U: embeddings of the support and unlabeled set which have been extracted by the pre-trained CNN
model (|S|=L, |U|=M)

begin:
    logits ← f(S) # support logits (L, c)
    loss ← CELoss(logits, targets) # CrossEntropy

    while True:
        # negative logits and negative label (M)
        neg_logits, neg_label ← get_neg_samples(Position, f, U, δ)
        if len(neg_label)==0:break # the condition to stop the iterations
        # NegCrossEntropy loss, cf. Eqn. (5); Minimum-Entropy loss, cf. Eqn. (6) of the paper
        loss ← NegCELoss(neg_logits, neg_label) + MiniEntropy(neg_logits)
    end

    pos_logits, pos_label ← get_pos_samples(Position)
    loss ← CELoss(pos_logits, pos_label) + MiniEntropy(pos_logits)
end
```

3 — Contributions

- ✓ We propose a simple but effective approach, i.e., MUSIC, to deal with semi-supervised few-shot classification tasks. To our best knowledge, MUSIC is the first approach to leverage negative learning as a straightforward way to provide pseudo-labels with as much confidence as possible in such extremely label-constrained scenarios.
- ✓ We can implement the proposed approach using only off-the-shelf deep learning operations, and it can be implemented in just few lines of code. Besides, we also provide the default value recommendations of hyper-parameters in our MUSIC, and further validate its strong practicality and generalization ability via various SSFSL tasks.
- ✓ We conduct comprehensive experiments on four few-shot benchmark datasets, i.e., miniImageNet, tieredImageNet, CIFAR-FS and CUB, for demonstrating our superiority over state-of-the-art FSL and SSFSL methods. Moreover, a series of ablation studies and discussions are performed to explore working mechanism of each component in our approach.

4 — Experiments

Comparisons of classification accuracy on four benchmarks.

METHOD	BACKBONE	*miniImageNet*		*tieredImageNet*		*CIFAR-FS*		*CUB*	
		1-shot	5-shot	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
MatchingNet [33]	4 CONV	43.56	55.31	–	–	–	–	–	–
MAML [4]	4 CONV	48.70	63.11	51.67	70.30	58.90	71.50	54.73	75.75
ProtoNet [29]	4 CONV	49.42	68.20	53.31	72.69	55.50	72.00	50.46	76.39
LEO [28]	WRN-28-10	61.76	77.59	66.33	81.44	–	–	–	–
CAN [8]	ResNet-12	63.85	79.44	69.89	84.23	–	–	–	–
DeepEMD [45]	ResNet-12	65.91	82.41	71.16	86.03	74.58	86.92	75.65	88.69
FEAT [43]	ResNet-12	66.78	82.05	70.80	84.79	–	–	73.27	85.77
RENet [11]	ResNet-12	67.60	82.58	71.61	85.28	74.51	86.60	82.85	91.32
FRN [40]	ResNet-12	66.45	82.83	72.06	86.89	–	–	83.55	92.92
COSOC [21]	ResNet-12	69.28	85.16	73.57	87.57	–	–	–	–
SetFeat [1]	ResNet-12	68.32	82.71	73.63	87.59	–	–	79.60	90.48
MCL [20]	ResNet-12	69.31	85.11	73.62	86.29	–	–	85.63	93.18
STL DeepBDC [41]	ResNet-12	67.83	85.45	73.82	89.00	–	–	84.01	**94.02**
TPN [19]	4 CONV	52.78	66.42	55.74	71.01	–	–	–	–
TransMatch [44]	WRN-28-10	60.02	79.30	72.19	82.12	–	–	–	–
LST [17]	ResNet-12	70.01	78.70	77.70	85.20	–	–	–	–
EPNet [23]	ResNet-12	70.50	80.20	75.90	82.11	–	–	–	–
ICI [36]	ResNet-12	69.66	80.11	84.01	89.00	76.51	84.32	89.58	92.48
iLPC [15]	ResNet-12	70.99	81.06	85.04	89.63	78.57	85.84	90.11	–
PLCM [9]	ResNet-12	72.06	83.71	84.78	90.11	77.62	86.13	–	–
Ours	ResNet-12	**74.96**	**85.99**	**85.40**	**90.79**	**78.96**	**87.25**	**90.76**	**93.27**
Ours (only neg)	ResNet-12	73.86	85.11	84.91	90.29	78.26	86.53	89.91	92.46
Ours (only pos)	ResNet-12	**74.44**	**85.86**	**85.33**	**90.62**	**78.81**	**87.11**	**90.27**	93.11

5 — Ablation Studies

Our Music can get more trustworthy pseudo-label.

METRIC	DATASET	ICI [36]	iLPC [15]	Ours (w/o δ)	**Ours**
ERROR RATE	*miniImageNet*	24.72% (61.8/250)	23.92% (59.8/250)	14.92% (37.3/250)	**9.09% (18.0/198)**
	CUB	29.03% (56.6/195)	26.10% (50.9/195)	20.00% (39.0/195)	**11.91% (18.1/152)**
PROPORTION	*miniImageNet*	100%	100%	100%	79.20%
	CUB	100%	100%	100%	77.95%

SETTINGS	*miniImageNet*		*CUB*	
	1-shot	5-shot	1-shot	5-shot
neg $\rightarrow$ pos $\rightarrow \cdots$	**74.77**	**85.43**	**90.47**	**92.59**
pos $\rightarrow$ neg $\rightarrow \cdots$	74.54	85.09	90.29	92.24

Learning from negative pseudo-label first can get better results.

(a) ICI (b) iLPC (c) Our MUSIC

t-SNE visualization