Propagation Regularizer for Semi-supervised Learning with Extremely Scarce Labeled Samples [CVPR 2022]

Serious problems of scarce-label situations:

1. **Confirmation bias**:

The model learns incorrect predictions for unlabeled data, so that the confidence of the incorrect prediction is increased and the model has resistance to new (correct) information that can be corrected.

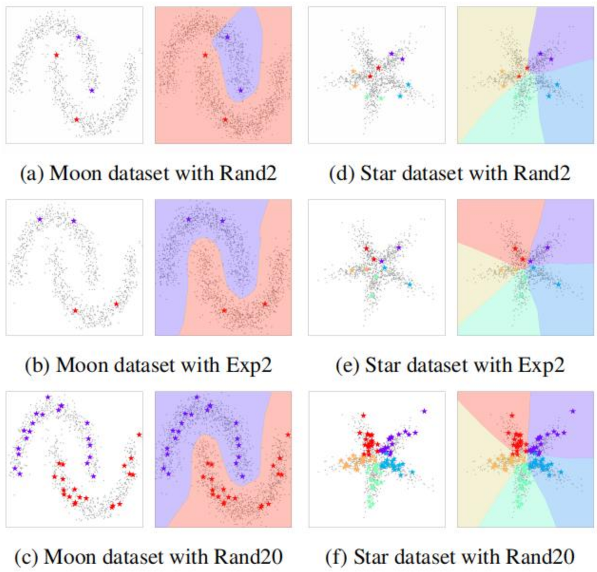


Figure 1. Class boundaries by FixMatch. Labeled samples are in colors and unlabeled samples are in grey. Each crescent is a class in moon dataset and each wing is a class in star dataset.

Method	Fold	Class										Entropy	Accuracy
		0	1	2	3	4	5	6	7	8	9		
FixMatch	Fold 1	0.11	0.09	0.02	0.00	0.45	0.00	0.09	0.01	0.11	0.11	0.72	62.29
	Fold 2	0.11	0.10	0.02	0.08	0.26	0.00	0.10	0.09	0.10	0.13	0.90	67.18
	Fold 3	0.22	0.01	0.08	0.00	0.36	0.00	0.07	0.08	0.00	0.17	0.72	53.05
	Fold 4	0.00	0.09	0.01	0.00	0.19	0.01	0.36	0.10	0.11	0.11	0.77	51.31
	Fold 5	0.10	0.09	0.01	0.00	0.01	0.17	0.13	0.30	0.10	0.09	0.84	66.23

Table 1. Class ratio and entropy of pseudo labels for CIFAR-10 dataset with 10 labeled samples.

2. **Model selection:**

In scarce-label situations, the learning of SSL can be very unstable because of confirmation bias. A model with a low training loss does not guarantee a good test accuracy.

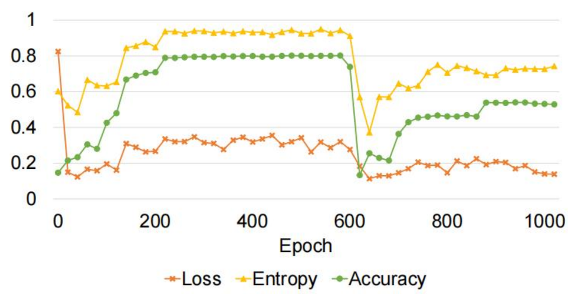


Figure 2. Training loss, entropy of pseudo-labels and test accuracy of FixMatch on CIFAR-10 with 10 labeled samples (Fold 3).

Pearson's Correlation Coefficient	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
Training Loss-Accuracy	0.175	0.364	0.747	0.257	0.406	0.390
Entropy-Accuracy	0.835	0.807	0.950	0.841	0.846	0.856

Table 2. Pearson's Correlation Coefficient of training loss-test accuracy and entropy-test accuracy during FixMatch training on the CIFAR-10 with 10 labeled samples.

Proposed method:

1. **Propagation Regularizer**

We need to keep the balance between pseudo-labels based on our observation that the correlation between test accuracy and the entropy of pseudo-classes is high as shown in Tab. 2. Learning imbalanced pseudo-labeled sample will augment confirmation bias.

To solve this problem, a regularization term is designed so that the pseudo-labeling for the unlabeled samples should be balanced for each class as follows:

$$L_{pr} = 1 - (-\mathbf{P}_{\mathrm{U}} \cdot \log_c(\mathbf{P}_{\mathrm{U}}))$$

$$\mathbf{P}_{\mathrm{U}} = \frac{1}{|\mathrm{U}|} \sum_{u \in \mathrm{U}} \mathbb{1}\left(\max\left(p\left(u\right)\right) \geq \tau\right) p\left(u\right)$$

If the pseudo-labels of unlabeled examples are evenly distributed, the value of L_{pr} will converge to 0.

2.**Model Selection based on Propagation Regularizer and Utilization**

To select an appropriate model, we propose a measure based on confirmation bias and utilization of unlabeled samples.

For the utilization measure of unlabeled samples:

$$T_{\mathrm{U}} = \frac{1}{|\mathrm{U}|} \sum_{u \in \mathrm{U}} \mathbb{1}(\max(p(u)) \geq \tau)$$

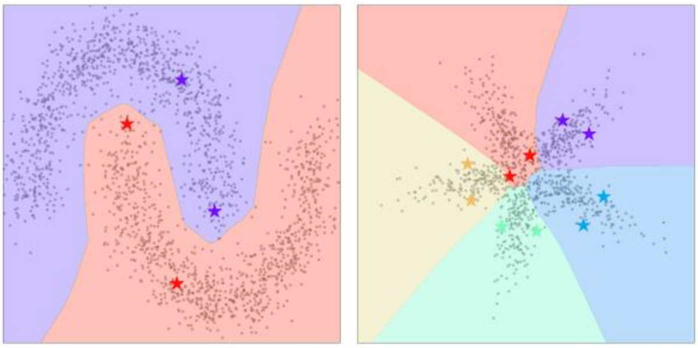
To measure the influence of confirmation bias:

$$L_{pr} = 1 - (-\mathbf{P}_{\mathrm{U}} \cdot \log_c(\mathbf{P}_{\mathrm{U}}))$$

Finally, we develop a metric for the model selection.

$$Sel = (1 - L_{pr}) + T_{\mathrm{U}}$$

A good SSL model utilizes most unlabeled samples and is less affected by confirmation bias, so the value will be max imized. In the training process, we evaluate Sel at each epoch, and choose the model with the maximum value of Sel as the final model.



(a) Moon dataset with Rand2 (b) Star dataset with Rand2

Figure 3. Datasets and class boundaries by FixMatch with the proposed method. In the dataset, labeled samples are in colors unlabeled samples are in grey. In moon datasets, each crescent is a class and in star datasets, each wing is a class.

Method	Fold	Class										Entropy	Accuracy
		0	1	2	3	4	5	6	7	8	9		
FixMatch **+ Sel + Reg**	Fold 1	0.10	0.12	0.07	0.08	0.18	0.11	0.12	0.03	0.11	0.09	0.97	68.60
	Fold 2	0.09	0.12	0.06	0.06	0.09	0.09	0.18	0.10	0.10	0.11	0.98	59.47
	Fold 3	0.07	0.17	0.04	0.09	0.13	0.08	0.12	0.03	0.09	0.17	0.95	72.89
	Fold 4	0.17	0.11	0.02	0.05	0.10	0.15	0.13	0.07	0.12	0.08	0.95	78.81
	Fold 5	0.12	0.11	0.01	0.10	0.03	0.10	0.12	0.21	0.12	0.08	0.94	74.58

Table 3. Class ratio and entropy of pseudo labels for CIFAR-10 dataset with 10 labeled samples. The proposed model selection and propagation regularizer are applied.