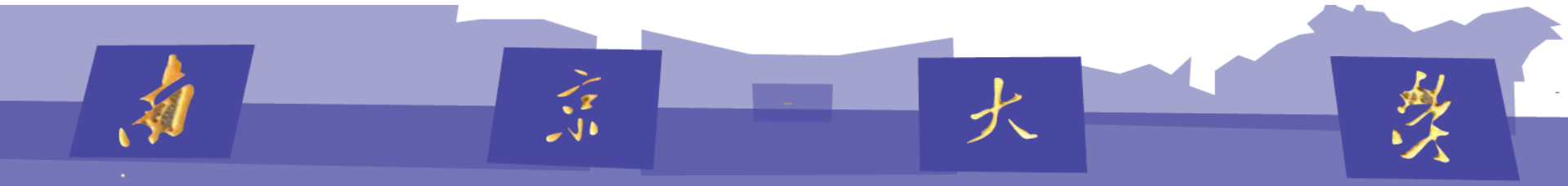
Safe Deep Semi-Supervised Learning for Unseen-Class Unlabeled Data (ICML2020)

Lan-Zhe Guo, Zhen-Yu Zhang, Yuan Jiang, Yu-Feng Li, Zhi-Hua Zhou

Nanjing University, China

http://www.lamda.nju.edu.cn



Semi-Supervised Learning

LAMDA
Learning And Mining from DatA
http://www.lamda.nju.edu.cn

Semi-supervised learning (SSL) aims to learn from both **labeled** and **unlabeled data**

Traditionally

labeled and unlabeled data come from the same distribution

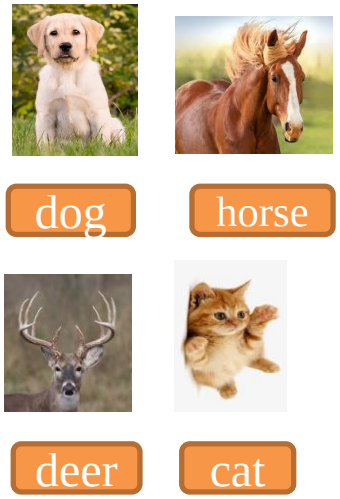
However

In real application

Class distribution mismatch often occurs

For example

Labeled Data



Unlabeled Data

Unlabeled data contains class that are not seen in labeled data

"Unseen-Class Unlabeled data"

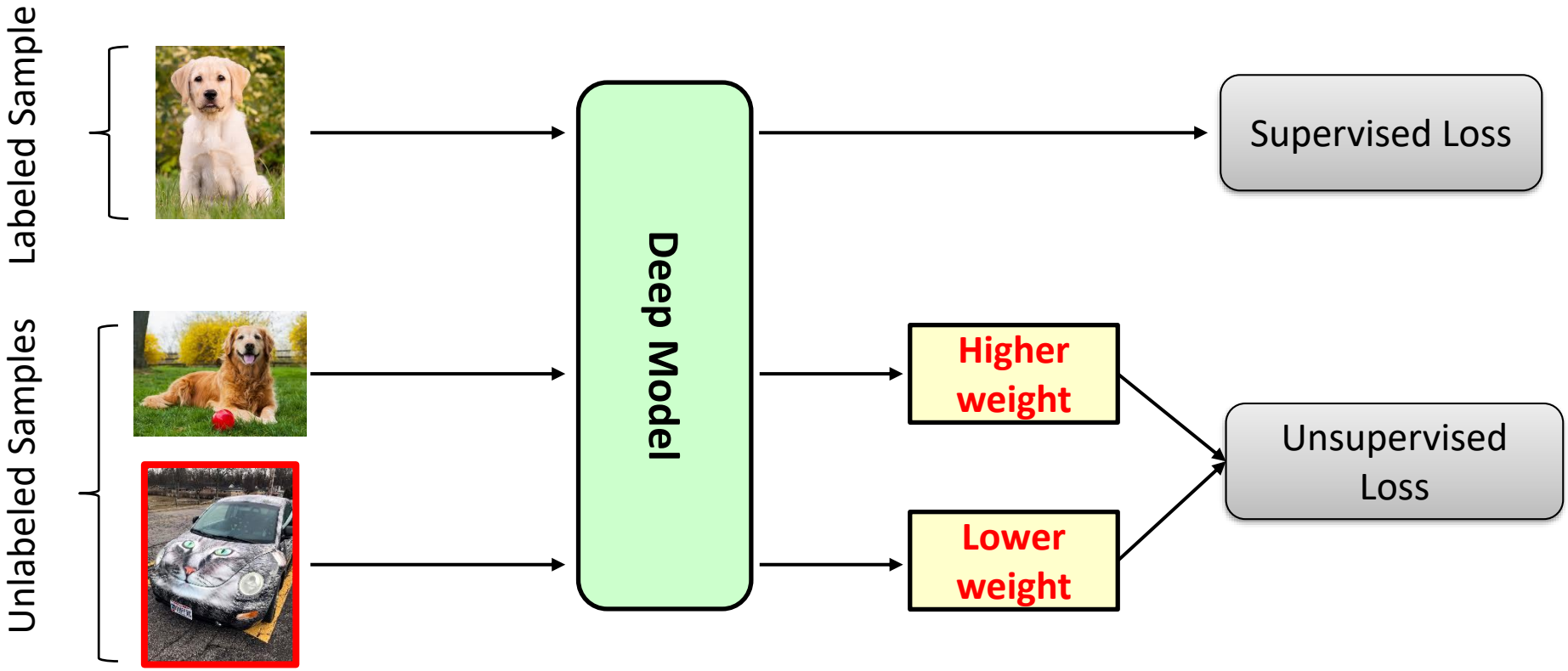
DS3L (Deep Safe SSL)

LAMDA
Learning And Mining from DatA
http://www.lamda.nju.edu.cn

DSSL : Treat all unlabeled instances equally —not suitable→ Faced with unseen classes

⇩

DS3L : Use unlabeled data selectively → Weighting unlabeled instances



DS3L (Deep Safe SSL)

Weighted Loss Function

$$\hat{\theta}(\alpha) = \min_{\theta \in \Theta} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \theta), \mathbf{y}_i) + \sum_{i=n+1}^{n+m} \underbrace{w(\mathbf{x}_i; \alpha)} \Omega(\mathbf{x}_i; \theta)$$

Weight function

Idea: The learned model should maximize the generalization performance

$$\alpha^* = \operatorname*{argmin}_{\alpha \in \mathbb{B}^d} \mathbb{E}_{(X,Y)}[\ell(h(X; \hat{\theta}(\alpha)), Y)]$$

Bi-Level Objective

Outer Level

$$\min_{\alpha \in \mathbb{B}^d} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \hat{\theta}), \mathbf{y}_i)$$

Inner Level

$$\hat{\theta}(\alpha) = \min_{\theta \in \Theta} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \theta), \mathbf{y}_i) + \sum_{i=n+1}^{n+m} w(\mathbf{x}_i; \alpha) \Omega(\mathbf{x}_i; \theta)$$

Theoretical Results

- **Safeness: DS³L can achieve safeness in terms of empirical performance**

The empirical risk of $\hat{\theta}$ learned by DS3L is never worse than the supervised model,

$$\hat{R}(\hat{\theta}) \leq \hat{R}(\theta^{SL})$$

- **Generalization: DS³L approaches the optimal weight in the order $O\left(\sqrt{\frac{d \ln n}{n}}\right)$**

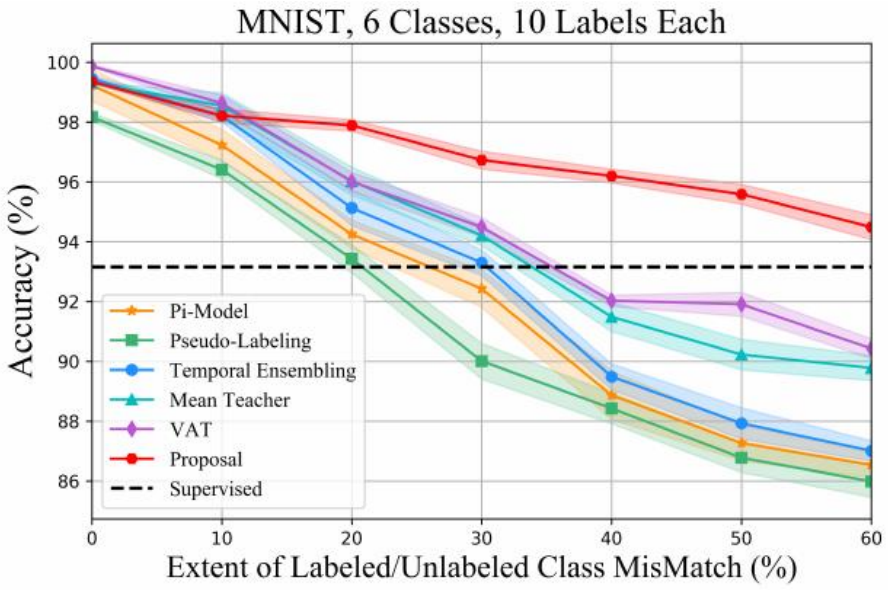
For any $\delta > 0$

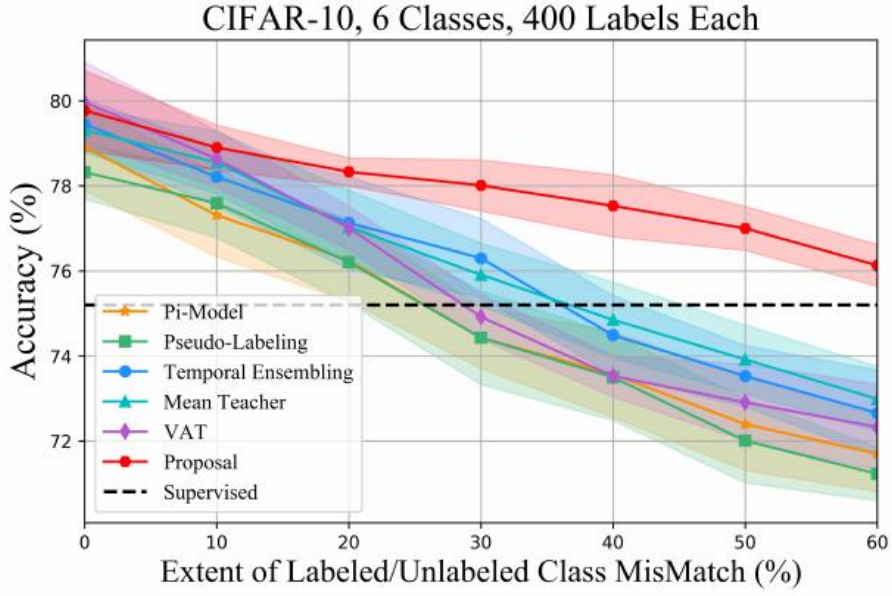
$$R(\hat{\theta}(\alpha^*)) \leq R(\hat{\theta}(\hat{\alpha})) + \frac{(3\lambda + \sqrt{4d\ln(n) + 8\ln(2/\delta)})}{\sqrt{n}}$$

Holds with probability at least $1 - \delta$

Experimental Results

Experimental results on MNIST and CIFAR-10 dataset with varying extent of labeled and unlabeled class mismatch





Existing DSSL methods are no longer as good as supervised learning in **40%** of unseen-class unlabeled data, the new method can still achieve performance gain in more than **60%** of unseen-class unlabeled data.

http://www.lamda.nju.edu.cn

Thank you!

If you are interested in, feel free to contact me:

Lan-Zhe Guo (guolz@lamda.nju.edu.cn)