Iterative Learning with Open-set Noisy Labels

Yisen Wang, Weiyang Liu, Xingjun Ma,

James Bailey, Le Song, Shu-Tao Xia

Tsinghua University, Georgia Tech, University of Melbourne

Types of Noisy Labels

- **Closed-set noisy labels**
  - *A noisy sample possesses a true class that is contained within the set of known classes in the training data*
- **Open-set noisy labels:**
  - *A noisy sample possesses a true class that is not contained within the set of known classes in the training data*

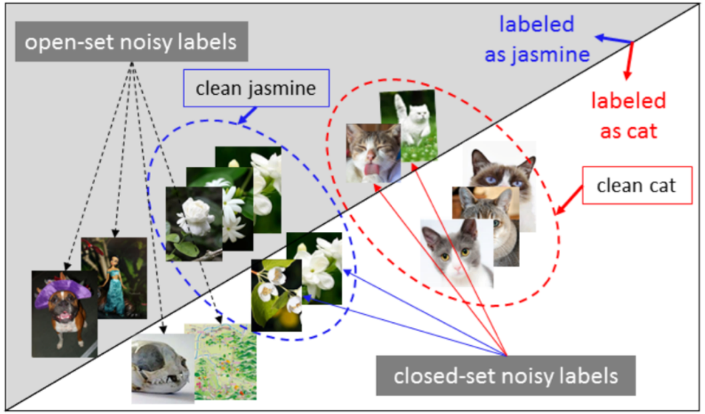


Figure 1. An illustration of closed-set vs open-set noisy labels.

Table 1. Types of labels for a "jasmine-cat" dataset.

	labeled as "jasmine"	labeled as "cat"
true "jasmine"	clean	closed-set
true "cat"	closed-set	clean
other class images	open-set	open-set

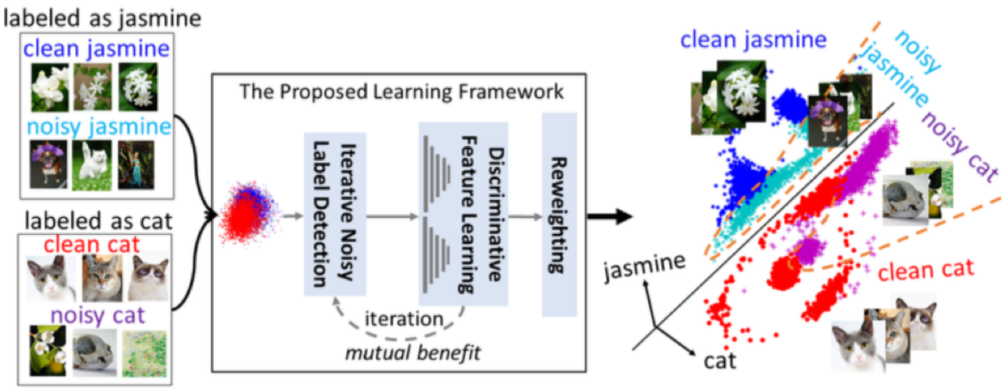
Different methods to handle different noise

- **Closed-set noisy labels**
  - ***loss correction***
  - ***noise model based clean label inferring***

- **Open-set noisy labels**
  - ***loss or label correction may be inaccurate since the true class may not exist in the dataset***
  - ***Iterative learning framework to pull away noisy samples from clean samples in the deep representation space***

Iterative learning framework

- **Iterative noisy label detection:**
  - ***iteratively detect noisy labels based on the features of the network***
- **Discriminative feature learning:**
  - ***imposes a representation constraint via contrastive loss to pull away noisy samples from clean samples in the deep representation space***
- **Reweighting:**
  - ***express a relative confidence of clean and noisy labels on the representation learning***



Experiments

- **True noisy label rate:**

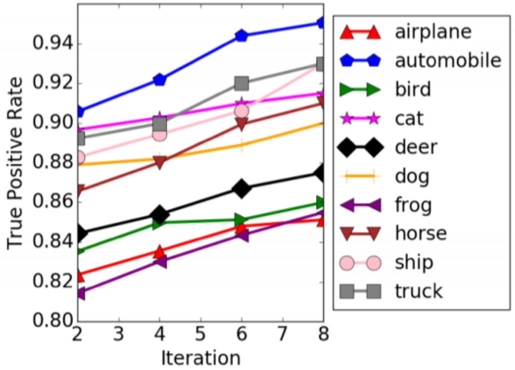


Figure 5. The true positive rate of the detected noisy labels over iteration on CIFAR-10+CIFAR-100 (40% open-set noise).

- **Visualization of the learned features:**

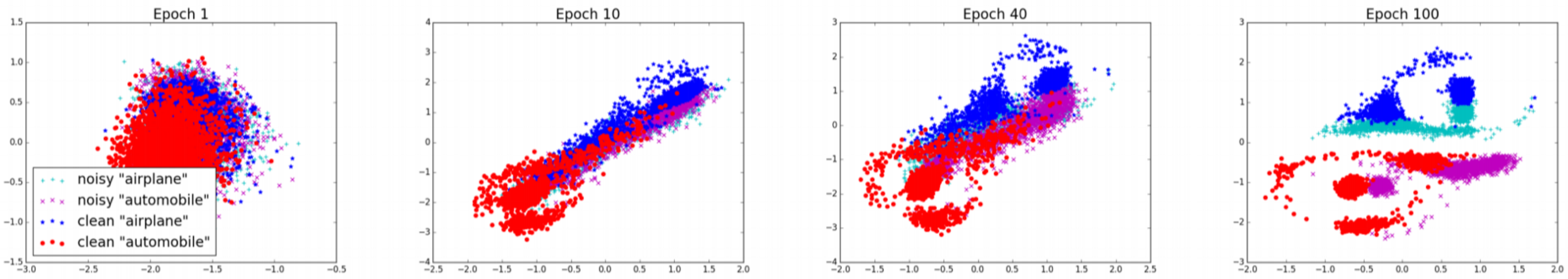


Figure 7. Visualization of the learned features. This visualization experiment uses a 2-class subset of CIFAR-10+CIFAR-100 (40% open-set noise) by setting the output feature dimension as 2.

Experiments

- **Ablation study:**

Table 4. Accuracies (%) on CIFAR-10+CIFAR-100 (20% & 40% open-set noise) after removing (w/o) each module from our model.

Method	CIFAR-10+CIFAR-100 20% noise	CIFAR-10+CIFAR-100 40% noise
Our model	81.96	79.28
(a) w/o reweighting		
-- case a_1: $\gamma = 1$	76.97	74.45
-- case a_2: $\gamma = 0$	79.27	76.03
(b) w/o discriminative learning		
-- case b_1: removing	76.22	68.40
-- case b_2: new class	78.34	73.11
(c) w/o iterative detection		
-- case c_1: only once	77.52	70.31
-- case c_2: no	76.17	63.50

Experiments

- **On ImageNet (200 classes):**

Table 6. Accuracies (%) of different models on the 200-class ImageNet with 20% open-set noise. The best results are in **bold**.

Method	ResNet-50		Inception-v3	
	Top-1	Top-5	Top-1	Top-5
Cross-entropy	58.51	75.62	60.73	76.75
Backward	59.32	75.61	61.27	76.74
Forward	64.17	79.43	65.48	80.68
Bootstrapping	59.05	75.00	61.50	76.13
CNN-CRF	66.54	82.37	67.23	84.12
Ours	**70.29**	**86.04**	**71.43**	**87.87**

Take Home Message

- **The open-set noisy label problem – a more complex noisy label scenario that commonly occurs in real-world datasets**
- **An iterative learning framework to address the problem with three powerful modules: iterative noisy label detection, discriminative feature learning, and reweighting**
- **Modules are benefited from each other and jointly improved over iterations**

Thank you!