Paper ID # 2486

Co-learning: Learning from Noisy Labels with Self-supervision

Cheng Tan, Jun Xia, Lirong Wu, Stan Z. Li
AI Lab, School of Engineering, Westlake University.
Institute of Advanced Technology, Westlake Institute for Advanced Study

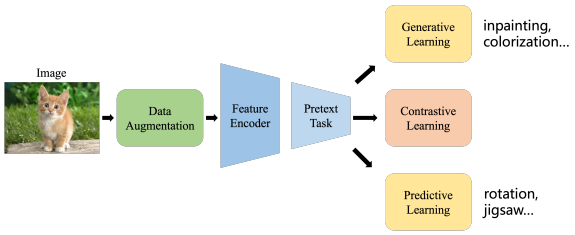
Section 1 Learning with noisy label:

- Manual labeling may produce some wrong labels.

- Data collected by web crawlers often contains incorrect labels.

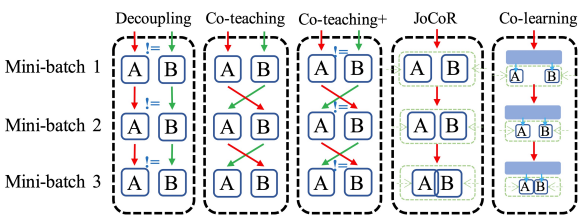
Section 2 Self-supervised learning:

- Self-supervised learning is a subset of unsupervised learning that designs pretext tasks to produce labels derived from the data.
- While solving pretext tasks, the model learns valuable representations. There are various pretext tasks that can help the model learn representations.
- Current state-of-the-art contrastive learning methods mainly rely on a combination of **intrinsic similarity** and a set of image transformations.



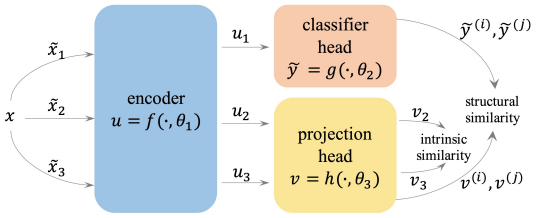
Section 3 Co-training based schemes:

- Decoupling updates the parameters with disagreed (!=) samples.
- Co-teaching uses small-loss samples to teach peer network.
- Co-teaching+ first predicts the two networks but uses disagreed samples only to compute the training loss.
- JoCoR trains the two networks as a whole with a joint loss.
- Co-learning trains a shared encoder network with two heads.

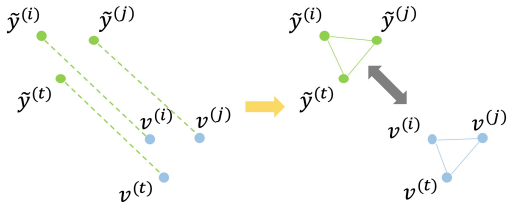


Section 4 Co-learning:

- The main learning task is to learn the encoder and the classifier, which is done with a vanilla cross-entropy loss defined on label-dependent information.



- A contrastive prediction-based pretext task is used to assist the main task with an **intrinsic similarity** loss defined on feature-dependent information.
- A **structural similarity** loss is introduced in this work to maximize agreement on the feature encoder shared by the two tasks.



Section 5 Experimental results:

- Average test accuracy (%) on CIFAR-10 over the last 10 epochs.

Flipping-Rate	Standard	Decoupling	Co-teaching	Co-teaching+	JoCoR	APL	Co-learning
Symmetric-20%	84.81±0.24	85.75±0.31	90.29±0.19	88.63±0.32	90.43±0.25	88.54±0.45	**92.21**±0.31
Symmetric-50%	61.49±0.58	61.93±0.82	63.45±3.89	76.27±2.80	66.00±0.53	76.51±1.73	**84.49**±0.34
Symmetric-80%	28.98±0.26	27.23±0.84	28.03±1.67	30.37±1.69	29.19±1.64	24.75±2.87	**61.20**±2.29
Asymmetric-20%	87.00±0.20	87.66±0.29	89.38±0.33	89.00±0.18	89.20±0.26	88.02±0.29	**91.07**±0.32
Asymmetric-30%	81.99±0.31	81.83±0.26	86.58±1.32	86.22±0.26	86.41±0.45	86.03±0.21	**86.89**±0.87
Asymmetric-40%	76.30±0.34	74.97±0.38	74.25±0.38	81.25±0.75	73.95±1.00	80.97±0.19	**81.42**±0.52

- Average test accuracy (%) on CIFAR-100 over the last 10 epochs.

Flipping-Rate	Standard	Decoupling	Co-teaching	Co-teaching+	JoCoR	APL	Co-learning
Symmetric-20%	57.79±0.44	56.18±0.32	64.28±0.32	55.40±0.71	62.29±0.71	59.21±0.50	**66.58**±0.15
Symmetric-50%	33.75±0.46	31.58±0.54	32.62±0.51	26.49±0.45	30.19±0.60	43.53±1.84	**54.54**±0.43
Symmetric-80%	8.64±0.22	7.71±0.23	6.65±0.71	8.57±1.55	6.84±0.92	13.97±0.53	**35.45**±0.79
Asymmetric-20%	59.36±0.36	57.97±0.24	59.76±0.53	56.11±0.60	58.58±0.51	58.89±0.40	**65.26**±0.76
Asymmetric-30%	51.06±0.44	49.86±0.54	49.53±0.79	47.12±0.73	49.04±0.91	51.46±0.15	**56.97**±1.22
Asymmetric-40%	42.49±0.23	41.51±0.67	40.62±0.79	38.98±0.54	39.72±0.76	41.96±0.92	**47.62**±0.79

- Accuracy (%) on Animal-10N dataset and Food-101N dataset.

Methods	best	last
Standard CE	82.68	81.10
Decoupling	79.22	78.24
Co-teaching	82.43	81.52
Co-teaching+	50.66	48.52
JoCoR	82.82	81.06
Co-learning	**82.95**	**82.18**

Methods	best	last
Standard CE	84.50	83.86
Decoupling	85.53	85.28
Co-teaching	91.91	91.86
Co-teaching+	81.61	81.24
JoCoR	77.94	77.86
Co-learning	**87.57**	**86.56**

Summary

- Point out the problems of the common co-training paradigm in noisy learning.
- Propose a new noisy learning method known as: Co-learning, which assisted supervised learning through self-supervised learning.
- Reproduce similar methods under a unified framework for fair comparison and obtain the best performance on multiple benchmark data sets. (See github.com/chengtan9907/Co-training-based_noisy-label-learning)