

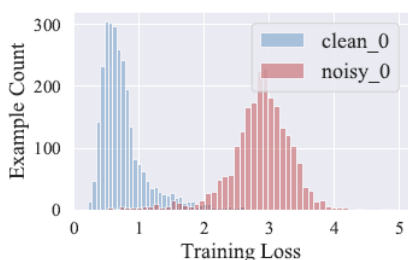
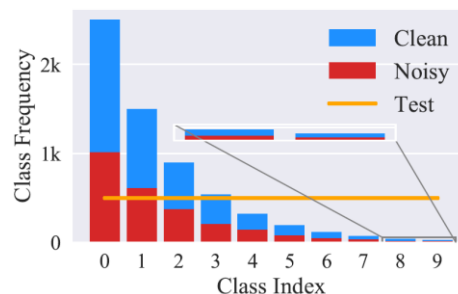
Robust Long-Tailed Learning Under Label Noise

Tong WEI, Jiang-Xin SHI, Min-Ling ZHANG, Yu-Feng LI

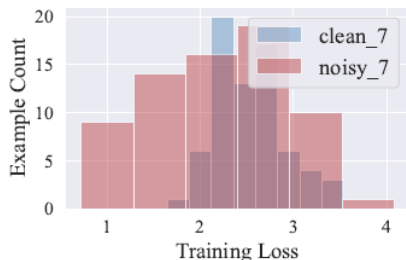
Frontiers of Computer Science, DOI: [10.1007/s11704-025-40860-0](https://doi.org/10.1007/s11704-025-40860-0)

Problems & Ideas

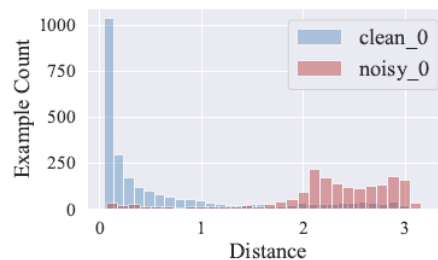
- Read-world data typically exhibits long-tail class distribution and contains label noise:
 - Previous long-tail learning methods overlooked the prevalence of noisy labels in training.
 - The commonly used small-loss noisy label detection criterion fails in long-tail data.
- Ideas: We find that the *class-wise small-distance* criterion is more robust than the small-loss criterion, and works well on both head and tail classes.



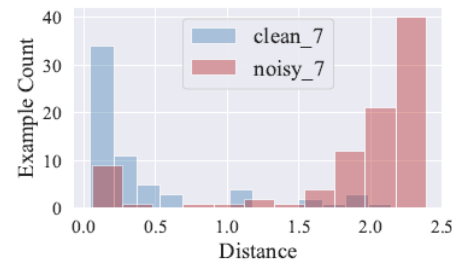
(a) loss of head class



(b) loss of tail class



(c) distance of head class



(d) distance of tail class

(a-b) Training losses head and tail class samples. (c-d) Distance distribution between samples and their class prototype for head and tail class. Experiments are conducted on CIFAR-10 with noise level $\gamma = 50\%$ and imbalance ratio $\rho = 100$.

Main Contributions

- Contributions:
 - We address the underexplored problem of learning from long-tail data in the presence of label noise, taking a significant step toward real-world applications.
 - We identify the limitations of the widely used small-loss criterion and introduce a novel criterion called *class-wise small-distance*. Building upon this, we propose a robust framework called RoLT.
 - Extensive experiments conducted on benchmark and real-world datasets demonstrate the superiority of our proposed method.

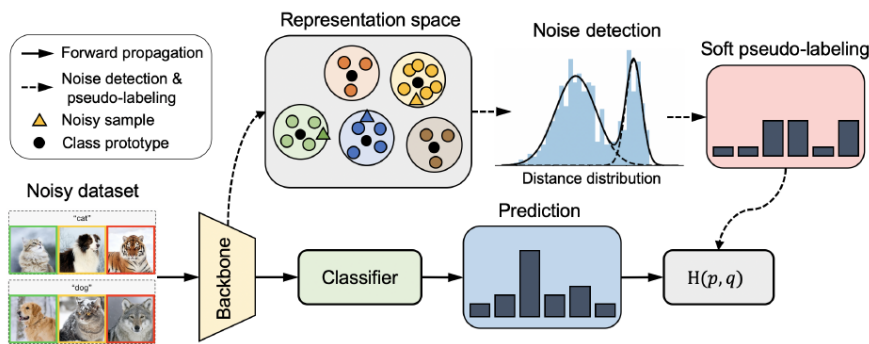


Fig. 3 The proposed framework RoLT.

		CIFAR-10						CIFAR-100					
Noise level		20%		50%				20%		50%			
Imbalance ratio		10	50	100	10	50	100	10	50	100	10	50	100
CE	Best	75.61	63.60	62.17	63.25	43.38	38.71	43.27	30.27	26.21	26.92	16.97	14.23
	Last	73.86	58.60	54.15	48.39	32.73	27.05	43.06	30.04	26.08	26.60	15.59	13.47
LDAM	Best	82.37	71.34	66.26	60.30	42.95	36.66	48.14	33.43	29.70	29.62	17.51	14.19
	Last	82.07	71.11	65.88	59.10	38.33	33.38	47.89	33.30	29.50	29.38	17.29	13.24
LDAM-DRW	Best	83.73	76.41	72.28	67.93	48.88	43.23	50.44	36.60	32.27	32.24	19.48	15.21
	Last	83.67	75.67	71.08	67.68	47.38	41.45	50.29	36.16	32.05	31.72	19.23	14.75
BBN	Best	77.81	68.01	64.51	64.71	46.22	36.72	47.60	31.07	28.79	30.01	19.75	14.56
	Last	76.81	67.48	64.24	53.76	43.35	34.83	47.26	30.76	28.56	29.42	19.55	14.34
cRT	Best	76.15	65.02	59.92	64.15	43.26	36.73	42.56	30.23	26.31	25.55	17.47	16.01
	Last	75.05	64.22	58.47	62.75	41.87	34.55	41.56	30.08	26.18	23.94	17.34	15.94
MW-Net	Best	82.19	71.63	67.26	72.12	56.09	46.36	50.20	36.68	31.77	37.50	23.99	21.24
	Last	77.67	64.12	58.23	59.68	45.39	37.05	47.82	34.45	29.57	33.14	20.33	18.82
HAR-DRW	Best	82.43	67.44	67.88	67.39	52.35	42.80	46.24	28.86	26.29	31.30	16.75	14.78
	Last	78.44	61.08	62.73	64.75	45.06	40.07	43.04	26.11	24.71	26.96	13.87	12.42
RoLT	Best	85.04	77.86	73.84	77.11	60.15	55.32	53.41	38.94	33.36	39.22	25.51	20.61
	Last	84.95	77.65	73.54	76.94	59.59	54.55	53.22	38.77	33.20	39.01	25.35	20.45

RoLT outperforms previous methods under different settings. The improvement becomes more significant at high noise levels.