# SSR: An Efficient and Robust Framework for Learning with Unknown Label Noise



# Chen Feng, Georgios Tzimiropoulos, Ioannis Patras

Queen Mary University of London, School of Electronic Engineering and Computer Science, London, UK

chen.feng@qmul.ac.uk, g.tzimiropoulos@qmul.ac.uk, i.patras@qmul.ac.uk



# Introduction

Despite the large progress in supervised learning with neural networks, there are significant challenges in obtaining high-quality, large-scale and accurately labelled datasets. In such a context, how to learn in the presence of noisy labels has received more and more attention. In this paper, we propose a simple, efficient and robust framework named Sample Selection and Relabelling (SSR) to learni when both the degree and the type of noise are unknown. At the heart of our method is a sample selection and relabelling mechanism based on a non-parametric KNN classifier (NPK)  $g_q$  and a parametric model classifier (PMC)  $g_p$ , respectively, to select the clean samples and gradually relabel the noisy samples. Without bells and whistles, such as model co-training, self-supervised pre-training and semi-supervised learning, and with robustness concerning the settings of its few hyper-parameters, our method significantly surpasses previous methods on both CIFAR10/CIFAR100 with synthetic noise and real-world noisy datasets such as WebVision, Clothing1M and ANIMAL-10N. Code is available at https://github.com/MrChenFeng/SSR\_BMVC2022.

## Insights

## Methods

- Instead of relying on a warm-up stage, our method can be **trained from scratch**.
- Instead of relying on semi-supervised learning methods to utilize the whole dataset, we simply identify and relabel closet-set noisy samples.
- Instead of complicated in-training tricks, we simply apply supervised cross-entropy loss only to achieve great performance.

# Ablations

Sample selection & Relabelling quality





- 1 while i < T do
- Generate relabeled dataset  $(\mathcal{X}, \mathcal{Y}^r)$  with  $\theta_r$ ; 2
- Generate selected clean subset  $(\mathcal{X}_c, \mathcal{Y}_c^r)$  with  $\theta_s$ ; 3
- Model training.

/\* Relabelling \*/

/\* Selection \*/



- The number of the relabeled samples is highly related to the value of  $\theta_r$  across different noise settings.
- The sample relabelling avoid open-set noise while identify and relabel closed-set noise effectively (Over 95% relabelling accuracy).
- High sample selection F-score across different noise settings (**Over 0.95**).

**Robustness w.r.t hyperparameters** 





 $\lambda = 1$  by default. For brevity, we name our method as **SSR** when  $\lambda = 0$ , and **SSR**+ when  $\lambda \neq 0$ .

Dataset	CIFAR10					CIFAR100			
Noise type	Symmetric			Assymetric	Symmetric				
Noise ratio	20%	50%	80%	90%	40%	20%	50%	80%	90%
Cross-Entropy	86.8	79.4	62.9	42.7	85.0	62.0	46.7	19.9	10.1
Co-teaching+ [36]	89.5	85.7	67.4	47.9	-	65.6	51.8	27.9	13.7
F-correction [21]	86.8	79.8	63.3	42.9	87.2	61.5	46.6	19.9	10.2
PENCIL [34]	92.4	89.1	77.5	58.9	88.5	69.4	57.5	31.1	15.3
LossModelling [2]	94.0	92.0	86.8	69.1	87.4	73.9	66.1	48.2	24.3
DivideMix* [14]	96.1	94.6	93.2	76.0	93.4	77.3	74.6	60.2	31.5
ELR+* [17]	95.8	94.8	93.3	78.7	93.0	77.6	73.6	60.8	33.4
RRL [15]	95.8	94.3	92.4	75.0	91.9	79.1	74.8	57.7	29.3
NGC [30]	95.9	94.5	91.6	80.5	90.6	79.3	75.9	62.7	29.8
AugDesc* [19]	96.3	95.4	93.8	91.9	94.6	79.5	77.2	66.4	41.2
C2D* [39]	96.4	95.3	94.4	93.6	93.5	78.7	76.4	67.8	58.7
SSR(ours)	96.3	95.7	95.2	94.6	95.1	79.0	75.9	69.5	61.8
SSR+(ours)	96.7	96.1	95.6	95.2	95.5	79.7	77.2	71.9	66.6

Experiments

- Unlike previous methods that try to integrate many different mechanisms and regularizations, we strive for a concise, simple and robust method.
- The proposed method does not utilize complicated

This work was supported by the EU H2020 AI4Media No. 951911 project.

Table 2: Evaluation on CIFAR-10 and CIFAR-100 with closed-set noise. Methods marked with an asterisk employ semi-supervised learning, model co-training or model pre-training.

CE	F-correction [21]	ELR [17]	RRL [15]	C2D* [39]	DivideMix* [14]	ELR+* [17]	AugDesc* [19]	SSR+(ours)
69.21	69.84	72.87	74.30	74.84	74.76	74.81	75.11	74.83

Table 3: Testing accuracy (%) on Clothing1M (methods with \* utilized model cotraining).

Method	Noise ratio	0.3		0.6	
method	Open ratio	0.5	1	0.5	1
	Best	87.4	90.4	80.5	83.4
LON [27]	Last	80.0	87.4	55.2	78.0
<b>D</b> <sub>2</sub> <b>C</b> [12]	Best	89.8	91.4	84.1	88.2
KOG [15]	Last	85.9	89.8	66.3	82.1
DivideNin [14]	Best	91.5	89.3	91.8	89.0
Dividemix [14]	Last	90.9	88.7	91.5	88.7
	Best	94.5	92.9	93.4	90.6
EDM [23]	Last	94.0	91.9	92.8	89.4
SSD (ours)	Best	96.0	95.7	93.8	93.1
SSK(ours)	Last	95.9	95.6	93.7	93.1
SSD ( (auro)	Best	96.3	96.1	95.2	94.0
55K+(ours)	Last	96.2	96.0	95.2	93.9

Methods	Weby	Vision	ILSVRC2012		
Wethous	Top1	Top5	Top1	Top5	
Co-teaching [10]	63.58	85.20	61.48	84.70	
DivideMix [14]	77.32	91.64	75.20	90.84	
ELR+ [17]	77.78	91.68	70.29	89.76	
NGC [30]	79.16	91.84	74.44	91.04	
LongReMix [7]	78.92	92.32	-	-	
RRL [15]	76.3	91.5	73.3	91.2	
SSR+(ours)	80.92	92.80	75.76	91.76	

Table 5: Testing accuracy (%) on Webvision.

Cross-Entropy	SELFIE [25]	PLC [37]	NCT [ <mark>6</mark> ]	SSR+(ours)
79.4	81.8	83.4	84.1	88.5

Table 6: Testing accuracy on ANIMAL-10N.

mechanisms such as semisupervised learning, model co-training and model pretraining, and is shown with extensive experiments and ablation studies to be robust to the values of its few hyper-parameters, and to consistently and by large surpass the state-of-the-art in various datasets.

• Please refer to the for more results and ablations.