Softmax-Weighted Pseudo-Label Refinement for Enhancing Robustness to Label Noise

Gouranga Bala[*1,2] GOURANGA.BALA23@GMAIL.COM
[1] *Department of Electrical Engineering, IIT Bombay*

Anuj Gupta[*1] GUPTAANUJ166@GMAIL.COM
[1] *Department of Electrical Engineering, IIT Bombay*

Amit Sethi[2] ASETHI@IITB.AC.IN
[2] *Department of Electrical Engineering, IIT Bombay*

Editors: Under Review for MIDL 2025

Abstract

Deep neural networks often require large-scale, accurately labelled datasets to perform well, but in practice, the labels are frequently corrupted by noise in medical imaging - especially instance-dependent noise. In this work, we propose a novel framework to address instance-dependent label noise by integrating three key components: (i) self-supervised pretraining using SimCLR to learn robust, noise-agnostic feature representations; (ii) an iterative pseudo-label refinement strategy employing a stage-wise consensus mechanism to progressively correct mislabeled samples; and (iii) a softmax-weighted cross-entropy loss that dynamically downweights uncertain predictions. We validate our approach on benchmark datasets such as CIFAR-10 and CIFAR-100 corrupted with synthetic noise at 20% and 30% levels, demonstrating significant improvements over state-of-the-art methods. We further validated our method on Chest X-rays and Chaoyang medical imaging datasets.

Keywords: Instance-dependent label noise, Self-supervised learning, Pseudo-label refinement, Dynamic loss weighting

1. Introduction

In real-world scenarios, label noise that depends on input features (instance-dependent) can significantly hinder generalization and reliability of deep learning models for medical images. Prior works in handling noisy labels, such as loss correction techniques (Patrini et al., 2017), small-loss selection (Jiang et al., 2018; Han et al., 2018), and methods specifically designed for instance-dependent noise (Wei et al., 2020) assume symmetric label noise. However, these techniques often fall short when faced with complex noise patterns where mislabeling is image-dependent. Our hybrid framework leverages self-supervised pretraining to establish a noise-agnostic feature foundation and employs iterative pseudo-label refinement with a confidence-weighted loss to correct mislabeled samples. Recent advances in self-supervised learning (e.g., MoCo (He et al., 2020)) and pseudo-label refinement (Li et al., 2020) have motivated our hybrid approach. Experimental results on CIFAR benchmarks and medical imaging datasets demonstrate the efficacy of our method in challenging noisy environments.

[*] Contributed equally

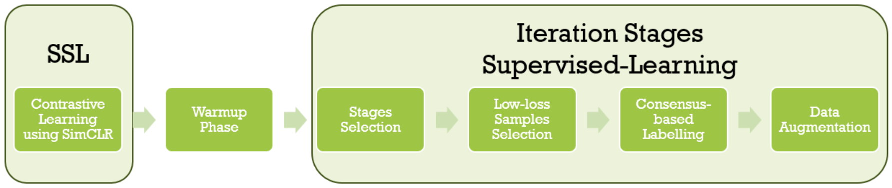

Figure 1: Proposed training method with self-supervised learning, warm-up phase, and iterative refinement

2. Methodology

The proposed pipeline has three key components as shown in Figure 1.

SimCLR-based Self-Supervised Pretraining: We leverage SimCLR (Chen et al., 2020) to learn robust, noise-agnostic features. Strong augmentations (RandomResizedCrop, ColorJitter, GaussianBlur, RandomGrayscale) were applied during pretraining, which is carried out for 30 epochs. The resulting embeddings serve as initial representations for subsequent supervised training.

Warmup Phase: A 5-epoch warmup using standard cross-entropy loss trains the model on the noisy data. This phase captures fundamental data patterns while mitigating overfitting to incorrect labels.

Iterative Pseudo-Label Refinement and Confidence-Weighted Loss: Our core contribution is the following iterative refinement mechanism:

- **Stage-Wise Training:** In the first iteration, epochs 2, 3, and 4 serve as monitoring stages; later iterations use epochs 2, 5, and 7 to capture both short- and long-term sample behavior.
- **Loss Threshold Filtering:** Samples with a loss below a preset threshold are considered reliable.
- **Consensus-Based Pseudo-Labeling:** Only the samples consistently deemed reliable across stages are updated with pseudo-labels, while the remaining samples retain their original labels.

Pseudo-labeled samples are re-augmented using the same strong transformations as in pretraining. Finally, the model is fine-tuned using a softmax-weighted cross-entropy loss that dynamically downweights uncertain predictions.

3. Experiments, Results, and Conclusion

We evaluated our approach on CIFAR-10 and CIFAR-100, and on Chest X-rays(Kermany, 2018) and Chaoyang(Zhu et al.) medical imaging datasets with 20% and 30% label noise.

Table 1 summarizes the classification accuracy (%) of various methods on CIFAR datasets.

Table 1: Comparison of methods on CIFAR-10 and CIFAR-100 for 20% and 30% noise.

Method	**CIFAR-10**		**CIFAR-100**	
	20% Noise	30% Noise	20% Noise	30% Noise
Co-teaching (Han et al., 2018)	80.96	78.56	63.58	33.43
MentorNet (Jiang et al., 2018)	81.03	77.22	63.51	34.23
JoCoR (Wei et al., 2020)	83.95	0.0	64.42	0.0
DivideMix (Li et al., 2020)	91.94	93.48	70.67	75.89
Our Method	**94.40**	**93.60**	**76.61**	**73.11**

Table 2: Performance on medical imaging datasets (Accuracy %).

Method	Chest X-rays		Chaoyang	
	20% Noise	30% Noise	20% Noise	30% Noise
DivideMix (Li et al., 2020)	73.24	73.88	34.46	20.91
Our Method	74.00	69.31	42.40	43.85

Table 2 shows the performance on medical imaging datasets for Chest X-rays(Kermany, 2018) and Chaoyang(Zhu et al.) under 20% and 30% noise levels.

Our experiments demonstrate that integrating self-supervised pretraining, iterative pseudo-label refinement, and a confidence-weighted loss creates a robust framework to mitigate instance-dependent label noise. The proposed method consistently outperforms state-of-the-art techniques on CIFAR-10 and CIFAR-100 at 20% and 30% noise levels. Moreover, the improvements observed on medical imaging datasets confirm the potential for clinical applications.

While our results are promising, the proposed framework relies on a fixed loss threshold for pseudo-label selection, while future work could explore adaptive thresholding strategies that adjust to varying noise conditions. Additionally, detailed ablation studies to isolate the contributions of each component would provide further insights into the mechanism's effectiveness. The framework should be tested on real-world noise scenarios where test data labels are curated while the training data is large but noisily-labeled.

References

T. Chen, S. Kornblith, M. Norouzi, and G. E. Hinton. A simple framework for contrastive learning of visual representations. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2020. URL `https://arxiv.org/abs/2002.05709`.

B. Han, Q. Yao, X. Yu, G. Niu, M. Xu, I. W. H. Tsang, and M. Sugiyama. Co-teaching: Robust training of deep neural networks with extremely noisy labels. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

L. Jiang, Z. Zhou, T. Leung, L. J. Li, and F. Fei-Fei. Mentornet: Learning data-driven curriculum for very deep neural networks on corrupted labels. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2018.

et. al. Kermany. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131.e9, 2018. doi: 10.1016/j.cell.2018.02.010. URL https://doi.org/10.1016/j.cell.2018.02.010.

J. Li, R. Socher, and S. C. Hoi. Dividemix: Learning with noisy labels as semi-supervised learning. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020.

G. Patrini, A. Rozza, A. K. Menon, R. Nock, and L. Qu. Making deep neural networks robust to label noise: A loss correction approach. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

J. Wei, X. Feng, L. An, and L. Ma. Combating noisy labels by agreement: Joint co-regularization with co-regularized loss. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

Chuang Zhu, Wenkai Chen, Ting Peng, Ying Wang, and Mulan Jin. Hard sample aware noise robust learning for histopathology image classification. *IEEE transactions on medical imaging*.