RSG: A Simple but Effective Module for Learning Imbalanced Datasets Supplementary Material

Jianfeng Wang[1], Thomas Lukasiewicz[1], Xiaolin Hu[2], Jianfei Cai[3], Zhenghua Xu[4*]
University of Oxford[1] Tsinghua University[2] Monash University[3] Hebei University of Technology[4]
{jianfeng.wang,thomas.lukasiewicz}@cs.ox.ac.uk, xlhu@tsinghua.edu.cn, jianfei.cai@monash.edu
zhenghua.xu@hebut.edu.cn

1. Ablation Studies

Hyperparameter search. We searched the number of centers, the transfer strength, the frequent-class ratio, λ_1, and λ_2 for RSG by conducting experiments on ResNet-32 [2] with LDAM-DRW [1]. The results are shown in Fig. 1. We initially set the frequent-class ratio to 0.5 for step distributions and 0.2 for long-tailed distributions. At first, we explored different numbers of centers in each class.

As Fig. 1 shows, we obtain the lowest error rate in most cases when the number of centers is equal to 15. Then, considering that the transfer strength determines the number of newly generated samples, we explored how error rates change with different settings of the transfer strength. We conclude from Fig. 1 that, in general, the error rates gradually decline with the increase of transfer strength for long-tailed imbalanced distributions. As for step imbalanced distributions, using a stronger transfer strength always leads to higher error rates. In addition, we explored how frequent-class ratio impacts the performance RSG. As Fig. 1 shows, RSG achieves the best performance when α is set to 0.2 for long-tailed imbalanced distributions and 0.5 for step-imbalanced distributions. Lastly, we also verified the impacts brought by different settings of λ_1 and λ_2 separately. The results show that a superior performance is obtained when λ_1 and λ_2 are equal to 0.1 and 0.01, respectively.

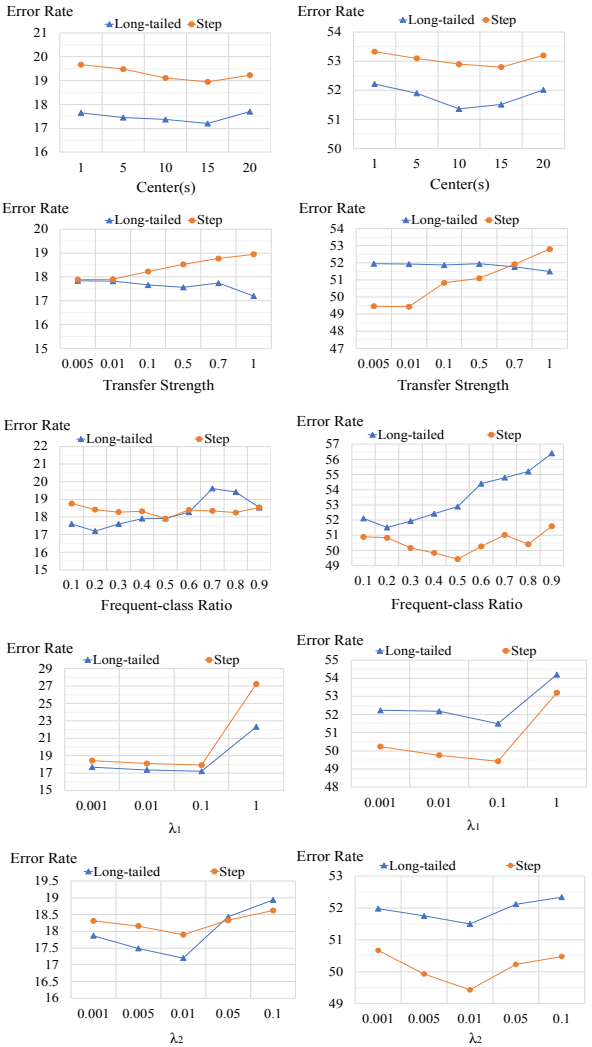

Figure 1. Top-1 error rates of ResNet-32 with RSG for different numbers of centers, transfer strengths, frequent-class ratios, λ_1, and λ_2 on Imbalanced CIFAR-10 (left column) and CIFAR-100 (right column) for $\rho = 50$.

References

[1] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma. Learning imbalanced datasets with label-distribution-aware margin loss. *Advances in Neural Information Processing Systems*, 2019. 1

[2] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Computer Vision and Pattern Recognition*, pages 770–778, 2016. 1

*Corresponding author.