Supplementary material for ”Zero-Shot Logit Adjustment”

Dubing Chen[1*] , **Yuming Shen**[2*] , **Haofeng Zhang**[1] , **Philip H.S. Torr**[2]
[1]Nanjing University of Science and Technology
[2]University of Oxford
{db.chen, zhanghf}@njust.edu.cn, ymcidence@gmail.com, philip.torr@eng.ox.ac.uk

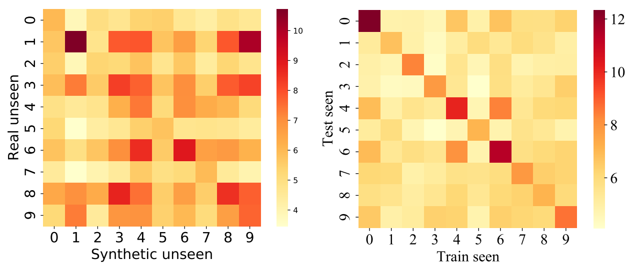

Figure A.1: Similarity hotmap on AWA2 [Lampert *et al.*, 2013]. **Left:** synthetic-real unseen class similarity. **Right:** train-test seen class similarity.

Experimental settings. We conduct the empirical analysis in Sec. 3.2 by comparing three methods, i.e., the single resampling technique, VAE [Kingma and Welling, 2013], and WGAN [Gulrajani *et al.*, 2017]. For the resampling technique, we train a semantic to visual-center mapping net with MSE loss to obtain the pseudo unseen class centers. The center of each unseen class were then replicated in 4000 copies. Similarly, we train the two generative networks (with similar architectures in [Xian *et al.*, 2018] and [Xian *et al.*, 2019]) to generate 4000 samples for each unseen class. The results are then compared using the classifier trained on real seen and generated unseen samples. The Adam optimizer is employed with a learning rate of 1×10^{-3}, and the batch size is set at 512 for all three methods.

Extra Discussion on Bias Problem. In Figure A.1 , we showcase the similarities between generating samples and real samples in the training domain and test domain. The heat map of class similarity between the generated unseen samples and the real sample is irregularly distributed. In contrast, the train-test class similarity of seen classes reflects the consistency of class relations in the training test phase.

Discriminability on Unseen Classes. Although we focus on Generalized Zero-Shot Learning, we display the discriminability of unseen classes as an additional measure for our method. Specifically, we train the classifier under the GZSL setting and use the trained classifier to perform ZSL classification. The confusion matrix is then plotted in Figure A.2 (a), revealing the outstanding discriminability of our method, even though it is not specially designed for ZSL. We conduct this experiment on WGAN, with a generation number of 10, and the σ is set to 1000.

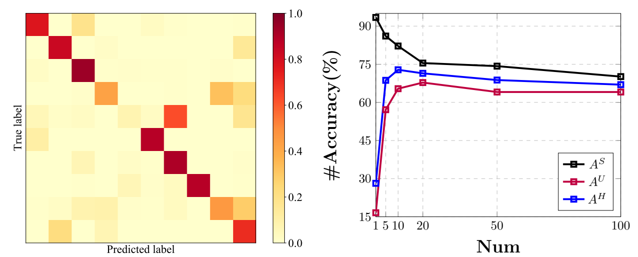

Figure A.2: **(a)** Confusion matrix on AWA2 [Lampert *et al.*, 2013]. We trained the classifier in GZSL setting and evaluate its discriminability on unseen classes. **(b)** Effect of the unseen generation numbers on AWA2.

Effect of generation numbers. As shown in Figure A.2 (b), the generation number plays a role in controlling the balance of accuracy between seen and unseen classes. As it increases from zero, $\mathcal{A}^H$ gradually rises, but soon starts to drop. This reflects the positive effect of generating unseen class samples, as well as the harm of over-generation, as discussed in main paper.

References

[Gulrajani *et al.*, 2017] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved training of wasserstein gans. *arXiv preprint arXiv:1704.00028*, 2017.

[Kingma and Welling, 2013] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *ICLR*, 2013.

[Lampert *et al.*, 2013] Christoph H Lampert, Hannes Nickisch, and Stefan Harmeling. Attribute-based classification for zero-shot visual object categorization. *IEEE TPAMI*, 36(3):453–465, 2013.

[Xian *et al.*, 2018] Yongqin Xian, Tobias Lorenz, Bernt Schiele, and Zeynep Akata. Feature generating networks for zero-shot learning. In *CVPR*, pages 5542–5551, 2018.

[Xian *et al.*, 2019] Yongqin Xian, Saurabh Sharma, Bernt Schiele, and Zeynep Akata. f-gan-d2: A feature generating framework for any-shot learning. In *CVPR*, pages 10275–10284, 2019.