Feature Space Augmentation for Long-Tailed Data: Supplementary Materials

Peng Chu[1⋆], Xiao Bian[2⋆], Shaopeng Liu[3], and Haibin Ling[4,1]

[1] Temple University, USA pchu@temple.edu
[2] Google Inc., USA xbian@google.com
[3] GE Research, USA sliu@ge.com
[4] Stony Brook University, USA hling@cs.stonybrook.edu

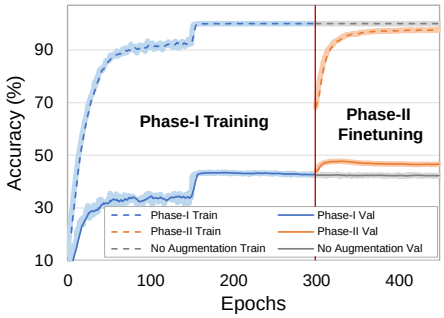
1 Learning Curve

In the supplementary material, we first present the learning curve of different network architectures on CIFAR-10 and CIFAR-100. We have observed a significant improvement compared to the baseline model training using conventional data augmentation (used in ImageNet ResNet model training) across different network architectures and dataset.

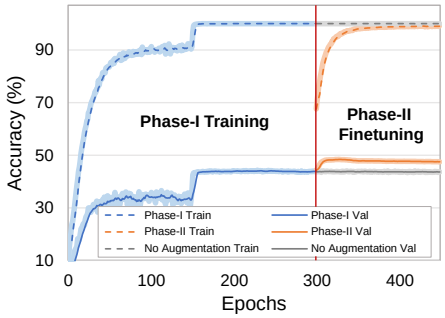
2 Feature Space Visualization

As discussed in Section 3.2, We present the scatter plot of feature decomposition of different network architectures on CIFAR-10 and CIFAR-100. It shows a generic trend that the class-specific features are significantly more separated compared to class-generic features from the same class. This result further substantiate the assumption that after separating the class-specific features, the remaining class-generic features from the head classes can be helpful to recover the loss information of the tail classes.

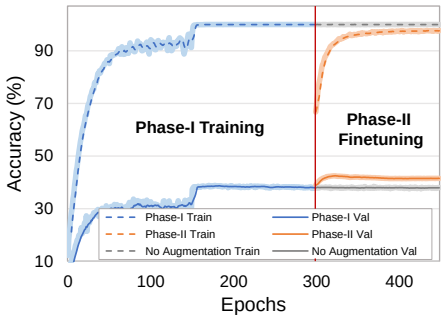
⋆ Work was done at GE Research.



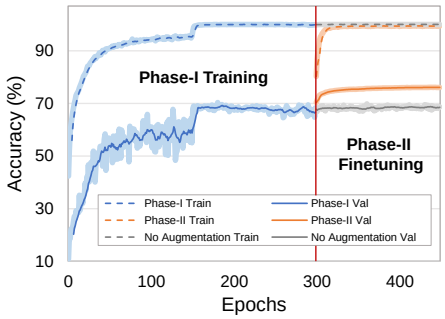
(a) CIFAR-100 with IM factor of 100 using ResNet-18



(b) CIFAR-100 with IM factor of 100 using ResNet-34

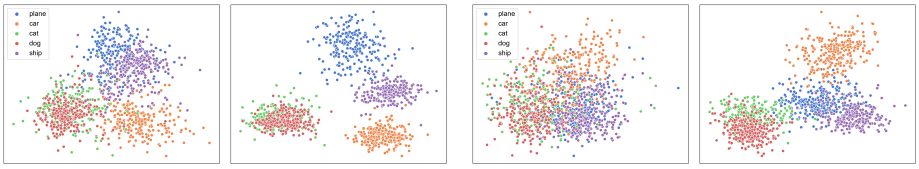


(c) CIFAR-100 with IM factor of 200 using ResNet-18



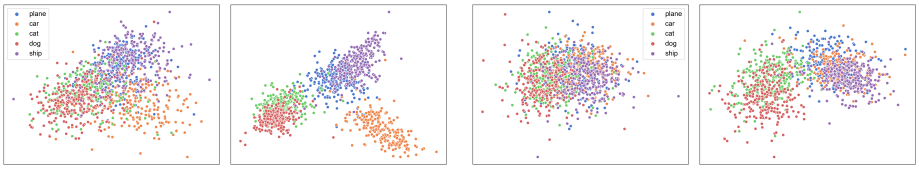
(d) CIFAR-10 with IM factor of 200 using ResNet-34

Fig. 1: Learning Curve



Class-Generic Class-Specific Class-Generic Class-Specific

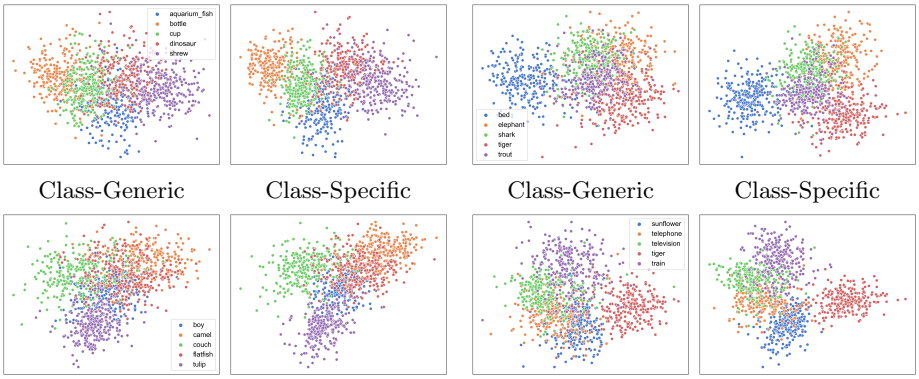
(a) ResNet-34 (b) DenseNet-121



Class-Generic Class-Specific Class-Generic Class-Specific

(c) MobileNetv2 (d) EfficientNet-B0

Fig. 2: CIFAR-10 Feature space visualization for different network architectures.



Class-Generic Class-Specific Class-Generic Class-Specific

Class-Generic Class-Specific Class-Generic Class-Specific

Fig. 3: Feature space visualization for different subset of CIFAR-100 using ResNet-18.