

Figure 4. Comparison of error curves with a batch size of 32 images/GPU. We show the ImageNet training error (left) and validation

Batch Norm (BN) Group Norm (GN) BN, 32 ims/gpu GN, 32 ims/gpu BN. 16 ims/gpu GN. 16 ims/gpu 55 55 BN, 8 ims/gpu GN, 8 ims/gpu 50 BN, 4 ims/gpu 50 GN, 4 ims/gpu BN, 2 ims/gpu GN, 2 ims/gpu 45 45 popolitic de la constitución de

error (right) vs. numbers of training epochs. The model is ResNet-50. 35 35 30 25

60

epochs

10 20 30

Figure 5. Sensitivity to batch sizes: ResNet-50's validation error of BN (left) and GN (right) trained with 32, 16, 8, 4, and 2 images/GPU.

10 20

epochs