



Figure 5. Training curves on ImageNet-1K. **(Left)**: ResNet/ResNeXt-50 with preserved complexity ( $\sim 4.1$  billion FLOPs,  $\sim 25$  million parameters); **(Right)**: ResNet/ResNeXt-101 with preserved complexity ( $\sim 7.8$  billion FLOPs,  $\sim 44$  million parameters).