

Evolution of loss during training

accuracy=0.8309

lr=0.01, betas=(0.9, 0.999), weight_decay=0.0005, hessian_power=1

