

Evolution of loss during training

accuracy=0.497

lr=0.001, betas=(0.9, 0.999), weight_decay=0.0001, hessian_power=0.5

