

Training on dataset: mnist with batch size: 128

