

**Pretraining loss after 10 epochs**

