Training Loss Curves (Representative Hyperparameters)

Legend:
- NP | batch=128 | lr=0.001
- NP | batch=128 | lr=0.01
- NP | batch=128 | lr=0.1
- NP | batch=32 | lr=0.001
- NP | batch=32 | lr=0.01
- NP | batch=32 | lr=0.1
- NP | batch=512 | lr=0.001
- NP | batch=512 | lr=0.01
- NP | batch=512 | lr=0.1
- TF | batch=128 | lr=0.001
- TF | batch=128 | lr=0.01
- TF | batch=128 | lr=0.1
- TF | batch=32 | lr=0.001
- TF | batch=32 | lr=0.01
- TF | batch=32 | lr=0.1
- TF | batch=512 | lr=0.001
- TF | batch=512 | lr=0.01
- TF | batch=512 | lr=0.1