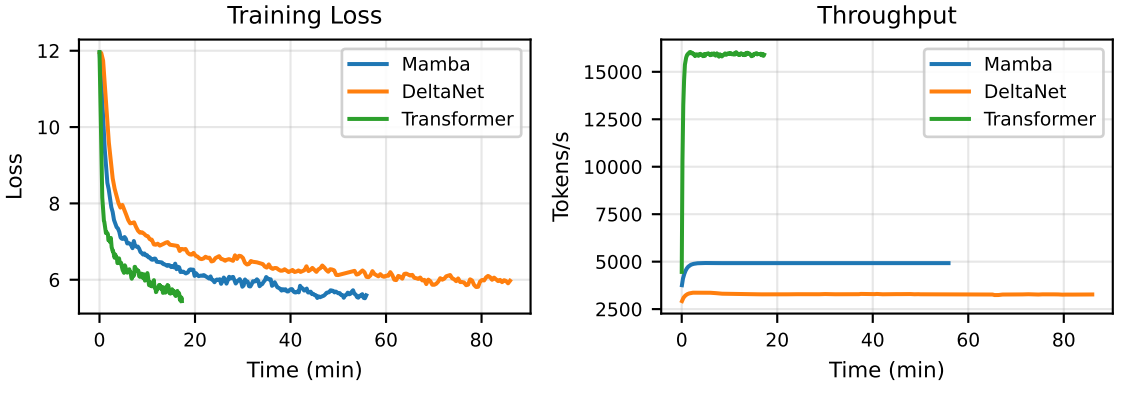
Training Loss
Loss
Time (min)
Mamba
DeltaNet
Transformer
12
10
8
6
0
20
40
60
80
Throughput
Tokens/s
Time (min)
Mamba
DeltaNet
Transformer
15000
12500
10000
7500
5000
2500
0
20
40
60
80