

Training loss

0.12
0.10
0.08
0.06
0.04
0.02

0

500

1000

1500

Number of parameters

