

Training loss

0.07  
0.06  
0.05  
0.04  
0.03  
0.02  
0.01

0

500

1000

1500

Number of parameters

