

Batch size

$2^{7.5}$

$2^7$

$2^{6.5}$

$2^6$

PadaDamp  
PadaDamp LR  
Adagrad

0

50,000

100,000

Model updates

