

$\times 10^{-7} + 7 \times 10^{-1}$

gradnorm

global\_gradnorm

$\text{lr} = 1.00 \times 10^{-07}$

$\text{lr} = 5.00 \times 10^{-08}$

$\text{lr} = 1.00 \times 10^{-08}$

Epoch

