

RMSNorm removed @ lr_max=3e-3 (diverges)

