三个网络参数初始化一致

Bn层在relu前后差不多，但比没有bn收敛要快很多, 放relu之后震荡稍小一点。

3769k，三个网络都带dropout









