

optimizer: SGD, $l_r = 0.01$, momentum = 0.2

