训练集损失函数 1.0 loss decode. loss_ce aux. loss_ce 0.8 0.6 0.4 0. 2 0.0 10000 20000 30000 40000 50000 60000 step