**Training with:**
**base_lr = 0.000100; stepsize = 50000; lr_mul = 1.000000 for new_mf**
**Finetuning: trial_5; Iter = 50000 replicating weights for higher dimensinality**

**Overall loss on all stages (min = 47.9456; iter = 9740)**

**Loss per stage (min = 3.1816; stage = 4)**