These models are all from tmux session 4, after about 144000 epoches training.