

Test loss for different models and sizes

