

Capacity during training process
for different data size, model size, and layers amount

