

capacity during training process for different data sizes, numbers of parameters, and numbers of layers

