Each parameter takes 4bytes =32bit float number.

1B parameter takes=4GB memory in GPU.

If you want to train each parameter will take extra 20byte

So each parameter takes 24byte to train.

1B Parameters will require GPU with memory size of 24GB.