Dropout(p=0.5)
[forward_compute_time=1.417,backward_compute_time=1.236,activation_size=2097152.0,parameter_size=0.0]

Linear(in_features=4096, out_features=1000, bias=True)
[forward_compute_time=1.045,backward_compute_time=1.014,activation_size=512000.0,parameter_size=16388000.0]

Linear(in_features=4096, out_features=4096, bias=True)
[forward_compute_time=2.593,backward_compute_time=1.488,activation_size=2097152.0,parameter_size=67125248.0]

ReLU(inplace)
[forward compute time=1.320,backward compute time=0.823,activation size=2097152.0,parameter size=0.0]