```
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0]
                                                      Conv2d(3, 64, kernel_size=(11, 11), stride=(4, 4), padding=(2, 2))
                             [forward_compute_time=40.304,backward_compute_time=15.142,activation_size=396492800.0,parameter_size=93184.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=1.737,backward_compute_time=2.213,activation_size=396492800.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=2.332,backward_compute_time=6.543,activation_size=95551488.0,parameter_size=0.0]
                                                      Conv2d(64, 192, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
                            [forward_compute_time=8.012,backward_compute_time=16.143,activation_size=286654464.0,parameter_size=1229568.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=1.252,backward_compute_time=1.533,activation_size=286654464.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=2.711,backward_compute_time=6.213,activation_size=66453504.0,parameter_size=0.0]
                                                      Conv2d(192, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=7.374,backward_compute_time=16.846,activation_size=132907008.0,parameter_size=2655744.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.615,backward_compute_time=0.715,activation_size=132907008.0,parameter_size=0.0]
                                                      Conv2d(384, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=13.925,backward_compute_time=20.786,activation_size=88604672.0,parameter_size=3539968.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.747,backward_compute_time=0.493,activation_size=88604672.0,parameter_size=0.0]
                                                      Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=11.384,backward_compute_time=16.060,activation_size=88604672.0,parameter_size=2360320.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.744,backward_compute_time=0.505,activation_size=88604672.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=1.129,backward_compute_time=1.724,activation_size=18874368.0,parameter_size=0.0]
                                            Size(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=4.0,parameter_size=0.0]
                                                                                View(9216)
                                 [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=18874368.0,parameter_size=0.0]
                                                                              Dropout(p=0.5)
                                 [forward_compute_time=0.161,backward_compute_time=0.904,activation_size=18874368.0,parameter_size=0.0]
                                                           Linear(in_features=9216, out_features=4096, bias=True)
                              [forward_compute_time=4.608,backward_compute_time=8.777,activation_size=8388608.0,parameter_size=151011328.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.096,backward_compute_time=0.087,activation_size=8388608.0,parameter_size=0.0]
                                                                              Dropout(p=0.5)
                                 [forward_compute_time=0.100,backward_compute_time=0.395,activation_size=8388608.0,parameter_size=0.0]
                                                           Linear(in_features=4096, out_features=4096, bias=True)
                              [forward_compute_time=2.187,backward_compute_time=3.940,activation_size=8388608.0,parameter_size=67125248.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.076,backward_compute_time=0.144,activation_size=8388608.0,parameter_size=0.0]
```

Linear(in\_features=4096, out\_features=1000, bias=True) [forward\_compute\_time=0.648,backward\_compute\_time=1.363,activation\_size=2048000.0,parameter\_size=16388000.0]

Input0