```
Input0
                                    [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=21.693,backward_compute_time=7.776,activation_size=198246400.0,parameter_size=93184.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.916,backward_compute_time=1.104,activation_size=198246400.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=1.068,backward_compute_time=3.273,activation_size=47775744.0,parameter_size=0.0]
                                                      Conv2d(64, 192, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
                             [forward_compute_time=4.168,backward_compute_time=8.841,activation_size=143327232.0,parameter_size=1229568.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.647,backward_compute_time=0.752,activation_size=143327232.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.783,backward_compute_time=2.507,activation_size=33226752.0,parameter_size=0.0]
                                                      Conv2d(192, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=3.730,backward_compute_time=9.255,activation_size=66453504.0,parameter_size=2655744.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.326,backward_compute_time=0.361,activation_size=66453504.0,parameter_size=0.0]
                                                      Conv2d(384, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=4.358,backward_compute_time=10.550,activation_size=44302336.0,parameter_size=3539968.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.234,backward_compute_time=0.254,activation_size=44302336.0,parameter_size=0.0]
                                                     Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                              [forward_compute_time=3.161,backward_compute_time=7.586,activation_size=44302336.0,parameter_size=2360320.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.235,backward_compute_time=0.262,activation_size=44302336.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.591,backward_compute_time=0.884,activation_size=9437184.0,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=9437184.0,parameter_size=0.0]
                                                                             Dropout(p=0.5)
                                 [forward_compute_time=0.559,backward_compute_time=0.870,activation_size=9437184.0,parameter_size=0.0]
                                                          Linear(in_features=9216, out_features=4096, bias=True)
                             [forward_compute_time=2.462,backward_compute_time=5.181,activation_size=4194304.0,parameter_size=151011328.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.058,backward_compute_time=0.055,activation_size=4194304.0,parameter_size=0.0]
                                                                             Dropout(p=0.5)
                                 [forward_compute_time=0.074,backward_compute_time=0.396,activation_size=4194304.0,parameter_size=0.0]
                                                          Linear(in features=4096, out features=4096, bias=True)
                             [forward_compute_time=1.118,backward_compute_time=2.051,activation_size=4194304.0,parameter_size=67125248.0]
                                                                              ReLU(inplace)
                                 [forward compute time=0.057,backward compute time=0.722,activation size=4194304.0,parameter size=0.0]
```

Linear(in_features=4096, out_features=1000, bias=True)
[forward_compute_time=0.458,backward_compute_time=1.023,activation_size=1024000.0,parameter_size=16388000.0]