```
Input
                               [forward_compute_time=635.902,backward_compute_time=0.000,activation_size=154140672.0,parameter_size=0.0]
                                                      Conv2d(3, 64, kernel_size=(11, 11), stride=(4, 4), padding=(2, 2))
                             [forward_compute_time=22.316,backward_compute_time=6.405,activation_size=198246400.0,parameter_size=93184.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.813,backward_compute_time=0.733,activation_size=198246400.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.756,backward_compute_time=1.148,activation_size=47775744.0,parameter_size=0.0]
                                                      Conv2d(64, 192, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
                             [forward_compute_time=3.128,backward_compute_time=5.612,activation_size=143327232.0,parameter_size=1229568.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.811,backward_compute_time=0.524,activation_size=143327232.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.579,backward_compute_time=0.861,activation_size=33226752.0,parameter_size=0.0]
                                                     Conv2d(192, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=3.260,backward_compute_time=6.610,activation_size=66453504.0,parameter_size=2655744.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.427,backward_compute_time=0.269,activation_size=66453504.0,parameter_size=0.0]
                                                     Conv2d(384, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                              [forward_compute_time=4.423,backward_compute_time=7.197,activation_size=44302336.0,parameter_size=3539968.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.539,backward_compute_time=0.188,activation_size=44302336.0,parameter_size=0.0]
                                                     Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=2.815,backward_compute_time=5.497,activation_size=44302336.0,parameter_size=2360320.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.524,backward_compute_time=0.176,activation_size=44302336.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.391,backward_compute_time=0.306,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=0.000,backward_compute_time=0.000,activation_size=9437184.0,parameter_size=0.0]
                                                          Linear(in_features=9216, out_features=4096, bias=True)
                             [forward_compute_time=1.737,backward_compute_time=3.072,activation_size=4194304.0,parameter_size=151011328.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.471,backward_compute_time=0.039,activation_size=4194304.0,parameter_size=0.0]
                                                                             Dropout(p=0.5)
                                 [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=4194304.0,parameter_size=0.0]
                                                           Linear(in features=4096, out features=4096, bias=True)
                              [forward_compute_time=1.075,backward_compute_time=1.366,activation_size=4194304.0,parameter_size=67125248.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.283,backward_compute_time=0.043,activation_size=4194304.0,parameter_size=0.0]
                                                           Linear(in_features=4096, out_features=1000, bias=True)
```

[forward_compute_time=0.453,backward_compute_time=0.474,activation_size=1024000.0,parameter_size=16388000.0]