```
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=2.767,backward compute time=0.060,activation size=774400.0,parameter size=93184.0]
                                                     ReLU()
    [forward compute time=0.981,backward compute time=0.522,activation size=774400.0,parameter size=0.0]
                   MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
    [forward compute time=1.042,backward compute time=1.260,activation size=186624.0,parameter size=0.0]
                         Conv2d(64, 192, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
[forward compute time=1.912,backward compute time=10.460,activation size=559872.0,parameter size=1229568.0]
                                                     ReLU()
    [forward compute time=0.839,backward compute time=0.498,activation size=559872.0,parameter size=0.0]
                   MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
    [forward compute time=1.014,backward compute time=1.137,activation size=129792.0,parameter size=0.0]
                         Conv2d(192, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward compute time=9.672,backward compute time=10.581,activation size=259584.0,parameter size=2655744.0]
                                                     ReLU()
    [forward compute time=0.830,backward compute time=0.844,activation size=259584.0,parameter size=0.0]
Conv2d(384, 256, kernel\_size=(3, 3), stride=(1, 1), padding=(1, 1)) \\ [forward\_compute\_time=2.734, backward\_compute\_time=12.222, activation\_size=173056.0, parameter\_size=3539968.0]
    [forward compute time=0.768,backward compute time=0.832,activation size=173056.0,parameter size=0.0]
                         Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=2.232,backward_compute_time=365.779,activation_size=173056.0,parameter_size=2360320.0]
                                                     ReLU()
    [forward compute time=0.939,backward compute time=0.504,activation size=173056.0,parameter size=0.0]
                   MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
     [forward compute time=0.983,backward compute time=0.618,activation size=36864.0,parameter size=0.0]
                                      AdaptiveAvgPool2d(output_size=(6, 6))
     [forward compute time=1.029,backward compute time=0.975,activation size=36864.0,parameter size=0.0]
                                                   View(9216)
     [forward compute time=0.000,backward compute time=0.000,activation size=36864.0,parameter size=0.0]
                                          Dropout(p=0.5, inplace=False)
     [forward compute time=0.928,backward compute time=0.857,activation size=36864.0,parameter size=0.0]
                             Linear(in features=9216, out features=4096, bias=True)
[forward compute time=6.004,backward compute time=58.731,activation size=16384.0,parameter size=151011328.0]
                                                     ReLU()
     [forward compute time=0.453,backward compute time=0.483,activation size=16384.0,parameter size=0.0]
                                          Dropout(p=0.5, inplace=False)
     [forward compute time=0.920,backward compute time=0.858,activation size=16384.0,parameter size=0.0]
                             Linear(in features=4096, out features=4096, bias=True)
[forward compute time=6.028,backward compute time=27.029,activation size=16384.0,parameter size=67125248.0]
                                                     ReLU()
     [forward compute time=0.931,backward compute time=0.845,activation size=16384.0,parameter size=0.0]
                             Linear(in features=4096, out features=1000, bias=True)
```

[forward compute time=1.604,backward compute time=11.693,activation size=4000.0,parameter size=16388000.0]