```
Input0
                                     [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0]
                                                       Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=58.933,backward_compute_time=38.616,activation_size=822083584.0,parameter_size=7168.0]
                                                                           ReLU(inplace=True)
                               [forward_compute_time=11.547,backward_compute_time=14.869,activation_size=822083584.0,parameter_size=0.0]
                                                      Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=82.672,backward_compute_time=211.782,activation_size=822083584.0,parameter_size=147712.0]
                                                                           ReLU(inplace=True)
                               [forward compute time=12.481,backward compute time=15.067,activation size=822083584.0,parameter size=0.0]
                                                MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                               [forward_compute_time=11.405,backward_compute_time=27.718,activation_size=205520896.0,parameter_size=0.0]
                                                      Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=37.589,backward_compute_time=87.315,activation_size=411041792.0,parameter_size=295424.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=6.557,backward_compute_time=7.574,activation_size=411041792.0,parameter_size=0.0]
                                                      Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward compute time=52.021,backward compute time=145.583,activation size=411041792.0,parameter size=590336.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=5.952,backward_compute_time=7.425,activation_size=411041792.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                               [forward_compute_time=6.237,backward_compute_time=14.165,activation_size=102760448.0,parameter_size=0.0]
                                                      Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=25.814,backward_compute_time=64.631,activation_size=205520896.0,parameter_size=1180672.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=3.184,backward_compute_time=3.735,activation_size=205520896.0,parameter_size=0.0]
                                                      Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=37.672,backward_compute_time=74.974,activation_size=205520896.0,parameter_size=2360320.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=2.757,backward_compute_time=3.817,activation_size=205520896.0,parameter_size=0.0]
                                                      Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=36.954,backward_compute_time=94.136,activation_size=205520896.0,parameter_size=2360320.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=2.998,backward_compute_time=3.882,activation_size=205520896.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=2.699,backward_compute_time=7.635,activation_size=51380224.0,parameter_size=0.0]
                                                     Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=19.080,backward_compute_time=42.599,activation_size=102760448.0,parameter_size=4720640.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=1.389,backward_compute_time=2.013,activation_size=102760448.0,parameter_size=0.0]
                                                     Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=31.508,backward_compute_time=55.320,activation_size=102760448.0,parameter_size=9439232.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=1.391,backward_compute_time=2.016,activation_size=102760448.0,parameter_size=0.0]
                                                     Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                           [forward_compute_time=30.900,backward_compute_time=72.992,activation_size=102760448.0,parameter_size=9439232.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=1.874,backward_compute_time=1.904,activation_size=102760448.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=1.410,backward_compute_time=4.336,activation_size=25690112.0,parameter_size=0.0]
                                                     Conv2d(512, 512, kernel size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=9.751,backward_compute_time=18.226,activation_size=25690112.0,parameter_size=9439232.0]
                                                                           ReLU(inplace=True)
                                 [forward_compute_time=0.390,backward_compute_time=0.633,activation_size=25690112.0,parameter_size=0.0]
                                                     Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=9.540,backward_compute_time=18.215,activation_size=25690112.0,parameter_size=9439232.0]
                                                                           ReLU(inplace=True)
                                 [forward_compute_time=0.389,backward_compute_time=0.632,activation_size=25690112.0,parameter_size=0.0]
                                                     Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                             [forward_compute_time=9.680,backward_compute_time=21.762,activation_size=25690112.0,parameter_size=9439232.0]
                                                                           ReLU(inplace=True)
                                [forward_compute_time=0.388,backward_compute_time=0.522,activation_size=25690112.0,parameter_size=0.0]
                                                MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.418,backward_compute_time=8.582,activation_size=6422528.0,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(-1)
                                 [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=6422528.0,parameter_size=0.0]
                                                          Linear(in features=25088, out features=4096, bias=True)
                            [forward compute time=9.999,backward compute time=15.379,activation size=1048576.0,parameter size=411058176.0]
                                                                           ReLU(inplace=True)
                                 [forward_compute_time=0.080,backward_compute_time=0.082,activation_size=1048576.0,parameter_size=0.0]
                                                                      Dropout(p=0.5, inplace=False)
                                 [forward_compute_time=0.165,backward_compute_time=1.228,activation_size=1048576.0,parameter_size=0.0]
                                                          Linear(in_features=4096, out_features=4096, bias=True)
                             [forward compute time=1.743,backward compute time=2.657,activation size=1048576.0,parameter size=67125248.0]
                                                                           ReLU(inplace=True)
                                 [forward_compute_time=0.076,backward_compute_time=0.086,activation_size=1048576.0,parameter_size=0.0]
```

Dropout(p=0.5, inplace=False) [forward_compute_time=0.114,backward_compute_time=0.323,activation_size=1048576.0,parameter_size=0.0]

Linear(in_features=4096, out_features=1000, bias=True) [forward_compute_time=0.778,backward_compute_time=3.336,activation_size=256000.0,parameter_size=16388000.0]