```
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=18.470,backward_compute_time=18.086,activation_size=822083584.0,parameter_size=7168.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=3.445,backward_compute_time=4.405,activation_size=822083584.0,parameter_size=0.0]
                                                       Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=24.239,backward_compute_time=66.530,activation_size=822083584.0,parameter_size=147712.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=3.432,backward_compute_time=4.141,activation_size=822083584.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                [forward_compute_time=3.209,backward_compute_time=7.185,activation_size=205520896.0,parameter_size=0.0]
                                                       Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=11.263,backward_compute_time=28.048,activation_size=411041792.0,parameter_size=295424.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=1.736,backward_compute_time=2.019,activation_size=411041792.0,parameter_size=0.0]
                                                      Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=15.709,backward_compute_time=36.826,activation_size=411041792.0,parameter_size=590336.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=1.736,backward_compute_time=2.050,activation_size=411041792.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                [forward_compute_time=1.555,backward_compute_time=3.607,activation_size=102760448.0,parameter_size=0.0]
                                                      Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=7.195,backward_compute_time=14.548,activation_size=205520896.0,parameter_size=1180672.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.888,backward_compute_time=1.022,activation_size=205520896.0,parameter_size=0.0]
                                                      Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=11.011,backward_compute_time=21.543,activation_size=205520896.0,parameter_size=2360320.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.979,backward_compute_time=1.015,activation_size=205520896.0,parameter_size=0.0]
                                                      Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=12.230,backward_compute_time=24.646,activation_size=205520896.0,parameter_size=2360320.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.927,backward_compute_time=1.024,activation_size=205520896.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.809,backward_compute_time=1.832,activation_size=51380224.0,parameter_size=0.0]
                                                      Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=5.243,backward_compute_time=10.490,activation_size=102760448.0,parameter_size=4720640.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.471,backward_compute_time=0.545,activation_size=102760448.0,parameter_size=0.0]
                                                      Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=8.738,backward_compute_time=15.974,activation_size=102760448.0,parameter_size=9439232.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.466,backward_compute_time=0.554,activation_size=102760448.0,parameter_size=0.0]
                                                      Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                            [forward_compute_time=8.730,backward_compute_time=19.089,activation_size=102760448.0,parameter_size=9439232.0]
                                                                              ReLU(inplace)
                                [forward_compute_time=0.460,backward_compute_time=0.529,activation_size=102760448.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                 [forward_compute_time=0.430,backward_compute_time=0.994,activation_size=25690112.0,parameter_size=0.0]
                                                      Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                              [forward_compute_time=2.843,backward_compute_time=5.328,activation_size=25690112.0,parameter_size=9439232.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.277,backward_compute_time=0.176,activation_size=25690112.0,parameter_size=0.0]
                                                      Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                              [forward_compute_time=2.831,backward_compute_time=5.338,activation_size=25690112.0,parameter_size=9439232.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.140,backward_compute_time=0.176,activation_size=25690112.0,parameter_size=0.0]
                                                      Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
                              [forward_compute_time=2.829,backward_compute_time=6.305,activation_size=25690112.0,parameter_size=9439232.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.140,backward_compute_time=0.157,activation_size=25690112.0,parameter_size=0.0]
                                                 MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
                                  [forward_compute_time=0.137,backward_compute_time=2.501,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=3.375,backward_compute_time=3.424,activation_size=1048576.0,parameter_size=411058176.0]
                                                                              ReLU(inplace)
                                 [forward_compute_time=0.040,backward_compute_time=0.033,activation_size=1048576.0,parameter_size=0.0]
                                                                              Dropout(p=0.5)
                                  [forward_compute_time=0.087,backward_compute_time=0.378,activation_size=1048576.0,parameter_size=0.0]
                                                           Linear(in_features=4096, out_features=4096, bias=True)
                              [forward_compute_time=0.405,backward_compute_time=0.617,activation_size=1048576.0,parameter_size=67125248.0]
                                                                              ReLU(inplace)
                                  [forward compute time=0.037,backward compute time=0.037,activation size=1048576.0,parameter size=0.0]
```

Dropout(p=0.5)
[forward_compute_time=0.057,backward_compute_time=0.241,activation_size=1048576.0,parameter_size=0.0]

Linear(in_features=4096, out_features=1000, bias=True) [forward_compute_time=0.203,backward_compute_time=1.019,activation_size=256000.0,parameter_size=16388000.0]