```
Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 [forward_compute_time=22.307,backward_compute_time=24.613,activation_size=1644167168.0,parameter_size=7168.0]
                                                  ReLU(inplace)
    [forward_compute_time=4.446,backward_compute_time=5.553,activation_size=1644167168.0,parameter_size=0.0]
                           Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=46.201,backward_compute_time=113.330,activation_size=1644167168.0,parameter_size=147712.0]
                                                  ReLU(inplace)
    [forward_compute_time=4.455,backward_compute_time=5.311,activation_size=1644167168.0,parameter_size=0.0]
                     MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    [forward_compute_time=3.795,backward_compute_time=5.579,activation_size=411041792.0,parameter_size=0.0]
                          Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 [forward_compute_time=15.155,backward_compute_time=39.648,activation_size=822083584.0,parameter_size=295424.0]
                                                  ReLU(inplace)
    [forward_compute_time=2.359,backward_compute_time=2.738,activation_size=822083584.0,parameter_size=0.0]
                          Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward compute time=22.062,backward compute time=53.379,activation size=822083584.0,parameter size=590336.0]
                                                  ReLU(inplace)
    [forward compute time=2.513,backward compute time=2.620,activation size=822083584.0,parameter size=0.0]
                     MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    [forward_compute_time=2.211,backward_compute_time=2.788,activation_size=205520896.0,parameter_size=0.0]
                          Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=10.814,backward_compute_time=18.335,activation_size=411041792.0,parameter_size=1180672.0]
                                                  ReLU(inplace)
    [forward_compute_time=1.377,backward_compute_time=1.380,activation_size=411041792.0,parameter_size=0.0]
                          Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=17.458,backward_compute_time=27.023,activation_size=411041792.0,parameter_size=2360320.0]
                                                  ReLU(inplace)
    [forward_compute_time=1.640,backward_compute_time=1.463,activation_size=411041792.0,parameter_size=0.0]
                          Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=14.245,backward_compute_time=33.888,activation_size=411041792.0,parameter_size=2360320.0]
                                                  ReLU(inplace)
    [forward_compute_time=1.549,backward_compute_time=1.321,activation_size=411041792.0,parameter_size=0.0]
                     MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    [forward_compute_time=1.169,backward_compute_time=1.513,activation_size=102760448.0,parameter_size=0.0]
                          Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 [forward_compute_time=9.249,backward_compute_time=14.574,activation_size=205520896.0,parameter_size=4720640.0]
                                                  ReLU(inplace)
    [forward_compute_time=1.026,backward_compute_time=0.696,activation_size=205520896.0,parameter_size=0.0]
                          Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward compute time=15.359,backward compute time=21.900,activation size=205520896.0,parameter size=9439232.0]
                                                  ReLU(inplace)
    [forward_compute_time=0.848,backward_compute_time=0.679,activation_size=205520896.0,parameter_size=0.0]
                          Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=11.775,backward_compute_time=28.670,activation_size=205520896.0,parameter_size=9439232.0]
                                                  ReLU(inplace)
    [forward_compute_time=0.881,backward_compute_time=0.642,activation_size=205520896.0,parameter_size=0.0]
                     MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
     [forward_compute_time=0.802,backward_compute_time=0.781,activation_size=51380224.0,parameter_size=0.0]
                          Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  [forward_compute_time=5.608,backward_compute_time=7.381,activation_size=51380224.0,parameter_size=9439232.0]
                                                  ReLU(inplace)
     [forward_compute_time=0.784,backward_compute_time=0.266,activation_size=51380224.0,parameter_size=0.0]
                          Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 [forward_compute_time=4.070,backward_compute_time=7.396,activation_size=51380224.0,parameter_size=9439232.0]
                                                  ReLU(inplace)
     [forward_compute_time=0.463,backward_compute_time=0.190,activation_size=51380224.0,parameter_size=0.0]
                          Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  [forward_compute_time=4.042,backward_compute_time=8.648,activation_size=51380224.0,parameter_size=9439232.0]
                                                  ReLU(inplace)
     [forward_compute_time=0.455,backward_compute_time=0.175,activation_size=51380224.0,parameter_size=0.0]
                     MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
     [forward_compute_time=0.553,backward_compute_time=0.234,activation_size=12845056.0,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=12845056.0,parameter_size=0.0]
                               Linear(in_features=25088, out_features=4096, bias=True)
 [forward_compute_time=2.475,backward_compute_time=4.575,activation_size=2097152.0,parameter_size=411058176.0]
                                                  ReLU(inplace)
      [forward_compute_time=0.377,backward_compute_time=0.030,activation_size=2097152.0,parameter_size=0.0]
                                                  Dropout(p=0.5)
      [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=2097152.0,parameter_size=0.0]
                               Linear(in_features=4096, out_features=4096, bias=True)
  [forward_compute_time=0.709,backward_compute_time=0.734,activation_size=2097152.0,parameter_size=67125248.0]
                                                  ReLU(inplace)
      [forward_compute_time=0.277,backward_compute_time=0.032,activation_size=2097152.0,parameter_size=0.0]
                                                  Dropout(p=0.5)
      [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=2097152.0,parameter_size=0.0]
```

Linear(in\_features=4096, out\_features=1000, bias=True) [forward\_compute\_time=0.393,backward\_compute\_time=0.548,activation\_size=512000.0,parameter\_size=16388000.0]

Input [forward compute time=17.972,backward compute time=0.000,activation size=77070336.0,parameter size=0.0]