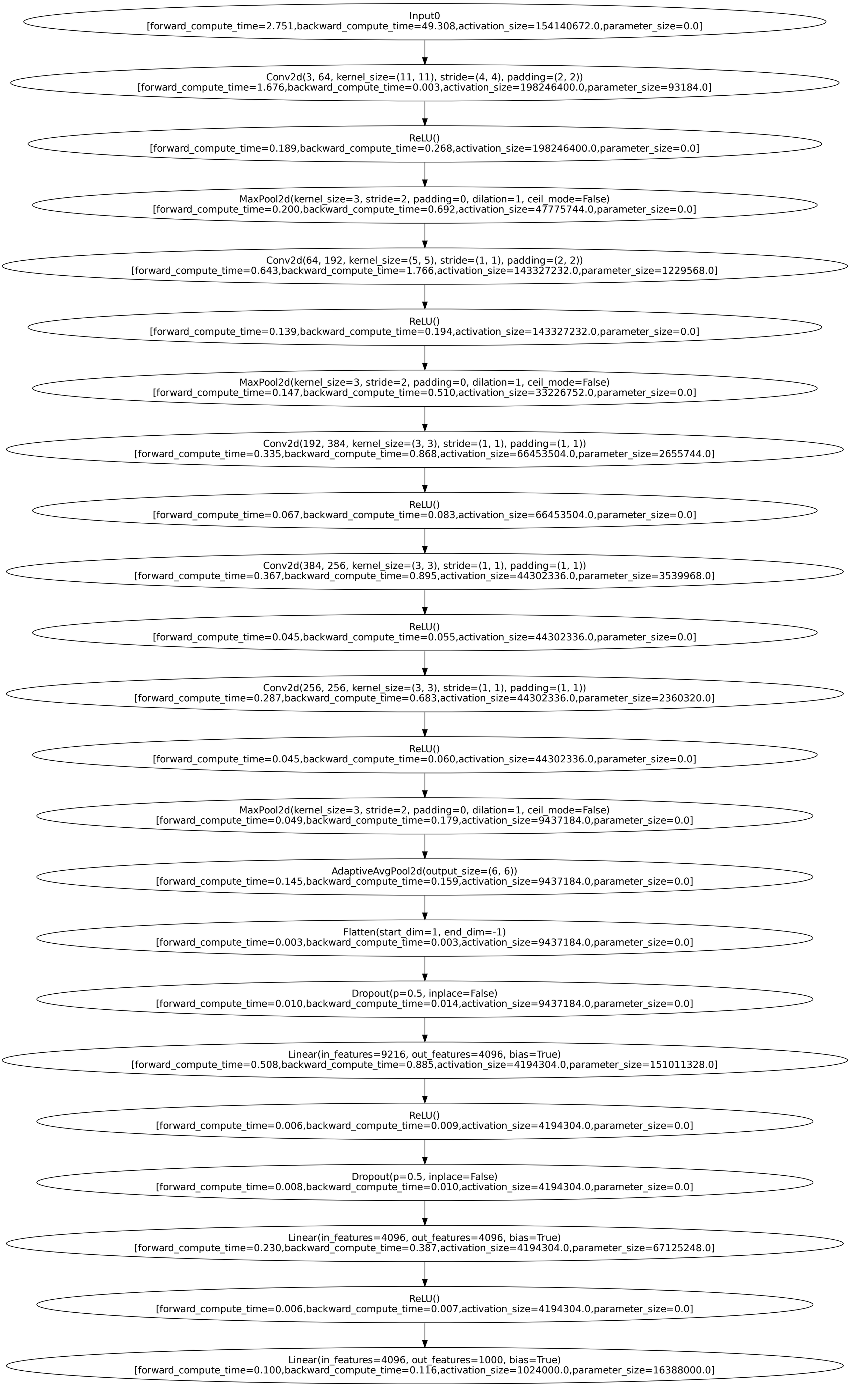
Input0
[forward_compute_time=2.751,backward_compute_time=49.308,activation_size=154140672.0,parameter_size=0.0]
Conv2d(3, 64, kernel_size=(11, 11), stride=(4, 4), padding=(2, 2))
[forward_compute_time=1.676,backward_compute_time=0.003,activation_size=198246400.0,parameter_size=93184.0]
ReLU()
[forward_compute_time=0.189,backward_compute_time=0.268,activation_size=198246400.0,parameter_size=0.0]
MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
[forward_compute_time=0.200,backward_compute_time=0.692,activation_size=47775744.0,parameter_size=0.0]
Conv2d(64, 192, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
[forward_compute_time=0.643,backward_compute_time=1.766,activation_size=143327232.0,parameter_size=1229568.0]
ReLU()
[forward_compute_time=0.139,backward_compute_time=0.194,activation_size=143327232.0,parameter_size=0.0]
MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
[forward_compute_time=0.147,backward_compute_time=0.510,activation_size=33226752.0,parameter_size=0.0]
Conv2d(192, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=0.335,backward_compute_time=0.868,activation_size=66453504.0,parameter_size=2655744.0]
ReLU()
[forward_compute_time=0.067,backward_compute_time=0.083,activation_size=66453504.0,parameter_size=0.0]
Conv2d(384, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=0.367,backward_compute_time=0.895,activation_size=44302336.0,parameter_size=3539968.0]
ReLU()
[forward_compute_time=0.045,backward_compute_time=0.055,activation_size=44302336.0,parameter_size=0.0]
Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
[forward_compute_time=0.287,backward_compute_time=0.683,activation_size=44302336.0,parameter_size=2360320.0]
ReLU()
[forward_compute_time=0.045,backward_compute_time=0.060,activation_size=44302336.0,parameter_size=0.0]
MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=False)
[forward_compute_time=0.049,backward_compute_time=0.179,activation_size=9437184.0,parameter_size=0.0]
AdaptiveAvgPool2d(output_size=(6, 6))
[forward_compute_time=0.145,backward_compute_time=0.159,activation_size=9437184.0,parameter_size=0.0]
Flatten(start_dim=1, end_dim=-1)
[forward_compute_time=0.003,backward_compute_time=0.003,activation_size=9437184.0,parameter_size=0.0]
Dropout(p=0.5, inplace=False)
[forward_compute_time=0.010,backward_compute_time=0.014,activation_size=9437184.0,parameter_size=0.0]
Linear(in_features=9216, out_features=4096, bias=True)
[forward_compute_time=0.508,backward_compute_time=0.885,activation_size=4194304.0,parameter_size=151011328.0]
ReLU()
[forward_compute_time=0.006,backward_compute_time=0.009,activation_size=4194304.0,parameter_size=0.0]
Dropout(p=0.5, inplace=False)
[forward_compute_time=0.008,backward_compute_time=0.010,activation_size=4194304.0,parameter_size=0.0]
Linear(in_features=4096, out_features=4096, bias=True)
[forward_compute_time=0.230,backward_compute_time=0.387,activation_size=4194304.0,parameter_size=67125248.0]
ReLU()
[forward_compute_time=0.006,backward_compute_time=0.007,activation_size=4194304.0,parameter_size=0.0]
Linear(in_features=4096, out_features=1000, bias=True)
[forward_compute_time=0.100,backward_compute_time=0.116,activation_size=1024000.0,parameter_size=16388000.0]