```
TSN(
  (base_model): InceptionV3(
    (conv_Conv2D): Conv2d(3, 32, kernel_size=(3, 3), stride=(2, 2))
    (conv_batchnorm): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv): ReLU(inplace=True)
    (conv_1_Conv2D): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1))
    (conv_1_batchnorm): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv_1): ReLU(inplace=True)
    (conv_2_Conv2D): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (conv_2_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv_2): ReLU(inplace=True)
    (pool): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=True)
    (conv_3_Conv2D): Conv2d(64, 80, kernel_size=(1, 1), stride=(1, 1))
    (conv_3_batchnorm): BatchNorm2d(80, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv_3): ReLU(inplace=True)
    (conv_4_Conv2D): Conv2d(80, 192, kernel_size=(3, 3), stride=(1, 1))
    (conv_4_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv_4): ReLU(inplace=True)
    (pool_1): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=True)
    (mixed_conv_Conv2D): Conv2d(192, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_conv): ReLU(inplace=True)
    (mixed_tower_conv_Conv2D): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1))
    (mixed_tower_conv_batchnorm): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_conv): ReLU(inplace=True)
    (mixed_tower_conv_1_Conv2D): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
    (mixed_tower_conv_1_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_conv_1): ReLU(inplace=True)
    (mixed_tower_1_conv_Conv2D): Conv2d(192, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_tower_1_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_1_conv): ReLU(inplace=True)
    (mixed_tower_1_conv_1_Conv2D): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_tower_1_conv_1_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_1_conv_1): ReLU(inplace=True)
    (mixed_tower_1_conv_2_Conv2D): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_tower_1_conv_2_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_1_conv_2): ReLU(inplace=True)
    (mixed_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
    (mixed_tower_2_conv_Conv2D): Conv2d(192, 32, kernel_size=(1, 1), stride=(1, 1))
    (mixed_tower_2_conv_batchnorm): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_tower_2_conv): ReLU(inplace=True)
    (mixed_1_conv_Conv2D): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_1_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_conv): ReLU(inplace=True)
    (mixed_1_tower_conv_Conv2D): Conv2d(256, 48, kernel_size=(1, 1), stride=(1, 1))
    (mixed_1_tower_conv_batchnorm): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_conv): ReLU(inplace=True)
    (mixed_1_tower_conv_1_Conv2D): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
    (mixed_1_tower_conv_1_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_conv_1): ReLU(inplace=True)
    (mixed_1_tower_1_conv_Conv2D): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_1_tower_1_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_1_conv): ReLU(inplace=True)
    (mixed_1_tower_1_conv_1_Conv2D): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_1_tower_1_conv_1_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_1_conv_1): ReLU(inplace=True)
    (mixed_1_tower_1_conv_2_Conv2D): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_1_tower_1_conv_2_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_1_conv_2): ReLU(inplace=True)
    (mixed_1_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
    (mixed_1_tower_2_conv_Conv2D): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_1_tower_2_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_1_tower_2_conv): ReLU(inplace=True)
    (mixed_2_conv_Conv2D): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1))
    (mixed_2_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_2_conv): ReLU(inplace=True)
    (mixed_2_tower_conv_Conv2D): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1))
    (mixed_2_tower_conv_batchnorm): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_2_tower_conv): ReLU(inplace=True)
    (mixed_2_tower_conv_1_Conv2D): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2))
    (mixed_2_tower_conv_1_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_2_tower_conv_1): ReLU(inplace=True)
```

(mixed_2_tower_1_conv_Conv2D): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1))
(mixed_2_tower_1_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_2_tower_1_conv): ReLU(inplace=True)
(mixed_2_tower_1_conv_1_Conv2D): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(mixed_2_tower_1_conv_1_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_2_tower_1_conv_1): ReLU(inplace=True)
(mixed_2_tower_1_conv_2_Conv2D): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(mixed_2_tower_1_conv_2_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_2_tower_1_conv_2): ReLU(inplace=True)
(mixed_2_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
(mixed_2_tower_2_conv_Conv2D): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1))
(mixed_2_tower_2_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_2_tower_2_conv): ReLU(inplace=True)
(mixed_3_conv_Conv2D): Conv2d(288, 384, kernel_size=(3, 3), stride=(2, 2))
(mixed_3_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_3_conv): ReLU(inplace=True)
(mixed_3_tower_conv_Conv2D): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1))
(mixed_3_tower_conv_batchnorm): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_3_tower_conv): ReLU(inplace=True)
(mixed_3_tower_conv_1_Conv2D): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(mixed_3_tower_conv_1_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_3_tower_conv_1): ReLU(inplace=True)
(mixed_3_tower_conv_2_Conv2D): Conv2d(96, 96, kernel_size=(3, 3), stride=(2, 2))
(mixed_3_tower_conv_2_batchnorm): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_3_tower_conv_2): ReLU(inplace=True)
(mixed_3_pool): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=True)
(mixed_4_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_4_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_conv): ReLU(inplace=True)
(mixed_4_tower_conv_Conv2D): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1))
(mixed_4_tower_conv_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_conv): ReLU(inplace=True)
(mixed_4_tower_conv_1_Conv2D): Conv2d(128, 128, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_4_tower_conv_1_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_conv_1): ReLU(inplace=True)
(mixed_4_tower_conv_2_Conv2D): Conv2d(128, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_4_tower_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_conv_2): ReLU(inplace=True)
(mixed_4_tower_1_conv_Conv2D): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1))
(mixed_4_tower_1_conv_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_1_conv): ReLU(inplace=True)
(mixed_4_tower_1_conv_1_Conv2D): Conv2d(128, 128, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_4_tower_1_conv_1_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_1_conv_1): ReLU(inplace=True)
(mixed_4_tower_1_conv_2_Conv2D): Conv2d(128, 128, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_4_tower_1_conv_2_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_1_conv_2): ReLU(inplace=True)
(mixed_4_tower_1_conv_3_Conv2D): Conv2d(128, 128, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_4_tower_1_conv_3_batchnorm): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_1_conv_3): ReLU(inplace=True)
(mixed_4_tower_1_conv_4_Conv2D): Conv2d(128, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_4_tower_1_conv_4_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_1_conv_4): ReLU(inplace=True)
(mixed_4_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
(mixed_4_tower_2_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_4_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_4_tower_2_conv): ReLU(inplace=True)
(mixed_5_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_5_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_conv): ReLU(inplace=True)
(mixed_5_tower_conv_Conv2D): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1))
(mixed_5_tower_conv_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_conv): ReLU(inplace=True)
(mixed_5_tower_conv_1_Conv2D): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_5_tower_conv_1_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_conv_1): ReLU(inplace=True)
(mixed_5_tower_conv_2_Conv2D): Conv2d(160, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_5_tower_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_conv_2): ReLU(inplace=True)
(mixed_5_tower_1_conv_Conv2D): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1))
(mixed_5_tower_1_conv_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_1_conv): ReLU(inplace=True)

(mixed_5_tower_1_conv_1_Conv2D): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_5_tower_1_conv_1_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_1_conv_1): ReLU(inplace=True)
(mixed_5_tower_1_conv_2_Conv2D): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_5_tower_1_conv_2_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_1_conv_2): ReLU(inplace=True)
(mixed_5_tower_1_conv_3_Conv2D): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_5_tower_1_conv_3_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_1_conv_3): ReLU(inplace=True)
(mixed_5_tower_1_conv_4_Conv2D): Conv2d(160, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_5_tower_1_conv_4_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_1_conv_4): ReLU(inplace=True)
(mixed_5_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
(mixed_5_tower_2_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_5_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_5_tower_2_conv): ReLU(inplace=True)
(mixed_6_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_6_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_conv): ReLU(inplace=True)
(mixed_6_tower_conv_Conv2D): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1))
(mixed_6_tower_conv_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_conv): ReLU(inplace=True)
(mixed_6_tower_conv_1_Conv2D): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_6_tower_conv_1_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_conv_1): ReLU(inplace=True)
(mixed_6_tower_conv_2_Conv2D): Conv2d(160, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_6_tower_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_conv_2): ReLU(inplace=True)
(mixed_6_tower_1_conv_Conv2D): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1))
(mixed_6_tower_1_conv_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_1_conv): ReLU(inplace=True)
(mixed_6_tower_1_conv_1_Conv2D): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_6_tower_1_conv_1_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_1_conv_1): ReLU(inplace=True)
(mixed_6_tower_1_conv_2_Conv2D): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_6_tower_1_conv_2_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_1_conv_2): ReLU(inplace=True)
(mixed_6_tower_1_conv_3_Conv2D): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_6_tower_1_conv_3_batchnorm): BatchNorm2d(160, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_1_conv_3): ReLU(inplace=True)
(mixed_6_tower_1_conv_4_Conv2D): Conv2d(160, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_6_tower_1_conv_4_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_1_conv_4): ReLU(inplace=True)
(mixed_6_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
(mixed_6_tower_2_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_6_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_6_tower_2_conv): ReLU(inplace=True)
(mixed_7_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_7_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_conv): ReLU(inplace=True)
(mixed_7_tower_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_7_tower_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_conv): ReLU(inplace=True)
(mixed_7_tower_conv_1_Conv2D): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_7_tower_conv_1_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_conv_1): ReLU(inplace=True)
(mixed_7_tower_conv_2_Conv2D): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_7_tower_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_conv_2): ReLU(inplace=True)
(mixed_7_tower_1_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
(mixed_7_tower_1_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_1_conv): ReLU(inplace=True)
(mixed_7_tower_1_conv_1_Conv2D): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_7_tower_1_conv_1_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_1_conv_1): ReLU(inplace=True)
(mixed_7_tower_1_conv_2_Conv2D): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
(mixed_7_tower_1_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_1_conv_2): ReLU(inplace=True)
(mixed_7_tower_1_conv_3_Conv2D): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
(mixed_7_tower_1_conv_3_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(mixed_7_tower_1_conv_3): ReLU(inplace=True)
(mixed_7_tower_1_conv_4_Conv2D): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))

```
    (mixed_7_tower_1_conv_4_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_7_tower_1_conv_4): ReLU(inplace=True)
    (mixed_7_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
    (mixed_7_tower_2_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
    (mixed_7_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_7_tower_2_conv): ReLU(inplace=True)
    (mixed_8_tower_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
    (mixed_8_tower_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_conv): ReLU(inplace=True)
    (mixed_8_tower_conv_1_Conv2D): Conv2d(192, 320, kernel_size=(3, 3), stride=(2, 2))
    (mixed_8_tower_conv_1_batchnorm): BatchNorm2d(320, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_conv_1): ReLU(inplace=True)
    (mixed_8_tower_1_conv_Conv2D): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1))
    (mixed_8_tower_1_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_1_conv): ReLU(inplace=True)
    (mixed_8_tower_1_conv_1_Conv2D): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3, 0))
    (mixed_8_tower_1_conv_1_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_1_conv_1): ReLU(inplace=True)
    (mixed_8_tower_1_conv_2_Conv2D): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0, 3))
    (mixed_8_tower_1_conv_2_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_1_conv_2): ReLU(inplace=True)
    (mixed_8_tower_1_conv_3_Conv2D): Conv2d(192, 192, kernel_size=(3, 3), stride=(2, 2))
    (mixed_8_tower_1_conv_3_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_8_tower_1_conv_3): ReLU(inplace=True)
    (mixed_8_pool): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1, ceil_mode=True)
    (mixed_9_conv_Conv2D): Conv2d(1280, 320, kernel_size=(1, 1), stride=(1, 1))
    (mixed_9_conv_batchnorm): BatchNorm2d(320, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_9_conv): ReLU(inplace=True)
    (mixed_9_tower_conv_Conv2D): Conv2d(1280, 384, kernel_size=(1, 1), stride=(1, 1))
    (mixed_9_tower_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_9_tower_conv): ReLU(inplace=True)
    (mixed_9_tower_mixed_conv_Conv2D): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1, 0))
    (mixed_9_tower_mixed_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_9_tower_mixed_conv): ReLU(inplace=True)
    (mixed_9_tower_mixed_conv_1_Conv2D): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0, 1))
    (mixed_9_tower_mixed_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_9_tower_mixed_conv_1): ReLU(inplace=True)
    (mixed_9_tower_1_conv_Conv2D): Conv2d(1280, 448, kernel_size=(1, 1), stride=(1, 1))
    (mixed_9_tower_1_conv_batchnorm): BatchNorm2d(448, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_9_tower_1_conv): ReLU(inplace=True)
    (mixed_9_tower_1_conv_1_Conv2D): Conv2d(448, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_9_tower_1_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_9_tower_1_conv_1): ReLU(inplace=True)
    (mixed_9_tower_1_mixed_conv_Conv2D): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1, 0))
    (mixed_9_tower_1_mixed_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_9_tower_1_mixed_conv): ReLU(inplace=True)
    (mixed_9_tower_1_mixed_conv_1_Conv2D): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0, 1))
    (mixed_9_tower_1_mixed_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_9_tower_1_mixed_conv_1): ReLU(inplace=True)
    (mixed_9_tower_2_pool): AvgPool2d(kernel_size=3, stride=1, padding=1)
    (mixed_9_tower_2_conv_Conv2D): Conv2d(1280, 192, kernel_size=(1, 1), stride=(1, 1))
    (mixed_9_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_9_tower_2_conv): ReLU(inplace=True)
    (mixed_10_conv_Conv2D): Conv2d(2048, 320, kernel_size=(1, 1), stride=(1, 1))
    (mixed_10_conv_batchnorm): BatchNorm2d(320, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_10_conv): ReLU(inplace=True)
    (mixed_10_tower_conv_Conv2D): Conv2d(2048, 384, kernel_size=(1, 1), stride=(1, 1))
    (mixed_10_tower_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_10_tower_conv): ReLU(inplace=True)
    (mixed_10_tower_mixed_conv_Conv2D): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1, 0))
    (mixed_10_tower_mixed_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_10_tower_mixed_conv): ReLU(inplace=True)
    (mixed_10_tower_mixed_conv_1_Conv2D): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0, 1))
    (mixed_10_tower_mixed_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_10_tower_mixed_conv_1): ReLU(inplace=True)
    (mixed_10_tower_1_conv_Conv2D): Conv2d(2048, 448, kernel_size=(1, 1), stride=(1, 1))
```

```
    (mixed_10_tower_1_conv_batchnorm): BatchNorm2d(448, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_10_tower_1_conv): ReLU(inplace=True)
    (mixed_10_tower_1_conv_1_Conv2D): Conv2d(448, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (mixed_10_tower_1_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_10_tower_1_conv_1): ReLU(inplace=True)
    (mixed_10_tower_1_mixed_conv_Conv2D): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1, 0))
    (mixed_10_tower_1_mixed_conv_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_10_tower_1_mixed_conv): ReLU(inplace=True)
    (mixed_10_tower_1_mixed_conv_1_Conv2D): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0, 1))
    (mixed_10_tower_1_mixed_conv_1_batchnorm): BatchNorm2d(384, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (mixed_10_tower_1_mixed_conv_1): ReLU(inplace=True)
    (mixed_10_tower_2_pool): MaxPool2d(kernel_size=3, stride=1, padding=1, dilation=1, ceil_mode=True)
    (mixed_10_tower_2_conv_Conv2D): Conv2d(2048, 192, kernel_size=(1, 1), stride=(1, 1))
    (mixed_10_tower_2_conv_batchnorm): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (mixed_10_tower_2_conv): ReLU(inplace=True)
    (top_cls_pool): AvgPool2d(kernel_size=8, stride=1, padding=0)
    (top_cls_fc): Dropout(p=0.8, inplace=False)
  )
  (new_fc): Linear(in_features=2048, out_features=256, bias=True)
  (consensus): RelationModuleMultiScale(
    (fc_fusion_scales): ModuleList(
      (0): Sequential(
        (0): ReLU()
        (1): Linear(in_features=2048, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (1): Sequential(
        (0): ReLU()
        (1): Linear(in_features=1792, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (2): Sequential(
        (0): ReLU()
        (1): Linear(in_features=1536, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (3): Sequential(
        (0): ReLU()
        (1): Linear(in_features=1280, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (4): Sequential(
        (0): ReLU()
        (1): Linear(in_features=1024, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (5): Sequential(
        (0): ReLU()
        (1): Linear(in_features=768, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
      (6): Sequential(
        (0): ReLU()
        (1): Linear(in_features=512, out_features=256, bias=True)
        (2): ReLU()
        (3): Linear(in_features=256, out_features=339, bias=True)
      )
    )
  )
)
```