Video frames
16 frames (224x224)

7x7x7 conv, *64*, stride *2,2,2*

batch normalization

ReLU

1x3x3 max-pool, stride *1,2,2*

1x1x1 conv, *64*, stride *1,1,1*

batch normalization

ReLU

3x3x3 conv, *192*, stride *1,1,1*

batch normalization

ReLU

1x3x3 max-pool, stride *1,2,2*

inc. [64, 96, 128, 16, 32, 32]

inc. [128, 128, 192, 32, 96, 64]

3x3x3 max-pool, stride *2,2,2*

inc. [192, 96, 208, 16, 48, 64]

inc. [160, 112, 224, 24, 64, 64]

inc. [128, 128, 256, 24, 64, 64]

inc. [112, 144, 288, 32, 64, 64]

inc. [256, 160, 320, 32, 128, 128]

2x2x2 max-pool, stride *2,2,2*

inc. [256, 160, 320, 32, 128, 128]

inc. [384, 192, 384, 48, 128, 128]

2x7x7 avg-pool, stride *1,1,1*

1x1x1 conv, *400*, stride *1,1,1*

Predictions

Inflated Inception-V1 Model