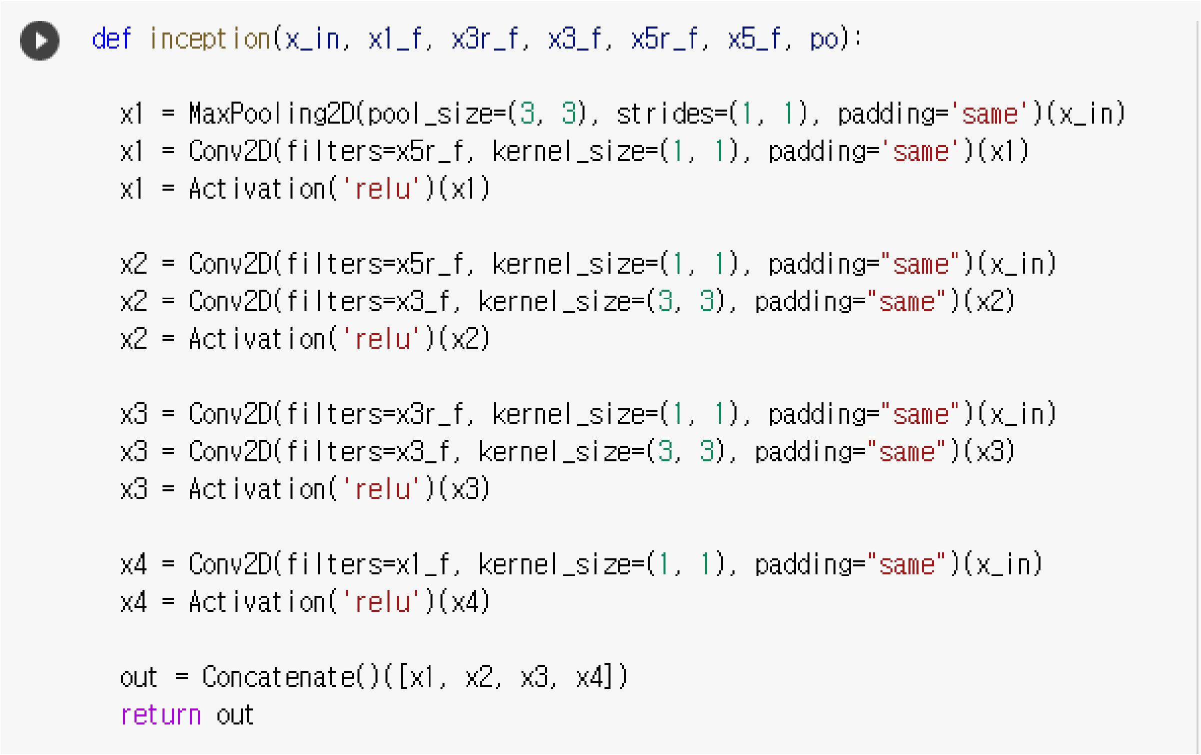
이름 : 강시온

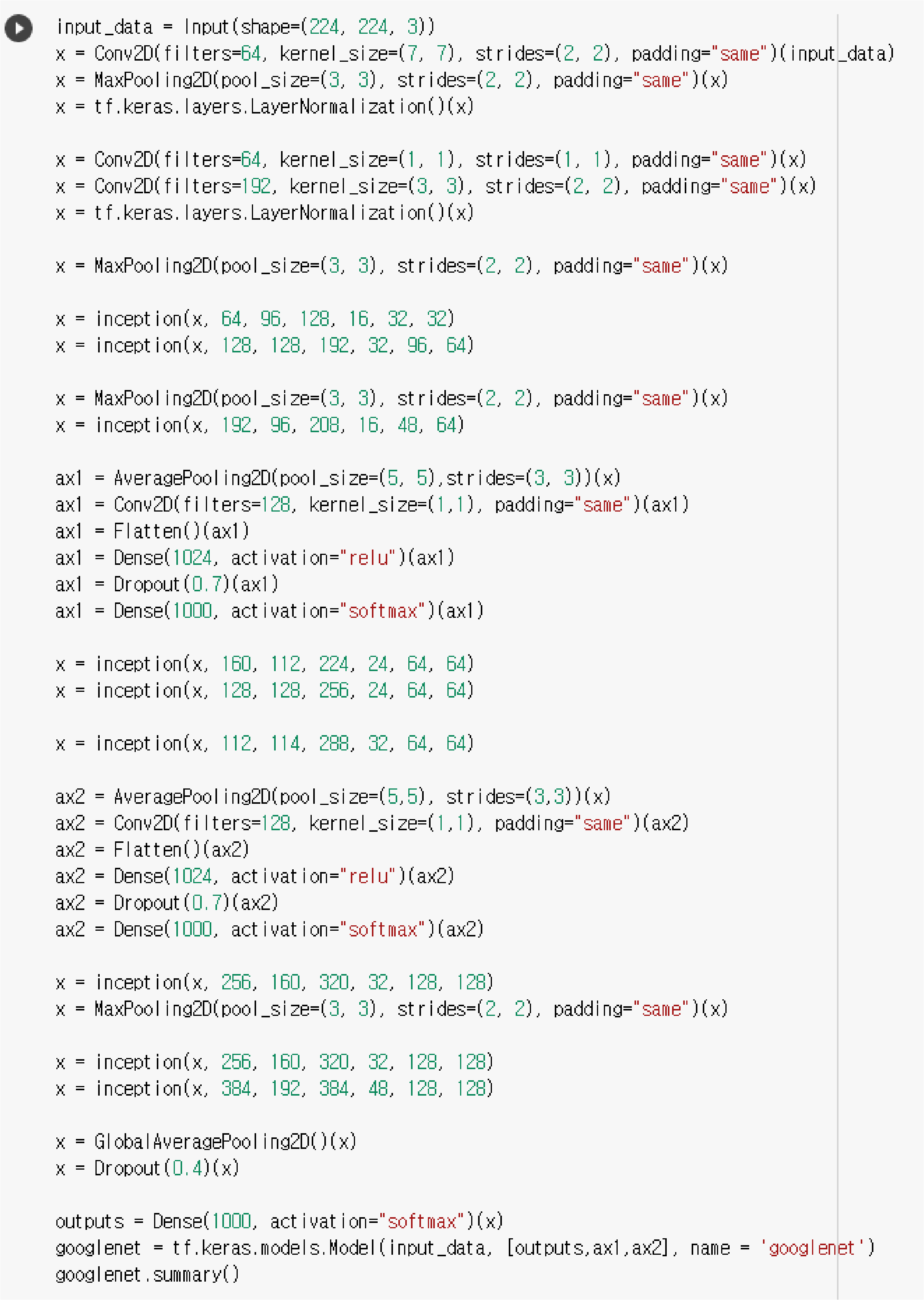
학번 : 201802045

학과 : 컴퓨터공학과

구현 코드



```python
def inception(x_in, x1_f, x3r_f, x3_f, x5r_f, x5_f, po):

    x1 = MaxPooling2D(pool_size=(3, 3), strides=(1, 1), padding='same')(x_in)
    x1 = Conv2D(filters=x5r_f, kernel_size=(1, 1), padding='same')(x1)
    x1 = Activation('relu')(x1)

    x2 = Conv2D(filters=x5r_f, kernel_size=(1, 1), padding="same")(x_in)
    x2 = Conv2D(filters=x3_f, kernel_size=(3, 3), padding="same")(x2)
    x2 = Activation('relu')(x2)

    x3 = Conv2D(filters=x3r_f, kernel_size=(1, 1), padding="same")(x_in)
    x3 = Conv2D(filters=x3_f, kernel_size=(3, 3), padding="same")(x3)
    x3 = Activation('relu')(x3)

    x4 = Conv2D(filters=x1_f, kernel_size=(1, 1), padding="same")(x_in)
    x4 = Activation('relu')(x4)

    out = Concatenate()([x1, x2, x3, x4])
    return out
```



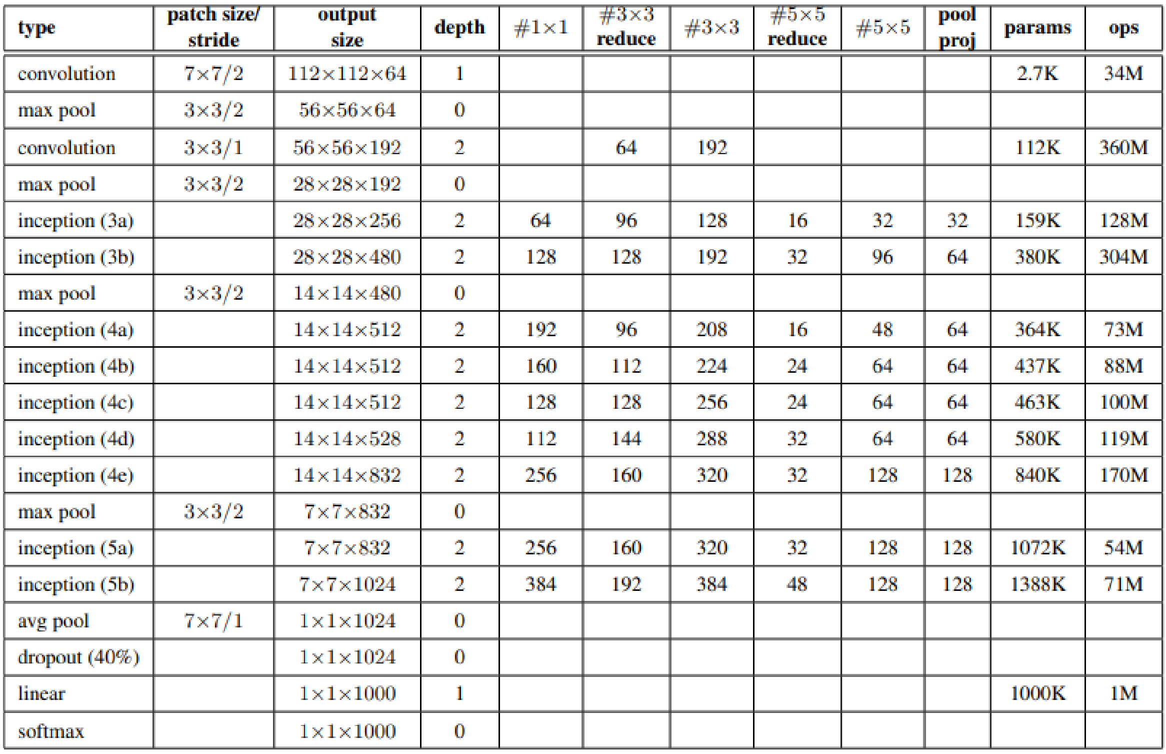
```
input_data = Input(shape=(224, 224, 3))
x = Conv2D(filters=64, kernel_size=(7, 7), strides=(2, 2), padding="same")(input_data)
x = MaxPooling2D(pool_size=(3, 3), strides=(2, 2), padding="same")(x)
x = tf.keras.layers.LayerNormalization()(x)

x = Conv2D(filters=64, kernel_size=(1, 1), strides=(1, 1), padding="same")(x)
x = Conv2D(filters=192, kernel_size=(3, 3), strides=(2, 2), padding="same")(x)
x = tf.keras.layers.LayerNormalization()(x)

x = MaxPooling2D(pool_size=(3, 3), strides=(2, 2), padding="same")(x)

x = inception(x, 64, 96, 128, 16, 32, 32)
x = inception(x, 128, 128, 192, 32, 96, 64)

x = MaxPooling2D(pool_size=(3, 3), strides=(2, 2), padding="same")(x)
x = inception(x, 192, 96, 208, 16, 48, 64)

ax1 = AveragePooling2D(pool_size=(5, 5),strides=(3, 3))(x)
ax1 = Conv2D(filters=128, kernel_size=(1,1), padding="same")(ax1)
ax1 = Flatten()(ax1)
ax1 = Dense(1024, activation="relu")(ax1)
ax1 = Dropout(0.7)(ax1)
ax1 = Dense(1000, activation="softmax")(ax1)

x = inception(x, 160, 112, 224, 24, 64, 64)
x = inception(x, 128, 128, 256, 24, 64, 64)

x = inception(x, 112, 114, 288, 32, 64, 64)

ax2 = AveragePooling2D(pool_size=(5,5), strides=(3,3))(x)
ax2 = Conv2D(filters=128, kernel_size=(1,1), padding="same")(ax2)
ax2 = Flatten()(ax2)
ax2 = Dense(1024, activation="relu")(ax2)
ax2 = Dropout(0.7)(ax2)
ax2 = Dense(1000, activation="softmax")(ax2)

x = inception(x, 256, 160, 320, 32, 128, 128)
x = MaxPooling2D(pool_size=(3, 3), strides=(2, 2), padding="same")(x)

x = inception(x, 256, 160, 320, 32, 128, 128)
x = inception(x, 384, 192, 384, 48, 128, 128)

x = GlobalAveragePooling2D()(x)
x = Dropout(0.4)(x)

outputs = Dense(1000, activation="softmax")(x)
googlenet = tf.keras.models.Model(input_data, [outputs,ax1,ax2], name = 'googlenet')
googlenet.summary()
```

코드 설명

먼저 공통적으로 반복되는 부분인 inception module을 함수로 만들어 분리하였다.

X_in 1개를 4개의 layer로 분기하고 마지막에 concatenate함수를 통해 하나로 합쳐 out으로 만들어줬다.

본 함수는 (224, 224, 3) shape를 가지는 input으로 시작하여

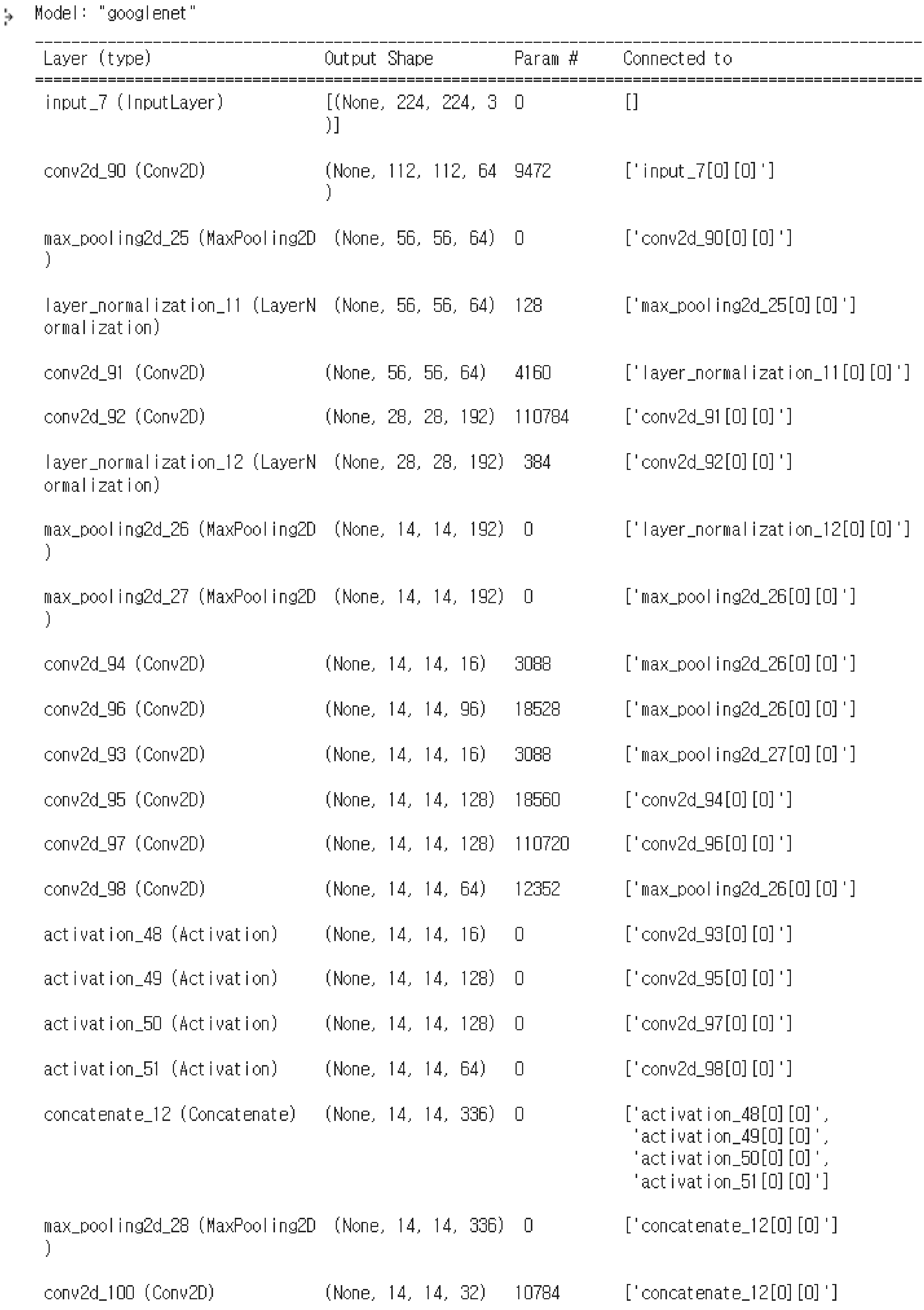


| type | patch size/ stride | output size | depth | #1×1 | #3×3 reduce | #3×3 | #5×5 reduce | #5×5 | pool proj | params | ops |
|---|---|---|---|---|---|---|---|---|---|---|---|
| convolution | 7×7/2 | 112×112×64 | 1 | | | | | | | 2.7K | 34M |
| max pool | 3×3/2 | 56×56×64 | 0 | | | | | | | | |
| convolution | 3×3/1 | 56×56×192 | 2 | | 64 | 192 | | | | 112K | 360M |
| max pool | 3×3/2 | 28×28×192 | 0 | | | | | | | | |
| inception (3a) | | 28×28×256 | 2 | 64 | 96 | 128 | 16 | 32 | 32 | 159K | 128M |
| inception (3b) | | 28×28×480 | 2 | 128 | 128 | 192 | 32 | 96 | 64 | 380K | 304M |
| max pool | 3×3/2 | 14×14×480 | 0 | | | | | | | | |
| inception (4a) | | 14×14×512 | 2 | 192 | 96 | 208 | 16 | 48 | 64 | 364K | 73M |
| inception (4b) | | 14×14×512 | 2 | 160 | 112 | 224 | 24 | 64 | 64 | 437K | 88M |
| inception (4c) | | 14×14×512 | 2 | 128 | 128 | 256 | 24 | 64 | 64 | 463K | 100M |
| inception (4d) | | 14×14×528 | 2 | 112 | 144 | 288 | 32 | 64 | 64 | 580K | 119M |
| inception (4e) | | 14×14×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 840K | 170M |
| max pool | 3×3/2 | 7×7×832 | 0 | | | | | | | | |
| inception (5a) | | 7×7×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 1072K | 54M |
| inception (5b) | | 7×7×1024 | 2 | 384 | 192 | 384 | 48 | 128 | 128 | 1388K | 71M |
| avg pool | 7×7/1 | 1×1×1024 | 0 | | | | | | | | |
| dropout (40%) | | 1×1×1024 | 0 | | | | | | | | |
| linear | | 1×1×1000 | 1 | | | | | | | 1000K | 1M |
| softmax | | 1×1×1000 | 0 | | | | | | | | |

해당 표에 맞는 인자들과 레이어층을 넣어 모델을 완성하였다.

Inception 함수의 인자는 해당 테이블의 순서대로 작성해주었다.

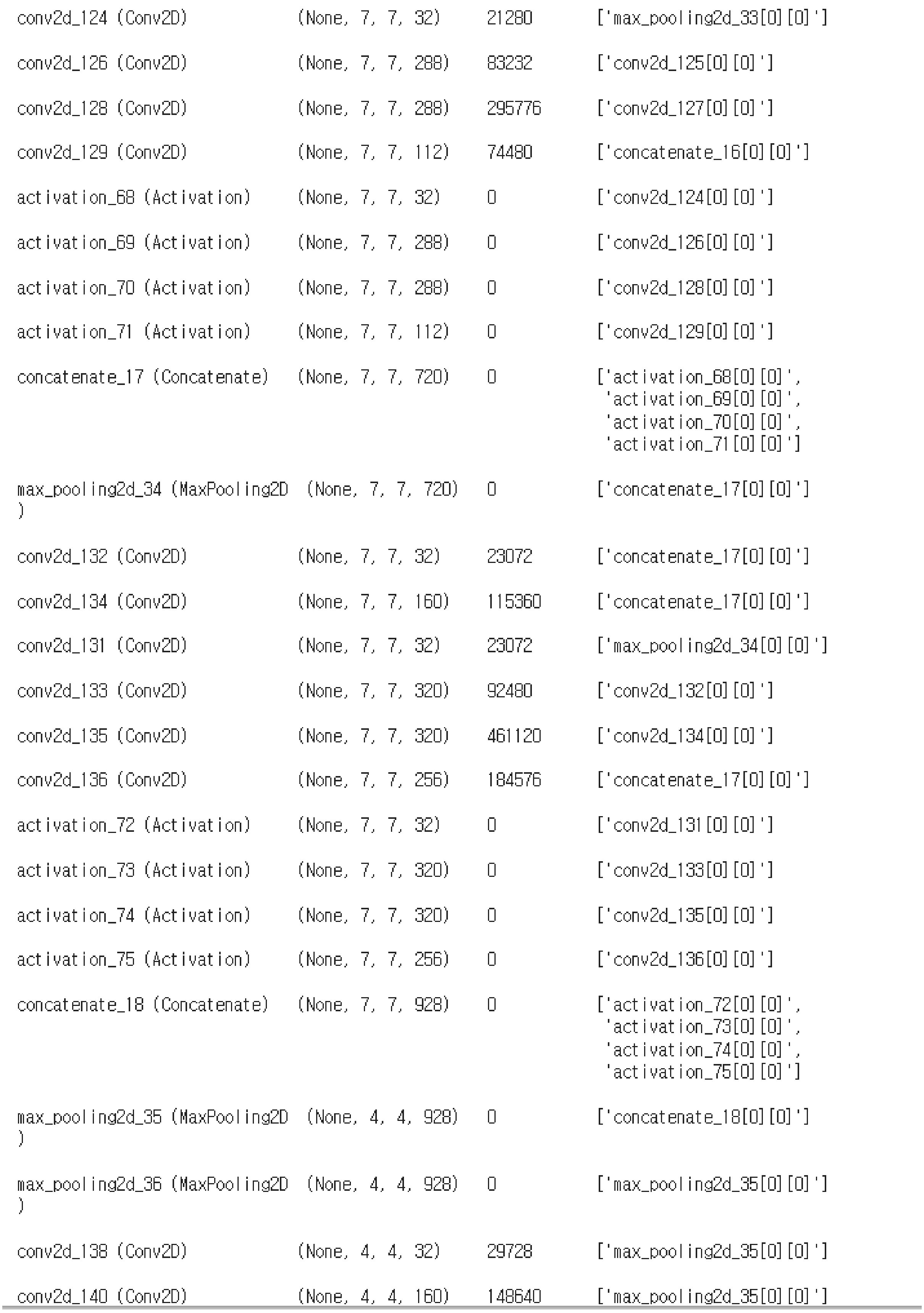
결과 이미지



```
Model: "googlenet"
--------------------------------------------------------------------------------------------------
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_7 (InputLayer)           [(None, 224, 224, 3  0           []
                                )]

 conv2d_90 (Conv2D)             (None, 112, 112, 64  9472        ['input_7[0][0]']
                                )

 max_pooling2d_25 (MaxPooling2D  (None, 56, 56, 64)  0           ['conv2d_90[0][0]']
 )

 layer_normalization_11 (LayerN  (None, 56, 56, 64)  128         ['max_pooling2d_25[0][0]']
 ormalization)

 conv2d_91 (Conv2D)             (None, 56, 56, 64)   4160        ['layer_normalization_11[0][0]']

 conv2d_92 (Conv2D)             (None, 28, 28, 192)  110784      ['conv2d_91[0][0]']

 layer_normalization_12 (LayerN  (None, 28, 28, 192)  384        ['conv2d_92[0][0]']
 ormalization)

 max_pooling2d_26 (MaxPooling2D  (None, 14, 14, 192)  0          ['layer_normalization_12[0][0]']
 )

 max_pooling2d_27 (MaxPooling2D  (None, 14, 14, 192)  0          ['max_pooling2d_26[0][0]']
 )

 conv2d_94 (Conv2D)             (None, 14, 14, 16)   3088        ['max_pooling2d_26[0][0]']

 conv2d_96 (Conv2D)             (None, 14, 14, 96)   18528       ['max_pooling2d_26[0][0]']

 conv2d_93 (Conv2D)             (None, 14, 14, 16)   3088        ['max_pooling2d_27[0][0]']

 conv2d_95 (Conv2D)             (None, 14, 14, 128)  18560       ['conv2d_94[0][0]']

 conv2d_97 (Conv2D)             (None, 14, 14, 128)  110720      ['conv2d_96[0][0]']

 conv2d_98 (Conv2D)             (None, 14, 14, 64)   12352       ['max_pooling2d_26[0][0]']

 activation_48 (Activation)     (None, 14, 14, 16)   0           ['conv2d_93[0][0]']

 activation_49 (Activation)     (None, 14, 14, 128)  0           ['conv2d_95[0][0]']

 activation_50 (Activation)     (None, 14, 14, 128)  0           ['conv2d_97[0][0]']

 activation_51 (Activation)     (None, 14, 14, 64)   0           ['conv2d_98[0][0]']

 concatenate_12 (Concatenate)   (None, 14, 14, 336)  0           ['activation_48[0][0]',
                                                                  'activation_49[0][0]',
                                                                  'activation_50[0][0]',
                                                                  'activation_51[0][0]']

 max_pooling2d_28 (MaxPooling2D  (None, 14, 14, 336)  0          ['concatenate_12[0][0]']
 )

 conv2d_100 (Conv2D)            (None, 14, 14, 32)   10784       ['concatenate_12[0][0]']
```

```
conv2d_102 (Conv2D)            (None, 14, 14, 128)  43136       ['concatenate_12[0][0]']

conv2d_99 (Conv2D)             (None, 14, 14, 32)   10784       ['max_pooling2d_28[0][0]']

conv2d_101 (Conv2D)            (None, 14, 14, 192)  55488       ['conv2d_100[0][0]']

conv2d_103 (Conv2D)            (None, 14, 14, 192)  221376      ['conv2d_102[0][0]']

conv2d_104 (Conv2D)            (None, 14, 14, 128)  43136       ['concatenate_12[0][0]']

activation_52 (Activation)     (None, 14, 14, 32)   0           ['conv2d_99[0][0]']

activation_53 (Activation)     (None, 14, 14, 192)  0           ['conv2d_101[0][0]']

activation_54 (Activation)     (None, 14, 14, 192)  0           ['conv2d_103[0][0]']

activation_55 (Activation)     (None, 14, 14, 128)  0           ['conv2d_104[0][0]']

concatenate_13 (Concatenate)   (None, 14, 14, 544)  0           ['activation_52[0][0]',
                                                                 'activation_53[0][0]',
                                                                 'activation_54[0][0]',
                                                                 'activation_55[0][0]']

max_pooling2d_29 (MaxPooling2D  (None, 7, 7, 544)   0           ['concatenate_13[0][0]']
)

max_pooling2d_30 (MaxPooling2D  (None, 7, 7, 544)   0           ['max_pooling2d_29[0][0]']
)

conv2d_106 (Conv2D)            (None, 7, 7, 16)     8720        ['max_pooling2d_29[0][0]']

conv2d_108 (Conv2D)            (None, 7, 7, 96)     52320       ['max_pooling2d_29[0][0]']

conv2d_105 (Conv2D)            (None, 7, 7, 16)     8720        ['max_pooling2d_30[0][0]']

conv2d_107 (Conv2D)            (None, 7, 7, 208)    30160       ['conv2d_106[0][0]']

conv2d_109 (Conv2D)            (None, 7, 7, 208)    179920      ['conv2d_108[0][0]']

conv2d_110 (Conv2D)            (None, 7, 7, 192)    104640      ['max_pooling2d_29[0][0]']

activation_56 (Activation)     (None, 7, 7, 16)     0           ['conv2d_105[0][0]']

activation_57 (Activation)     (None, 7, 7, 208)    0           ['conv2d_107[0][0]']

activation_58 (Activation)     (None, 7, 7, 208)    0           ['conv2d_109[0][0]']

activation_59 (Activation)     (None, 7, 7, 192)    0           ['conv2d_110[0][0]']

concatenate_14 (Concatenate)   (None, 7, 7, 624)    0           ['activation_56[0][0]',
                                                                 'activation_57[0][0]',
                                                                 'activation_58[0][0]',
                                                                 'activation_59[0][0]']

max_pooling2d_31 (MaxPooling2D  (None, 7, 7, 624)   0           ['concatenate_14[0][0]']
)
```

```
 conv2d_115 (Conv2D)            (None, 7, 7, 112)    70000       ['concatenate_14[0][0]']

 conv2d_112 (Conv2D)            (None, 7, 7, 24)     15000       ['max_pooling2d_31[0][0]']

 conv2d_114 (Conv2D)            (None, 7, 7, 224)    48608       ['conv2d_113[0][0]']

 conv2d_116 (Conv2D)            (None, 7, 7, 224)    226016      ['conv2d_115[0][0]']

 conv2d_117 (Conv2D)            (None, 7, 7, 160)    100000      ['concatenate_14[0][0]']

 activation_60 (Activation)     (None, 7, 7, 24)     0           ['conv2d_112[0][0]']

 activation_61 (Activation)     (None, 7, 7, 224)    0           ['conv2d_114[0][0]']

 activation_62 (Activation)     (None, 7, 7, 224)    0           ['conv2d_116[0][0]']

 activation_63 (Activation)     (None, 7, 7, 160)    0           ['conv2d_117[0][0]']

 concatenate_15 (Concatenate)   (None, 7, 7, 632)    0           ['activation_60[0][0]',
                                                                  'activation_61[0][0]',
                                                                  'activation_62[0][0]',
                                                                  'activation_63[0][0]']

 max_pooling2d_32 (MaxPooling2D  (None, 7, 7, 632)   0           ['concatenate_15[0][0]']
 )

 conv2d_119 (Conv2D)            (None, 7, 7, 24)     15192       ['concatenate_15[0][0]']

 conv2d_121 (Conv2D)            (None, 7, 7, 128)    81024       ['concatenate_15[0][0]']

 conv2d_118 (Conv2D)            (None, 7, 7, 24)     15192       ['max_pooling2d_32[0][0]']

 conv2d_120 (Conv2D)            (None, 7, 7, 256)    55552       ['conv2d_119[0][0]']

 conv2d_122 (Conv2D)            (None, 7, 7, 256)    295168      ['conv2d_121[0][0]']

 conv2d_123 (Conv2D)            (None, 7, 7, 128)    81024       ['concatenate_15[0][0]']

 activation_64 (Activation)     (None, 7, 7, 24)     0           ['conv2d_118[0][0]']

 activation_65 (Activation)     (None, 7, 7, 256)    0           ['conv2d_120[0][0]']

 activation_66 (Activation)     (None, 7, 7, 256)    0           ['conv2d_122[0][0]']

 activation_67 (Activation)     (None, 7, 7, 128)    0           ['conv2d_123[0][0]']

 concatenate_16 (Concatenate)   (None, 7, 7, 664)    0           ['activation_64[0][0]',
                                                                  'activation_65[0][0]',
                                                                  'activation_66[0][0]',
                                                                  'activation_67[0][0]']

 max_pooling2d_33 (MaxPooling2D  (None, 7, 7, 664)   0           ['concatenate_16[0][0]']
 )

 conv2d_125 (Conv2D)            (None, 7, 7, 32)     21280       ['concatenate_16[0][0]']

 conv2d_127 (Conv2D)            (None, 7, 7, 114)    75810       ['concatenate_16[0][0]']
```



```
conv2d_124 (Conv2D)            (None, 7, 7, 32)     21280       ['max_pooling2d_33[0][0]']

conv2d_126 (Conv2D)            (None, 7, 7, 288)    83232       ['conv2d_125[0][0]']

conv2d_128 (Conv2D)            (None, 7, 7, 288)    295776      ['conv2d_127[0][0]']

conv2d_129 (Conv2D)            (None, 7, 7, 112)    74480       ['concatenate_16[0][0]']

activation_68 (Activation)     (None, 7, 7, 32)     0           ['conv2d_124[0][0]']

activation_69 (Activation)     (None, 7, 7, 288)    0           ['conv2d_126[0][0]']

activation_70 (Activation)     (None, 7, 7, 288)    0           ['conv2d_128[0][0]']

activation_71 (Activation)     (None, 7, 7, 112)    0           ['conv2d_129[0][0]']

concatenate_17 (Concatenate)   (None, 7, 7, 720)    0           ['activation_68[0][0]',
                                                                 'activation_69[0][0]',
                                                                 'activation_70[0][0]',
                                                                 'activation_71[0][0]']

max_pooling2d_34 (MaxPooling2D  (None, 7, 7, 720)   0           ['concatenate_17[0][0]']
)

conv2d_132 (Conv2D)            (None, 7, 7, 32)     23072       ['concatenate_17[0][0]']

conv2d_134 (Conv2D)            (None, 7, 7, 160)    115360      ['concatenate_17[0][0]']

conv2d_131 (Conv2D)            (None, 7, 7, 32)     23072       ['max_pooling2d_34[0][0]']

conv2d_133 (Conv2D)            (None, 7, 7, 320)    92480       ['conv2d_132[0][0]']

conv2d_135 (Conv2D)            (None, 7, 7, 320)    461120      ['conv2d_134[0][0]']

conv2d_136 (Conv2D)            (None, 7, 7, 256)    184576      ['concatenate_17[0][0]']

activation_72 (Activation)     (None, 7, 7, 32)     0           ['conv2d_131[0][0]']

activation_73 (Activation)     (None, 7, 7, 320)    0           ['conv2d_133[0][0]']

activation_74 (Activation)     (None, 7, 7, 320)    0           ['conv2d_135[0][0]']

activation_75 (Activation)     (None, 7, 7, 256)    0           ['conv2d_136[0][0]']

concatenate_18 (Concatenate)   (None, 7, 7, 928)    0           ['activation_72[0][0]',
                                                                 'activation_73[0][0]',
                                                                 'activation_74[0][0]',
                                                                 'activation_75[0][0]']

max_pooling2d_35 (MaxPooling2D  (None, 4, 4, 928)   0           ['concatenate_18[0][0]']
)

max_pooling2d_36 (MaxPooling2D  (None, 4, 4, 928)   0           ['max_pooling2d_35[0][0]']
)

conv2d_138 (Conv2D)            (None, 4, 4, 32)     29728       ['max_pooling2d_35[0][0]']

conv2d_140 (Conv2D)            (None, 4, 4, 160)    148640      ['max_pooling2d_35[0][0]']
```

```
conv2d_137 (Conv2D)            (None, 4, 4, 32)     29728       ['max_pooling2d_36[0][0]']

conv2d_139 (Conv2D)            (None, 4, 4, 320)    92480       ['conv2d_138[0][0]']

conv2d_141 (Conv2D)            (None, 4, 4, 320)    461120      ['conv2d_140[0][0]']

conv2d_142 (Conv2D)            (None, 4, 4, 256)    237824      ['max_pooling2d_35[0][0]']

activation_76 (Activation)     (None, 4, 4, 32)     0           ['conv2d_137[0][0]']

activation_77 (Activation)     (None, 4, 4, 320)    0           ['conv2d_139[0][0]']

activation_78 (Activation)     (None, 4, 4, 320)    0           ['conv2d_141[0][0]']

activation_79 (Activation)     (None, 4, 4, 256)    0           ['conv2d_142[0][0]']

concatenate_19 (Concatenate)   (None, 4, 4, 928)    0           ['activation_76[0][0]',
                                                                 'activation_77[0][0]',
                                                                 'activation_78[0][0]',
                                                                 'activation_79[0][0]']

max_pooling2d_37 (MaxPooling2D  (None, 4, 4, 928)   0           ['concatenate_19[0][0]']
)

conv2d_144 (Conv2D)            (None, 4, 4, 48)     44592       ['concatenate_19[0][0]']

conv2d_146 (Conv2D)            (None, 4, 4, 192)    178368      ['concatenate_19[0][0]']

conv2d_143 (Conv2D)            (None, 4, 4, 48)     44592       ['max_pooling2d_37[0][0]']

conv2d_145 (Conv2D)            (None, 4, 4, 384)    166272      ['conv2d_144[0][0]']

conv2d_147 (Conv2D)            (None, 4, 4, 384)    663936      ['conv2d_146[0][0]']

conv2d_148 (Conv2D)            (None, 4, 4, 384)    356736      ['concatenate_19[0][0]']

average_pooling2d_2 (AveragePo  (None, 1, 1, 624)   0           ['concatenate_14[0][0]']
oling2D)

average_pooling2d_3 (AveragePo  (None, 1, 1, 720)   0           ['concatenate_17[0][0]']
oling2D)

activation_80 (Activation)     (None, 4, 4, 48)     0           ['conv2d_143[0][0]']

activation_81 (Activation)     (None, 4, 4, 384)    0           ['conv2d_145[0][0]']

activation_82 (Activation)     (None, 4, 4, 384)    0           ['conv2d_147[0][0]']

activation_83 (Activation)     (None, 4, 4, 384)    0           ['conv2d_148[0][0]']

conv2d_111 (Conv2D)            (None, 1, 1, 128)    80000       ['average_pooling2d_2[0][0]']

conv2d_130 (Conv2D)            (None, 1, 1, 128)    92288       ['average_pooling2d_3[0][0]']

concatenate_20 (Concatenate)   (None, 4, 4, 1200)   0           ['activation_80[0][0]',
                                                                 'activation_81[0][0]',
                                                                 'activation_82[0][0]',
                                                                 'activation_83[0][0]']
```

```
 flatten_2 (Flatten)            (None, 128)          0           ['conv2d_111[0][0]']

 flatten_3 (Flatten)            (None, 128)          0           ['conv2d_130[0][0]']

 global_average_pooling2d_1 (Gl  (None, 1200)        0           ['concatenate_20[0][0]']
 obalAveragePooling2D)

 dense_5 (Dense)                (None, 1024)         132096      ['flatten_2[0][0]']

 dense_7 (Dense)                (None, 1024)         132096      ['flatten_3[0][0]']

 dropout_5 (Dropout)            (None, 1200)         0           ['global_average_pooling2d_1[0][0
                                                                 ]']

 dropout_3 (Dropout)            (None, 1024)         0           ['dense_5[0][0]']

 dropout_4 (Dropout)            (None, 1024)         0           ['dense_7[0][0]']

 dense_9 (Dense)                (None, 1000)         1201000     ['dropout_5[0][0]']

 dense_6 (Dense)                (None, 1000)         1025000     ['dropout_3[0][0]']

 dense_8 (Dense)                (None, 1000)         1025000     ['dropout_4[0][0]']

==================================================================================================
Total params: 9,691,258
Trainable params: 9,691,258
Non-trainable params: 0
__________________________________________________________________________________________________
```

난이도

복잡한 모델을 구현한다고 하였는데 이정도로 복잡한 모델일줄은 몰라서 모델의 그림을 보고 처음에 굉장히 당황했다 하지만 반복되는 inception module을 함수로 작성해보니 또 그렇게 엄청나게 길진 않아 금방 구현할 수 있었다.