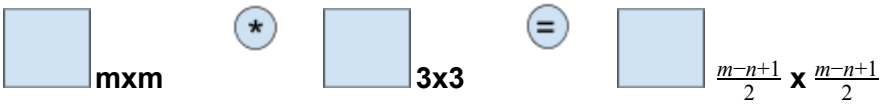
XCEPTION ARCHITECTURE- Documentation

The layers in Xception Architecture are as follows:

Convolution 2D with 3x3 filter with stride-2 :

Shape -



Params -

For n filters , The no.of parameters are nx3x3xdepth

Example: 32 filters of 3x3 Conv 2d for RGB image uses 32x3x3x3 = 864 params

Batch Normalization:

Shape - Remains the same

Params-

The Batch Normalization uses four parameters,

Mean - μ

Standard Deviation - σ

Normalization - z_{norm}($\frac{X-\mu}{\sigma}$)

Scaling - z'_{norm}($\gamma z_{norm}+\beta$)

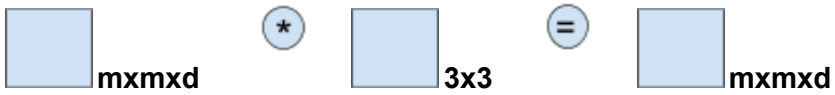
For depth N, The no.of parameters used are Nx4

Example: For depth 3, The no. of params = 3x4 =12

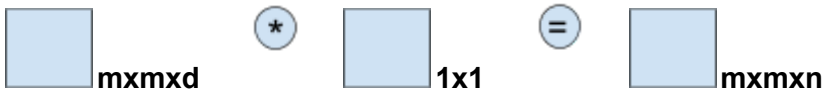
Separable Convolution:

Shape-

Depth-wise Convolution with 3x3 filter-



Point-wise Convolution with n filters-



Params-

Depth-wise Convolution with 3x3 filter-

For depth d, the no.of parameters used are 3x3xd

Example: for depth 64 , the no.of parameters is 3x3x64 = 576

Point-wise Convolution with n filters-

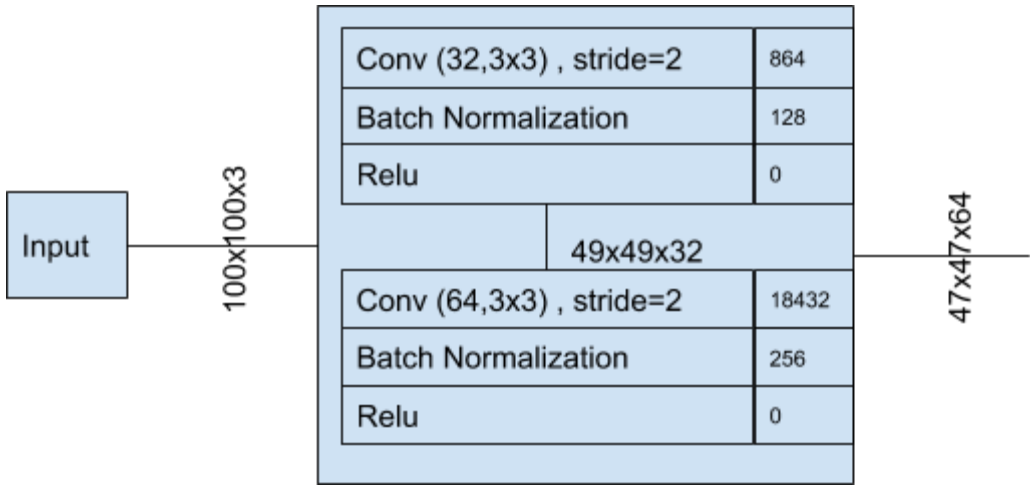
For n filters , the no.of parameters used are nxd

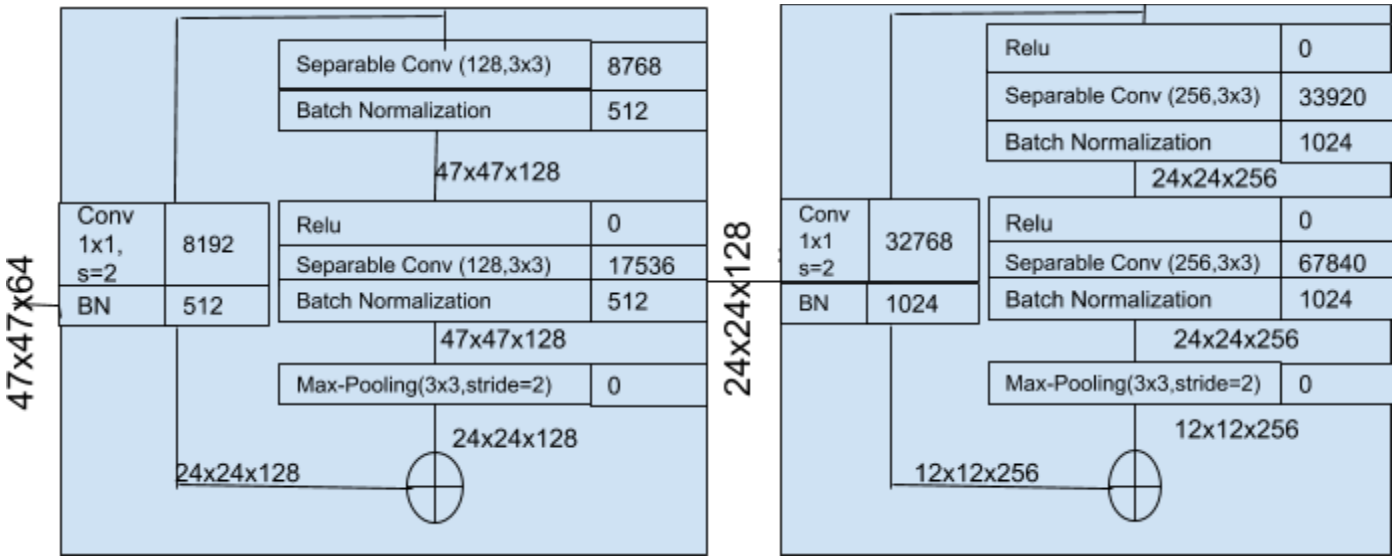
Example: For 128 filters, the no.of parameters are 128x64 = 8192

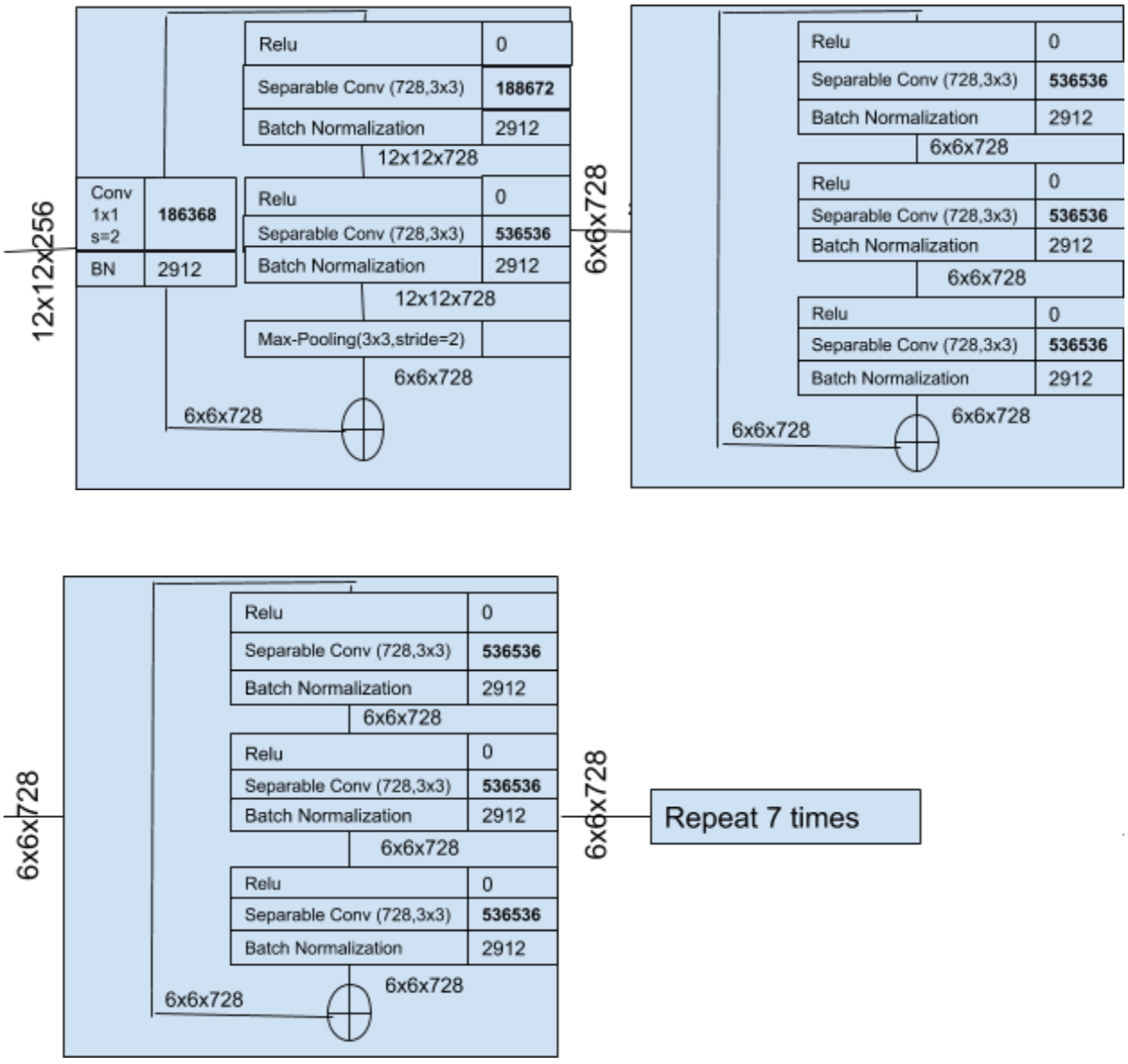
Total no.of parameters = Depth-wise + Point-wise

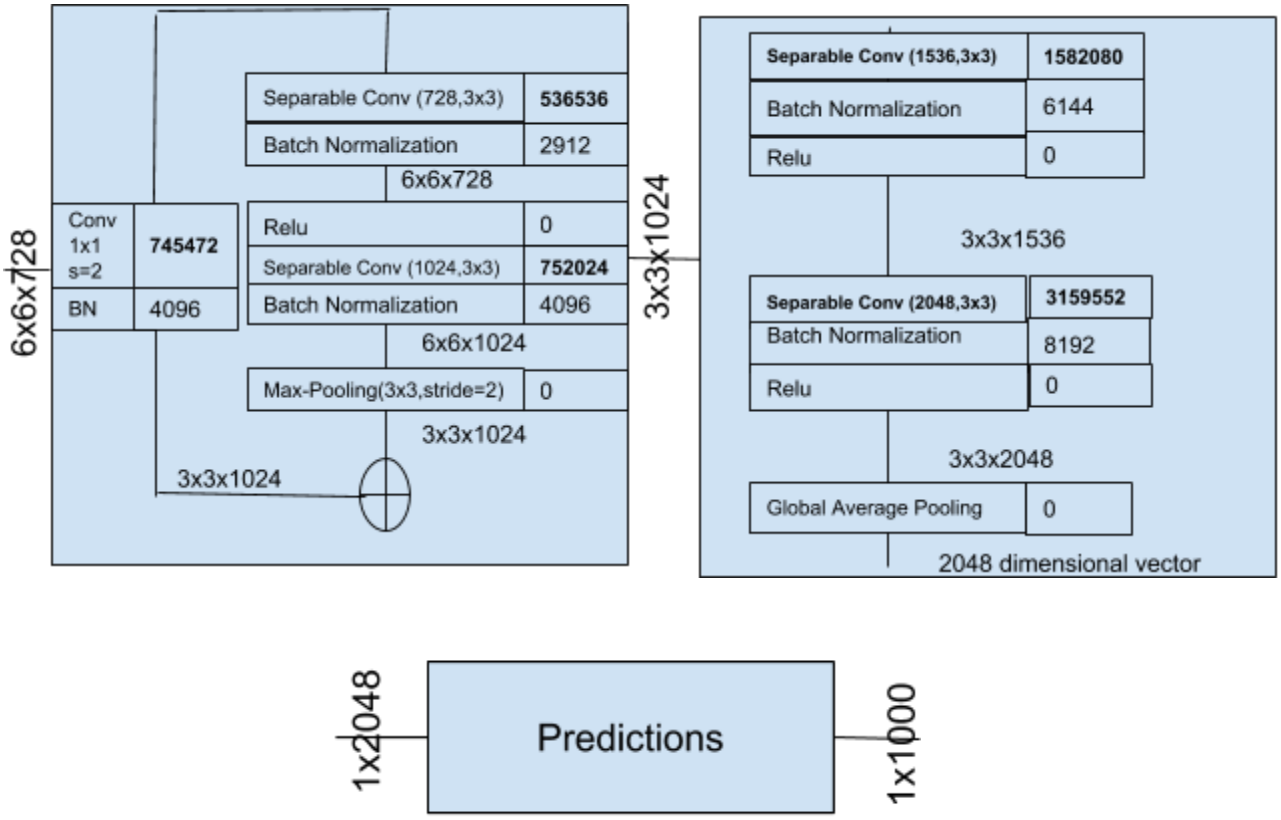
Example: 576+8192 = 8768

Xception Architecture:






12x12x256
Conv 1x1 s=2
186368
BN
2912
Relu
0
Separable Conv (728,3x3)
188672
Batch Normalization
2912
12x12x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
12x12x728
Max-Pooling(3x3,stride=2)
6x6x728
6x6x728
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
6x6x728
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
Relu
0
Separable Conv (728,3x3)
536536
Batch Normalization
2912
6x6x728
6x6x728
6x6x728
Repeat 7 times



The predictions here can be used for 1000 classes. It can be changed according to the application.