Deep Learning Architectures

Mohsen Kiskani

April 14, 2019

1 Description

1.1 Basic Formulas

$$\mathrm{Conv2D}(C_i, C_o, K, S = 1, P = 0, d = 1) \tag{1}$$
$$\mathrm{MaxPool2D}(K, S = None, P = 0, d = 1) \tag{2}$$
$$\text{MaxPool2d and Conv2D} \rightarrow L_{out} = \lfloor \frac{L_{in} + 2P - d(K-1) - 1}{S} + 1 \rfloor \tag{3}$$
$$\mathrm{Conv2D}(C_i, C_o, 1) \rightarrow L_{out} = L_{in} \tag{4}$$
$$\mathrm{Conv2D}(C_i, C_o, 3, 1, 1) \rightarrow L_{out} = L_{in} \tag{5}$$
$$\text{Tensor representation} \rightarrow N \times C \times H \times W \tag{6}$$

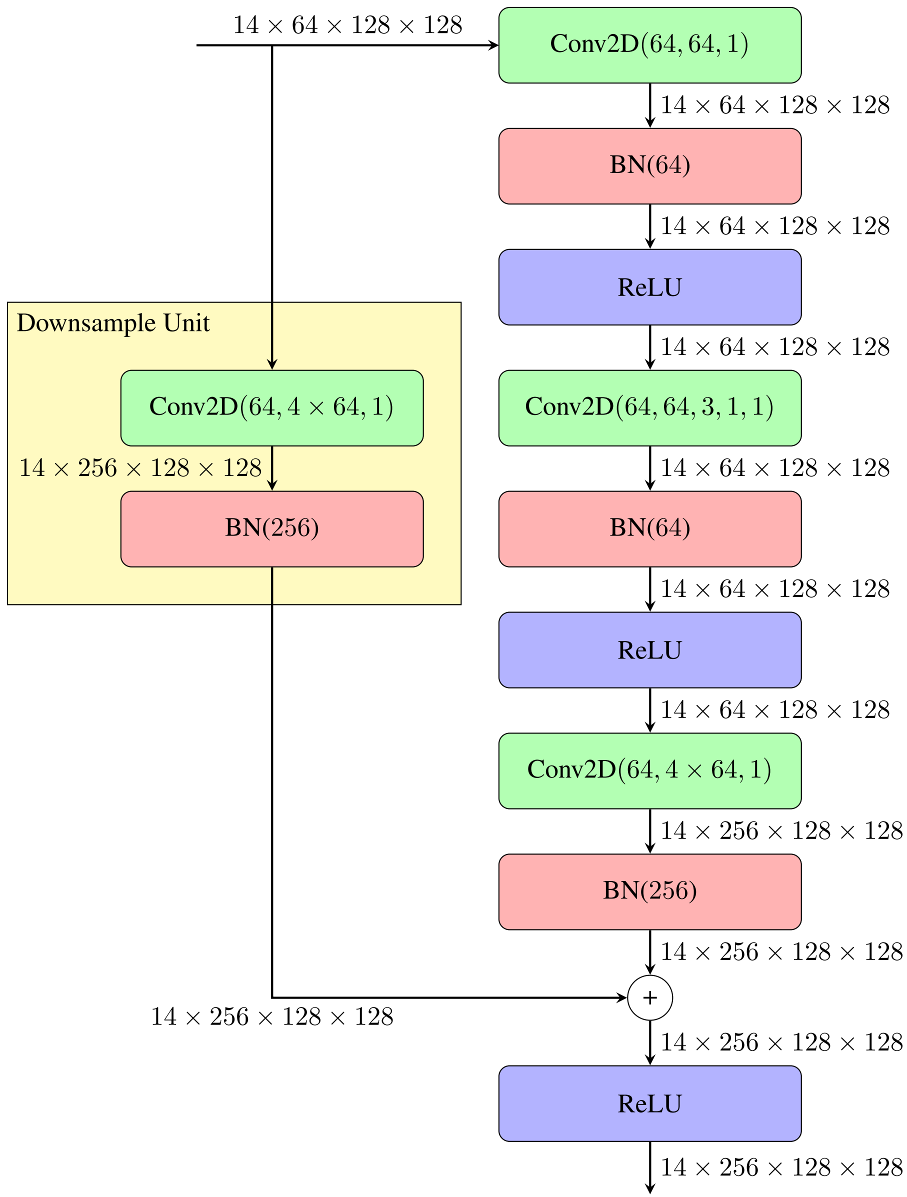

Figure 1: Bottleneck layer 1 with downsample unit

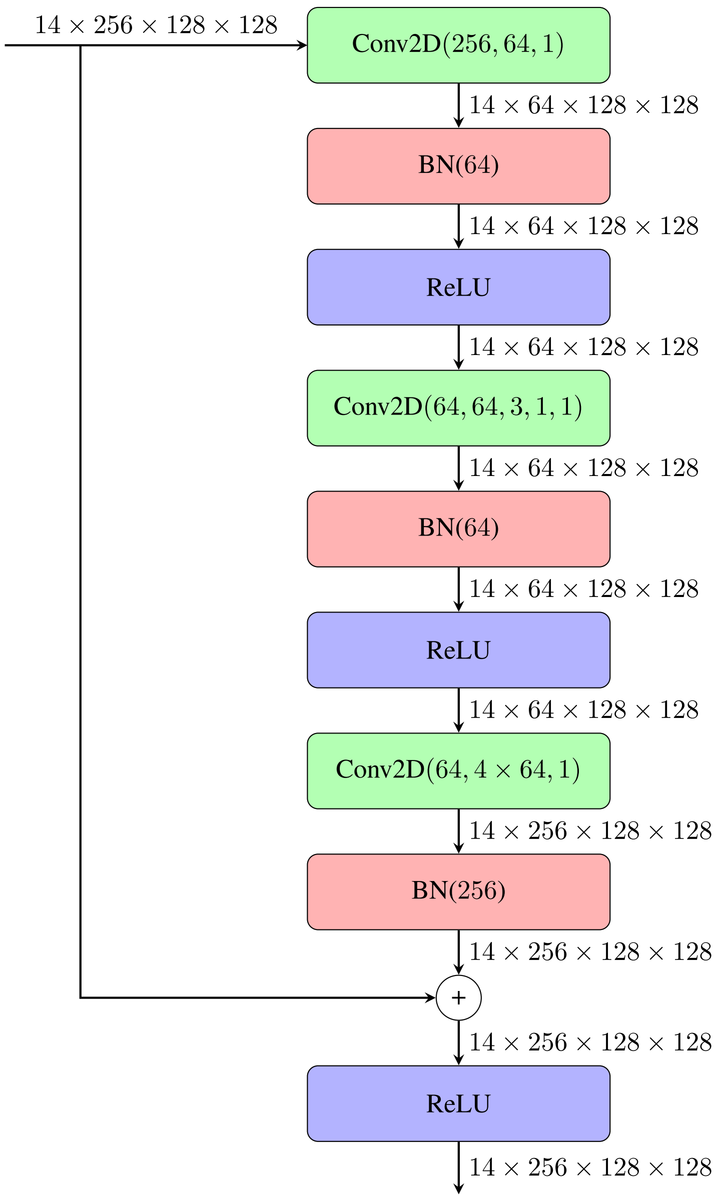

Figure 2: Bottleneck layer 1 without downsample unit

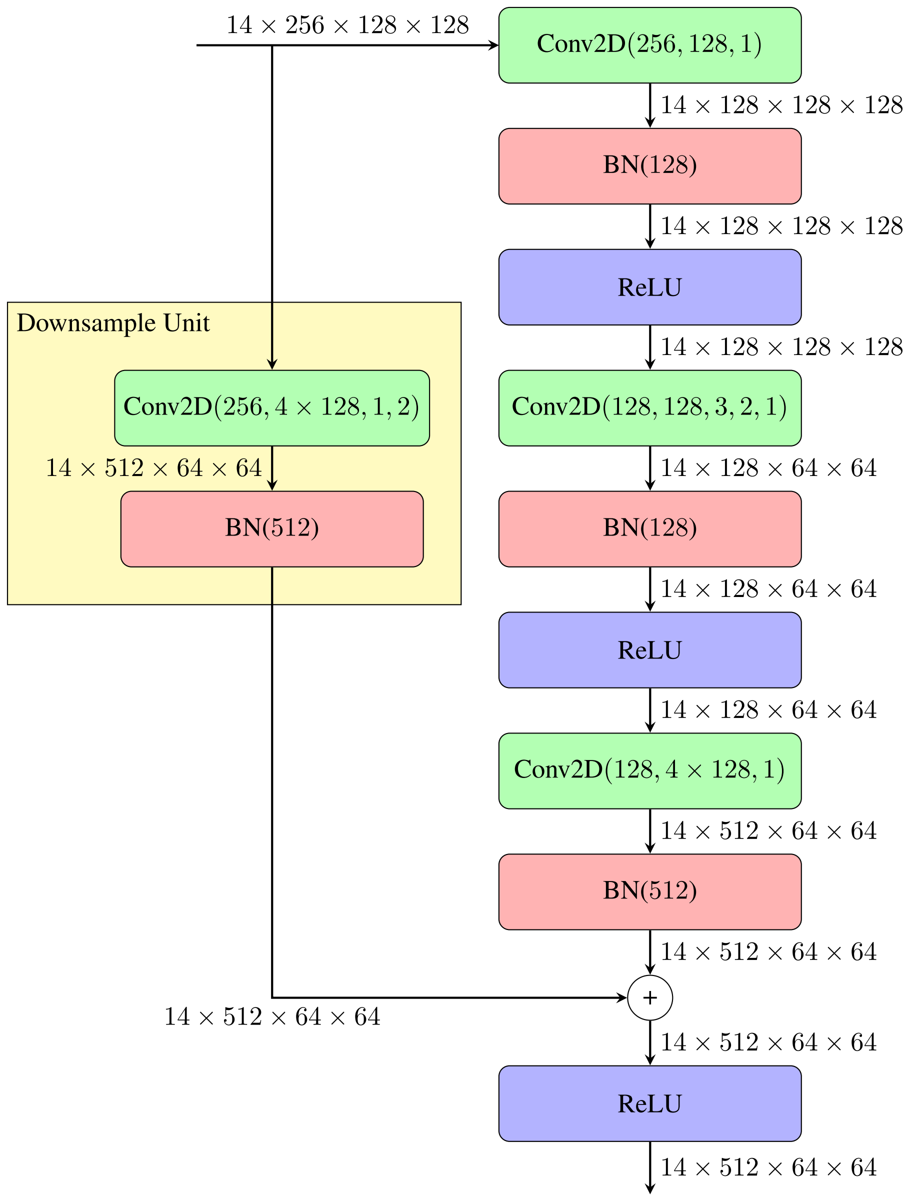

Figure 3: Bottleneck layer 2 with downsample unit

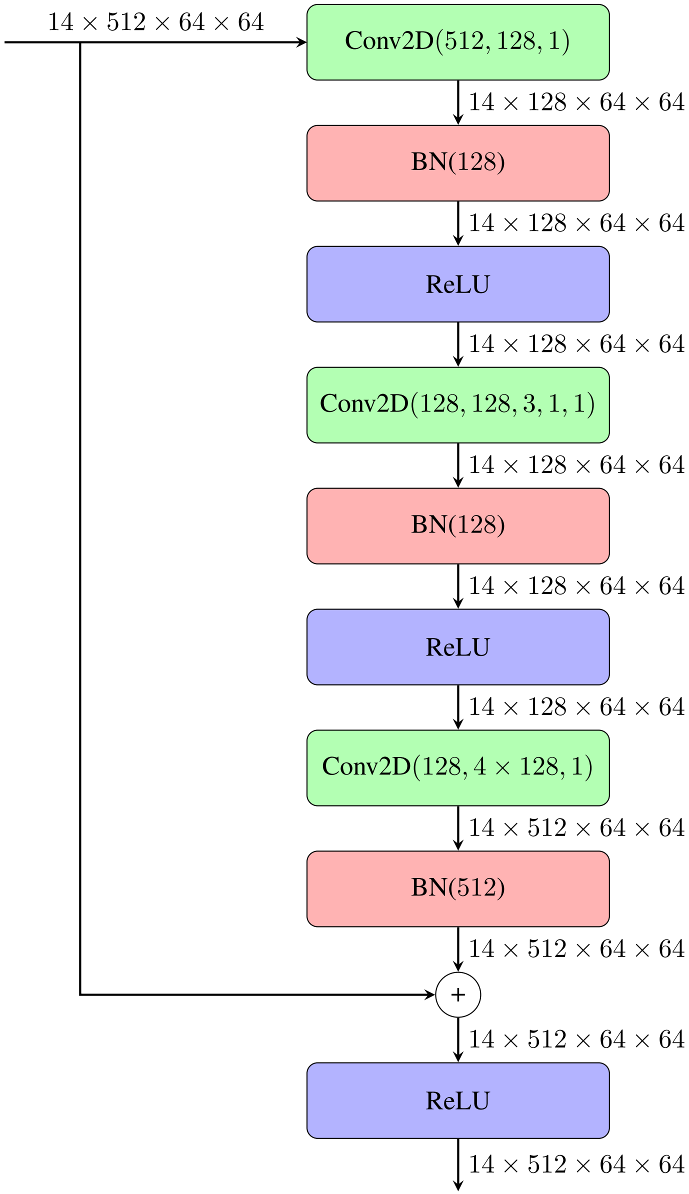

Figure 4: Bottleneck layer 2 without downsample unit

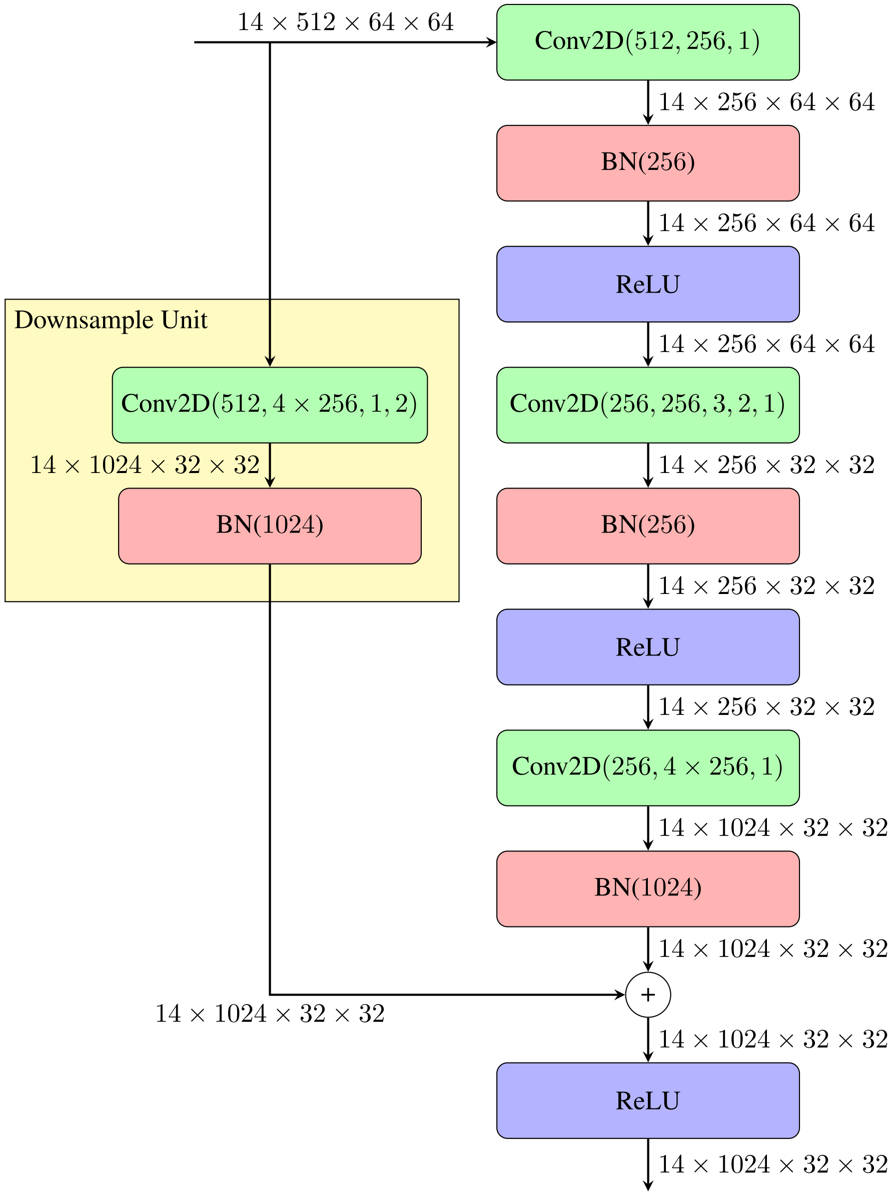

Figure 5: Bottleneck layer 3 with downsample unit

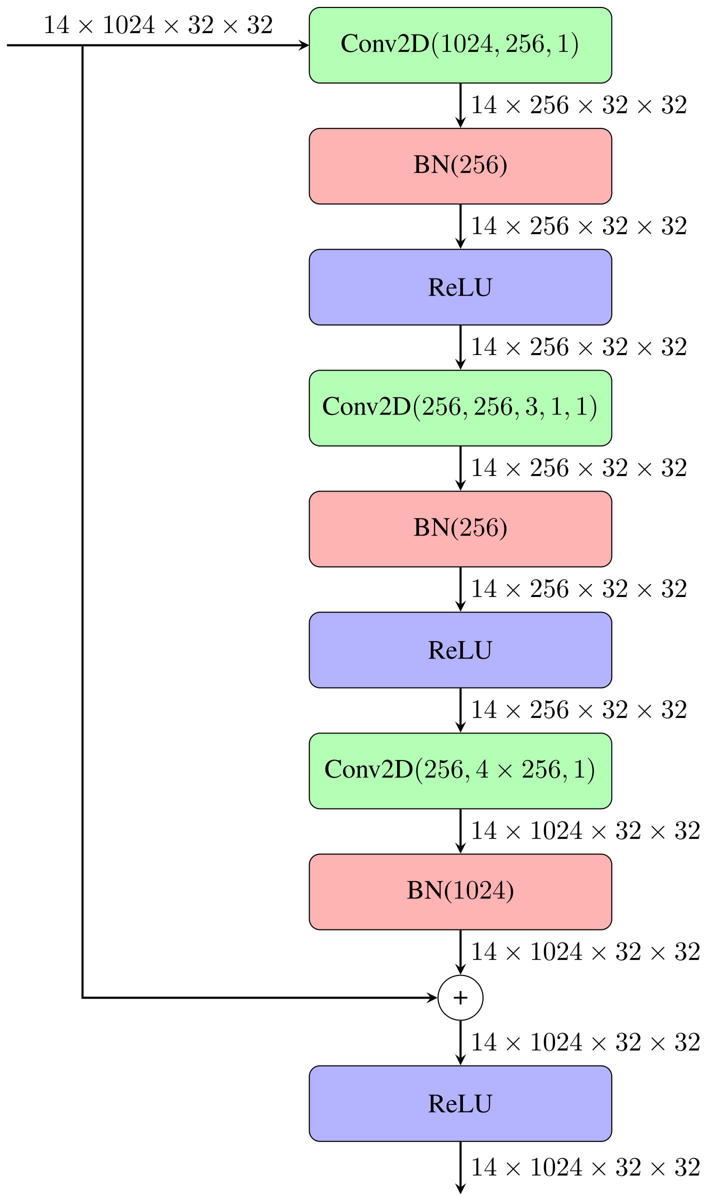

Figure 6: Bottleneck layer 3 without downsample unit

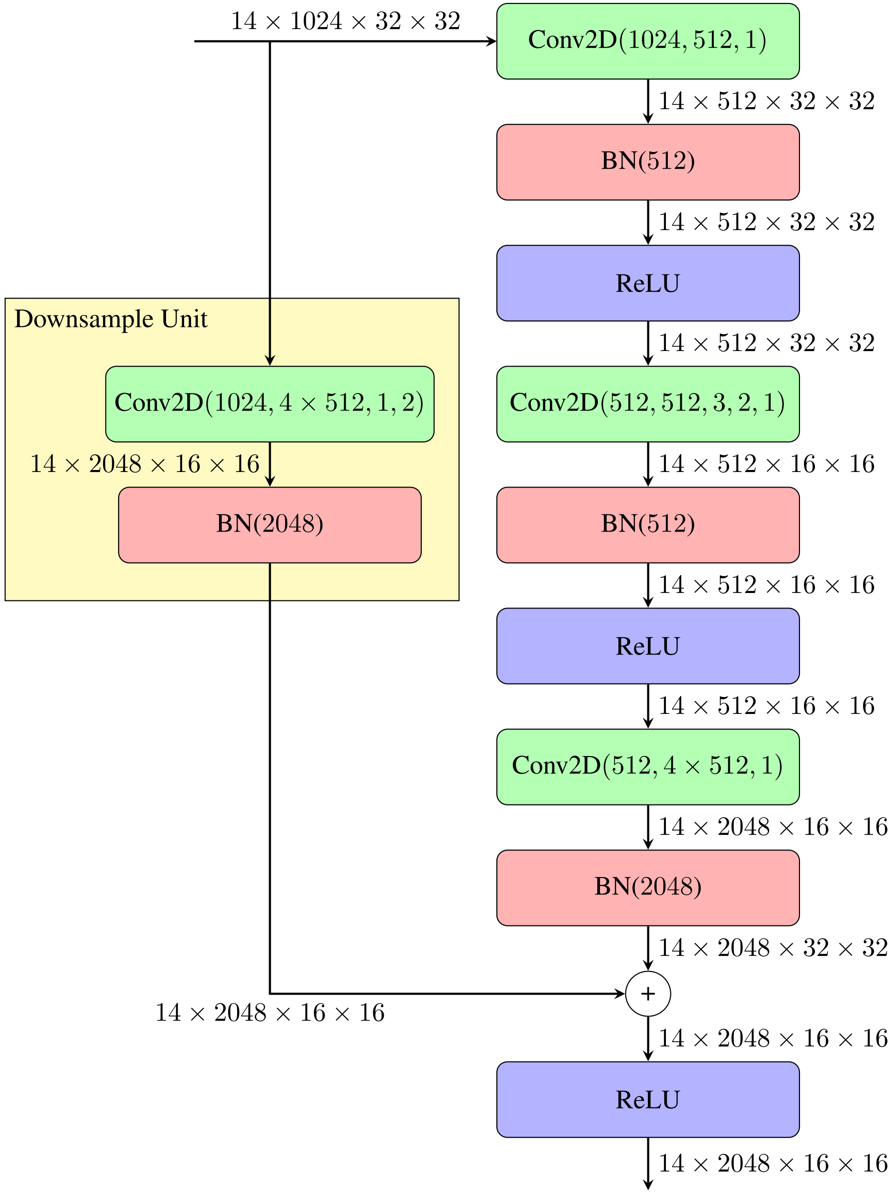

Figure 7: Bottleneck layer 4 with downsample unit

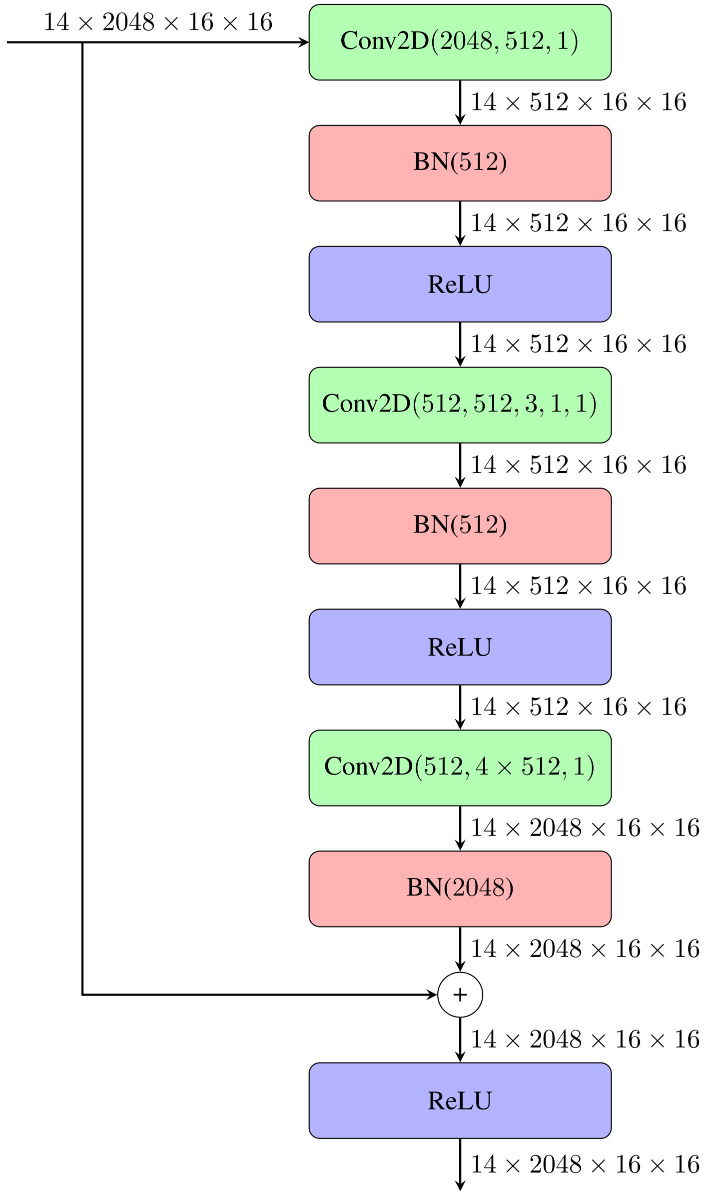

Figure 8: Bottleneck layer 4 without downsample unit

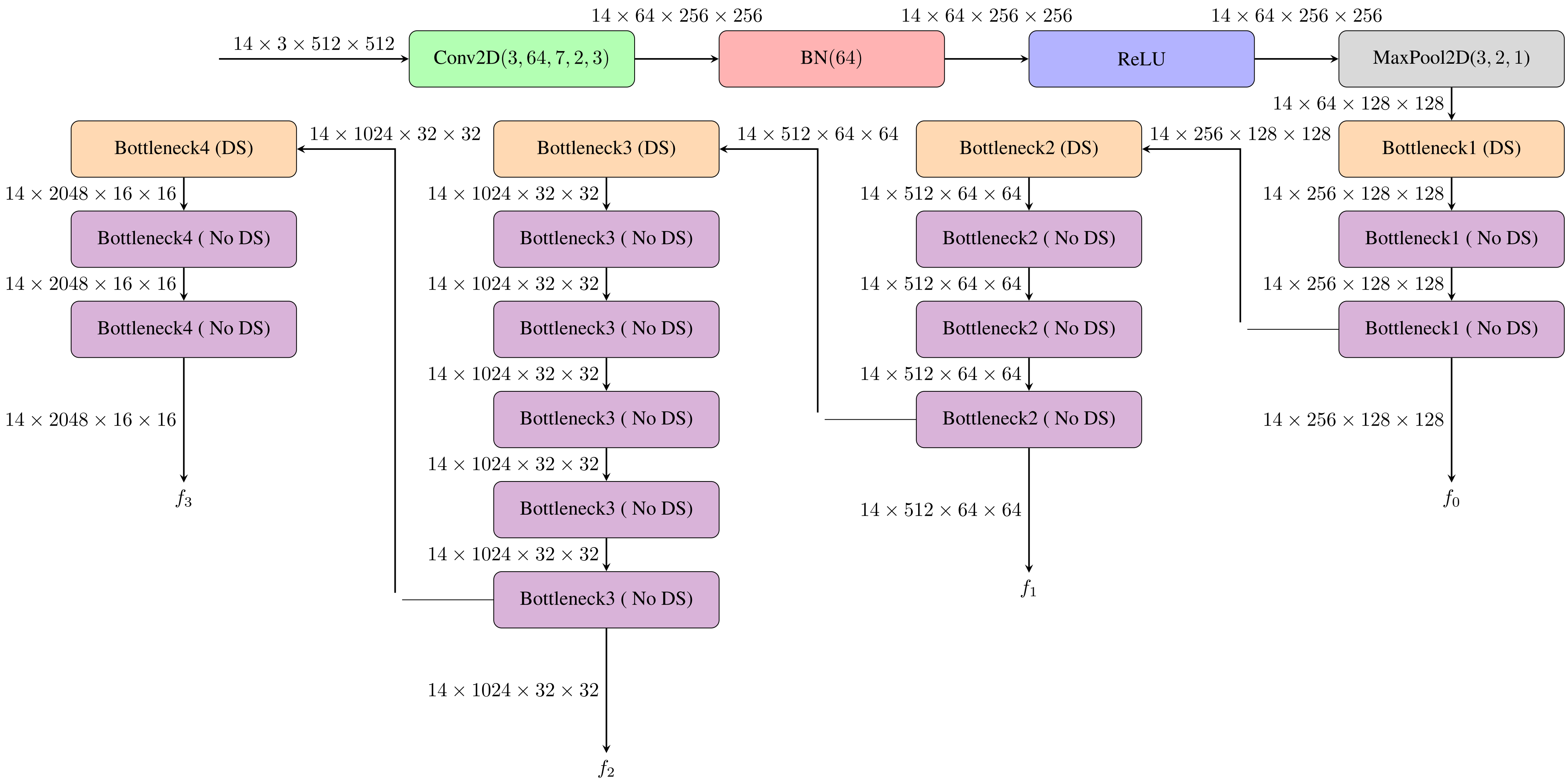

Figure 9: ResNet-50 architecture used.

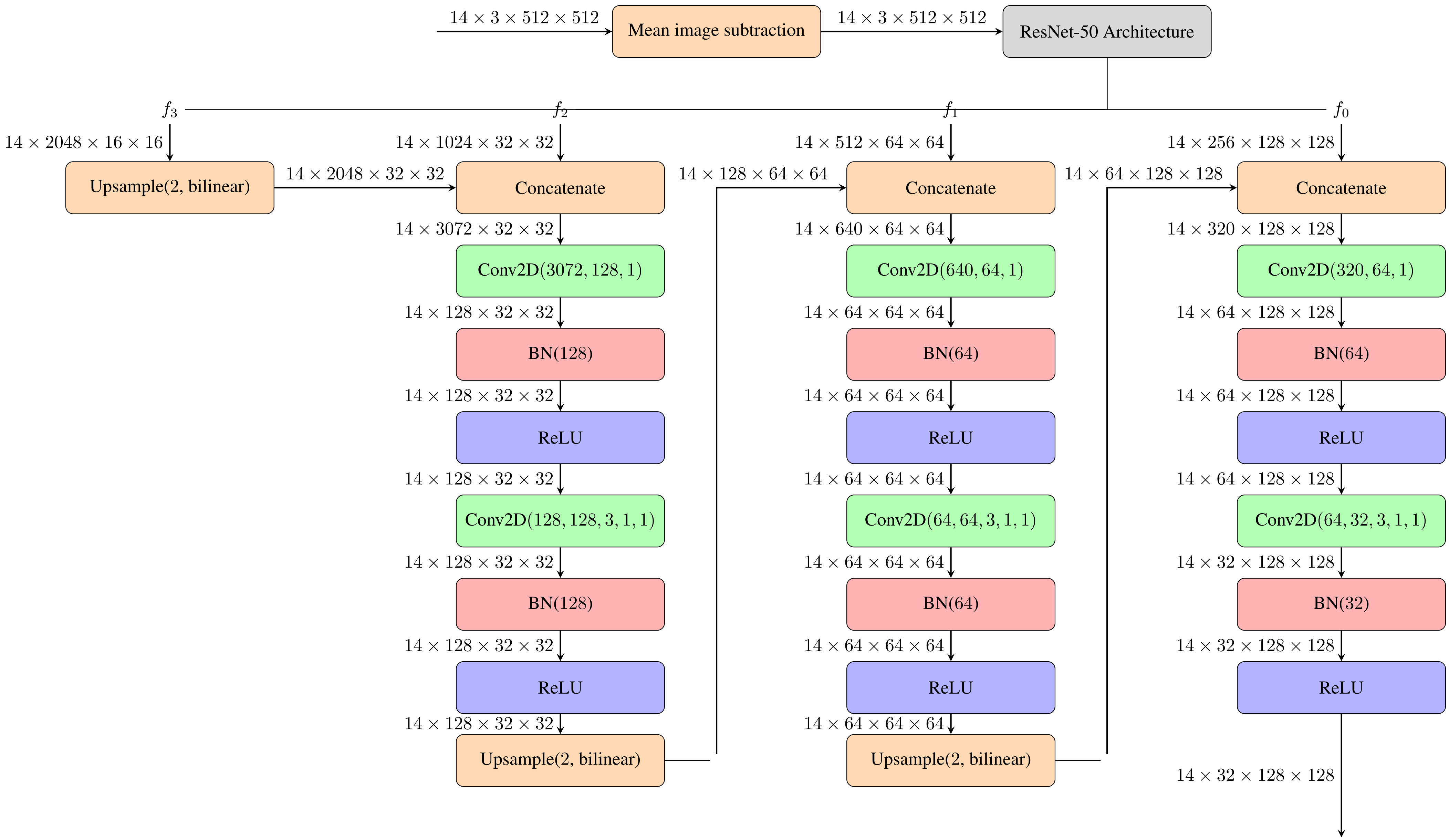

Figure 10: Architecture to concatenate ResNet outputs.

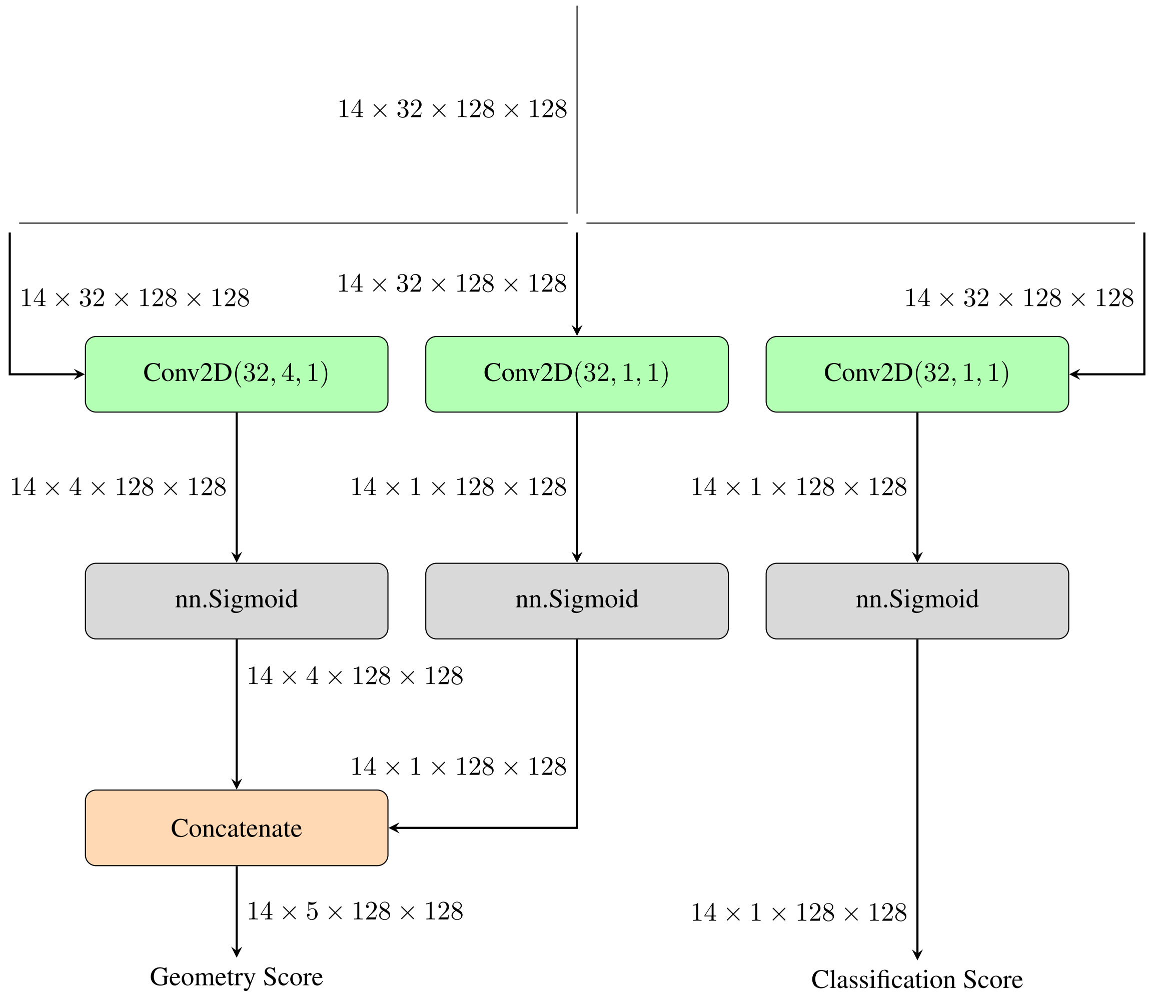

Figure 11: Final network layers.

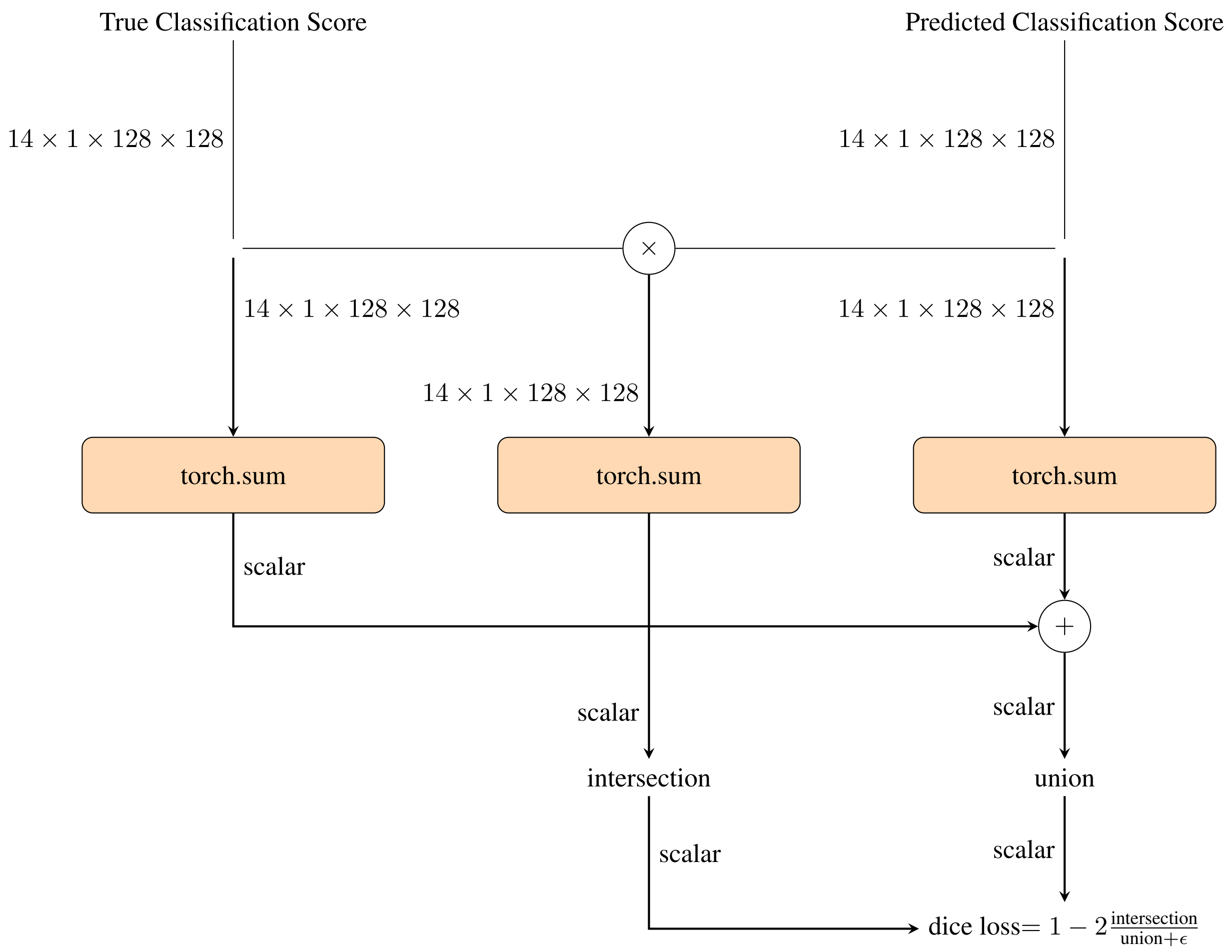

Figure 12: Dice loss calculation

2 Loss

2.1 Geometry loss calculation

Both predicted and true Geometry Score tensors of size $14 \times 5 \times 128 \times 128$ are split into 4 tensors each with sizes $14 \times 1 \times 128 \times 128$. These tensors are respectively called $d^1_{gt}, d^2_{gt}, d^3_{gt}, d^4_{gt}, \theta_{gt}$ and $d^1_{pr}, d^2_{pr}, d^3_{pr}, d^4_{pr}, \theta_{pr}$. Then the following element-wise operations are done to find tensors Area_{gt} and Area_{pr} of sizes $14 \times 1 \times 128 \times 128$,

$$\text{Area}_{gt} = (d^1_{gt} + d^3_{gt}) \odot (d^2_{gt} + d^4_{gt}) \tag{7}$$
$$\text{Area}_{pr} = (d^1_{pr} + d^3_{pr}) \odot (d^2_{pr} + d^4_{pr}). \tag{8}$$

To find the intersection area tensor, the following element-wise tensor operations are used which result is tensors w_{union} and h_{union} of sizes $14 \times 1 \times 128 \times 128$,

$$w_{\text{union}} = \min(d^2_{gt}, d^2_{pr}) + \min(d^4_{gt}, d^4_{pr}) \tag{9}$$
$$h_{\text{union}} = \min(d^1_{gt}, d^1_{pr}) + \min(d^3_{gt}, d^3_{pr}). \tag{10}$$

This allows us to compute the area intersection and union tensors of size $14 \times 1 \times 128 \times 128$ as follows,

$$\text{Area}_{\text{intersection}} = w_{\text{union}} \odot h_{\text{union}} \tag{11}$$
$$\text{Area}_{\text{union}} = \text{Area}_{gt} + \text{Area}_{pr} - \text{Area}_{\text{intersection}}. \tag{12}$$

Using these areas and based on element-wise tensor operations, the loss tensor L_{AABB} of size $14 \times 1 \times 128 \times 128$ is calculated as follows,

$$L_{\text{AABB}} = -\log\left(\frac{\text{Area}_{\text{intersection}} + 1}{\text{Area}_{\text{union}} + 1}\right). \tag{13}$$

The angle loss tensor L_θ and the overall loss tensor L_g of sizes $14 \times 1 \times 128 \times 128$ are also calculated as follows,

$$L_\theta = 1 - \cos(\theta_{pr} - \theta_{tr}) \tag{14}$$
$$L_g = L_{\text{AABB}} + 20L_\theta. \tag{15}$$

Finally, the overall loss value is calculates as

$$l = \text{torch.mean}(L_g \odot y_{tr}) + L_{\text{classification}} \tag{16}$$