**Question 1 [20+1+1+2 marks]**

A small CNN model consisting of 5 convolution layers. Each convolution layer would be followed by a ReLU activation and a max pooling layer.

Note: Stride is 2. Padding is 1.

1. Write down all missing values in the table.

| **Layer** | **Input** | **Output** | **Number of Parameters** | **Number of operations (multiplication)** |
| --- | --- | --- | --- | --- |
| Block1\_ Conv2D | (H,W,3) |  |  |  |
| Block1\_MaxPool |  |  |  |  |
| Block2\_Conv2D |  |  |  |  |
| Block2\_MaxPool |  |  |  |  |
| Block3\_Conv2D |  |  |  |  |
| Block3\_MaxPool |  |  |  |  |
| Block4\_Conv2D |  |  |  |  |
| Block4\_MaxPool |  |  |  |  |
| Block5\_Conv2D |  |  |  |  |
| Block5\_MaxPool |  |  |  |  |

1. What is the total number of computations done by your network? (assume m filters in each layer of size k×k. Exclude the dense layer.)
2. What is the total number of parameters in your network? (assume m filters in each layer of size k ×k. Exclude the dense layer.)
3. In the described CNN, how many neurons are in the input of the **first Dense layer**, assuming 5 max pooling operations follow conv layers with 32 filters?

**Question 2 [2+4 marks]**

1. Suppose your CNN model uses 6 convolution layers, each with 32 filters of size 3×3, stride 1, and “same” padding. If the input image is 256×256×3, what is the output feature map size after all 5 conv → ReLU → max pool blocks?
2. You apply a 3×3 filter (no padding, stride = 2) over a 5×5 input feature map to get a 3×3 output. A ReLU activation was applied after the convolution. During backpropagation, the δ-map from the next layer is:

δ=

The filter used during the forward pass was:

w=

**Tasks:**

a) Compute the gradient with respect to the weights (i.e., ∂L/∂w ​).