

Model Layers

Learn about the initial and final layers for ResNet.

Chapter Goals:

- Understand how the ResNet architecture is structured

A. Initial layers

Prior to the block layers, the ResNet model uses a convolution layer and max pooling layer on the input data. The convolution layer uses 7x7 kernels, which is larger than we've seen in previous models. The increased kernel size is specifically for the ImageNet dataset, which has larger images than our previous datasets.

The convolution layer uses a stride size of 2 for dimension reduction. The max pooling layer also uses a stride size of 2. In total, the height and width dimensions of the input are each reduced by a factor of 4 after the initial convolution and pooling layers. This reduces computational cost, since the input data for the first block layer is now 16 times smaller.

B. Final layers

The ResNet model ends with global average pooling and a fully-connected layer to obtain the logits. The fully-connected layer is the final weight layer of the model. So the weight layers of ResNet are the initial convolution layer, the final fully-connected layer, and the convolution layers within each of the blocks.

The number of weight layers in the model gives each variation of ResNet its name (e.g. ResNet-18, ResNet-50). For example, ResNet-18 has 2 blocks in each of its 4 block layers. The blocks are regular blocks, so they have 2 convolution layers. So ResNet-18 has 16 convolution layers across all its blocks, which gives it a total of 18 weight layers.

In the next chapter you'll see the full code for the ResNet model architecture

(including the block layers and logits layer).