

Pool5 is a global average pooling 2d layer, the input is the height, width, number of channels and batch size. This layer will compute the mean of the height and width dimension of the input and this layer will provide the mean of the height and weight value to the next layer fc.

Fc layer is a fully connected layer, the inputs are the height, width and number of channels that from the output of pool5. This layer will multiply the input by a weight matrix and then adds a bias vector. And it will provide output size vector to the next layer. And the the meaning that we change the output size is that we are going to output a size 5 vector.

Soft max is a soft max layer, this layer applies a soft max function to the input. And for soft max function, it's a function that takes S-by-Q matrix of the net input vector and return a S-by-Q matrix. The input is a 2D array that is about the output size. In the 2D array, there are number of class and batch size. So does the output, is is a 2D array that contain the number of class and batch size.

Finally, class output layer is a classification layer, this layer will compute the cross-entropy loss for classification and weighted classification tasks with mutually exclusive class. The input of the classification layer is the 2D array from soft max function along with the true labels that is a vector of true class label of batch size. For the output, it is the value representing the average cross-entropy loss.