Exercises

Operations in Convolutional Neural Networks

Francesca Galassi

Question 1.

Consider a Convolutional Neural Network (CNN) designed to process single-channel images with a size of 6×6.

This CNN consists of two convolutional layers, each followed by a ReLU activation function. We use a stride of 1 and no padding. The final layer is a fully connected layer, producing a single output unit. Bias terms are omitted for simplicity

Here's a detailed breakdown of the network's architecture :

— The first convolutional layer comprises three 3×3 filters. Let's consider the following values for each filter (these weights are typically learned during the training process of the CNN) :

$$W_1^{(1)} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}; \quad W_2^{(1)} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}; \quad W_3^{(1)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

— The second convolutional layer consists of a single $3 \times 3 \times 3$ filter. The weights are as follows :

$$W^{(2)}[i,j,k] = \begin{cases} 1, & \text{if } j = k = 2 \\ 0, & \text{otherwise} \end{cases} \qquad \text{for } i,j,k \in \{1,2,3\}$$

— The fully connected layer (third layer) consists of 4 units. The vector of weights associated with this fully connected layer is :

$$\mathbf{w}^{(3)} = \begin{bmatrix} 6 \\ 2 \\ 0 \\ -1 \end{bmatrix}.$$

1.1 Given the input image represented by the matrix X, where

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

compute the output of the first convolutional layer (or first hidden layer) $H_i^{(1)} = \mathrm{ReLU}(X * W_i^{(1)})$, where $i = 1, 2, 3$ refers to the channel index. Each channel represents a different feature map or filter applied to the input image.

1.2 Given the output of the first layer, $H_{1:3}^{(1)}$, i.e., a collection of three feature maps, compute the output of the second layer $H^{(2)} = \mathrm{ReLU}(H_{1:3}^{(1)} * W^{(2)})$.
The notation $H_{1:3}^{(1)}[i, j, k]$ refers to the elements of $H_{1:3}^{(1)}$, and the calculation is performed for i, j in $\{1, 2, 3, 4\}$, and for k in $\{1, 2, 3\}$.

1.3 Given the output of the second layer, $H^{(2)}$, compute the output of the network $y = \mathbf{w}^{(3)T} \cdot h^{(2)}$, where $h^{(2)}$ is the flattened output of the second layer.

Question 2.

Let's consider a Neural Network designed for an image classification task with the following output categories : car, bicycle, motorbike, and bus. The input to the network is a grayscale image of size 20×20 pixels.

Answer the following questions :

2.1 If the neural network comprises a fully connected layer only, with n units, how many parameters must be optimized by the gradient descent algorithm ?
Compute the number of parameters without and with bias terms.

2.2 Consider a neural network composed of the following layers :

1. Convolutional layer, with m filters of size 5×5.
2. Fully Connected layer with n units.

How many parameters does the gradient descent optimize, excluding bias terms ?

2.3 Consider a neural network composed of the following layers :

1. Convolutional layer, with m filters of size 5×5 ;
2. MaxPooling layer, with a window size of 4×4 and a stride of 4 ;
3. Fully Connected layer with n units.

How many parameters does the gradient descent optimize, excluding bias terms ?

2.4 Consider a neural network with the following layers :

1. Convolutional layer, consisting of m_1 filters of size 5×5.
2. Max Pooling layer, with window size of 2 and stride 2 ;
3. Convolutional layer, comprising m_2 filters of size $3 \times 3 \times m_1$;
4. Max Pooling layer, with window size of 2 and stride 2 ;
5. Fully Connected layer with n_1 units ;
6. Fully Connected layer with n_2 units.

Taking into account the bias for the output layer only, determine the total number of parameters optimized by gradient descent.

Question 3. Repeat the exercises in **Question 2** for an input RGB image.