

# Convolutions for neural networks

Arnaud Bergeron

August 10, 2015

Some vocabulary:

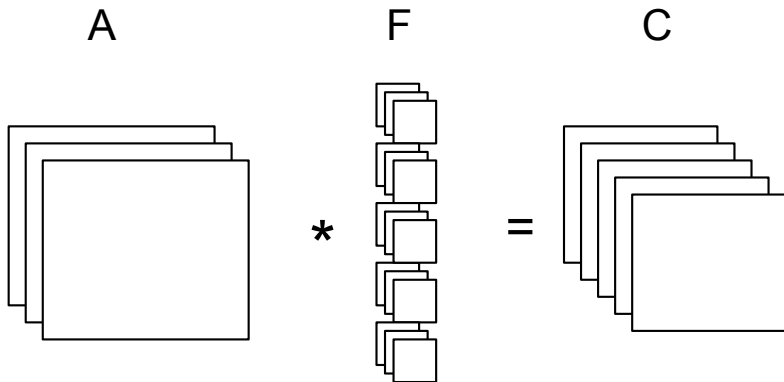
**feature map** analogue to layer (the activations, not the weights)  
for a neural network.

**channel** a single 2D "image" which composes a feature map.

**filter** weight (as in layer weights).

**stack** ???

The basic convolution operation (2D, discrete) for a single example



where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=-1}^{k=0} \sum_{m=-1}^{x=0} \sum_{n=-1}^{y=0} A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=-1}^{k=0} \sum_{m=-1}^{x=0} \sum_{n=-1}^{y=0} A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

- $k$  is the input channel.

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=1}^{k=0} \sum_{m=1}^{x=0} \sum_{n=1}^{y=0} A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

- ▶  $k$  is the input channel.
- ▶  $o$  is the output channel.

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=0}^{k-1} \sum_{m=0}^{x-1} \sum_{n=0}^{y-1} A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

- ▶  $k$  is the input channel.
- ▶  $o$  is the output channel.
- ▶  $l$  is the number of input channels.

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=0}^{k-1} \sum_{m=0}^{x-1} \sum_{n=0}^{y-1} A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

- ▶  $k$  is the input channel.
- ▶  $o$  is the output channel.
- ▶  $l$  is the number of input channels.
- ▶  $(m, n)$  is the size of the filters.

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.



The basic convolution operation (2D, discrete) for a single example

$$C_{bo}(i, j) = \sum_{l=1}^k \sum_{m=1}^x \sum_{n=1}^y A_{bk}(x + i - \frac{m}{2}, y + j - \frac{n}{2}) F_{ko}(x, y)$$

- ▶  $k$  is the input channel.
- ▶  $o$  is the output channel.
- ▶  $l$  is the number of input channels.
- ▶  $(m, n)$  is the size of the filters.
- ▶  $b$  is the batch.

where  $A$  is the input,  $F$  are the filters and  $C$  is the convolution result.

Memory layout for images: 'bc01'

- ▶ first dimension is the batch ('b')
- ▶ second dimension is the channel ('c')
- ▶ last two dimensions are the data ('0', '1')

Memory layout for images: 'bc01'

- ▶ first dimension is the batch ('b')
- ▶ second dimension is the channel ('c')
- ▶ last two dimensions are the data ('0', '1')

Memory layout for filters: 'nc01'

- ▶ first dimension is the output channel ('n')
- ▶ second dimension is the input channel ('c')
- ▶ last two dimensions are the data ('0', '1')

Memory layout for images: 'bc01'

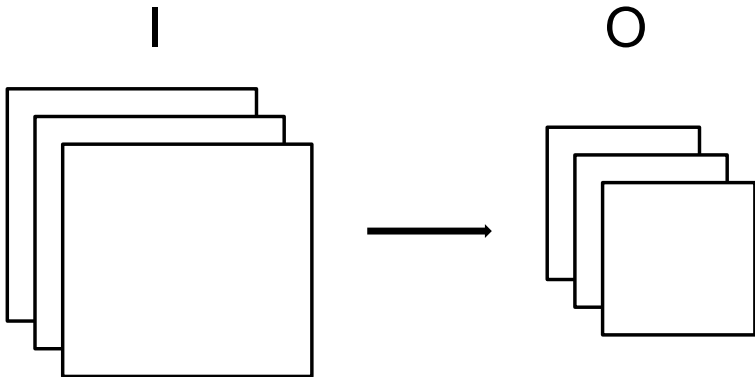
- ▶ first dimension is the batch ('b')
- ▶ second dimension is the channel ('c')
- ▶ last two dimensions are the data ('0', '1')

Memory layout for filters: 'nc01'

- ▶ first dimension is the output channel ('n')
- ▶ second dimension is the input channel ('c')
- ▶ last two dimensions are the data ('0', '1')

Some other packages may use different conventions.

Basic pooling operation (max)



Basic pooling operation (max)

$$O_k(i, j) = \max_{\substack{0 \leq x < m \\ 0 \leq y < n}} I_k(x + i - \frac{m}{2}, y + j - \frac{n}{2})$$

Basic pooling operation (max)

$$O_k(i, j) = \max_{\substack{0 \leq x < m \\ 0 \leq y < n}} I_k(x + i - \frac{m}{2}, y + j - \frac{n}{2})$$

- $k$  is the channel.

Basic pooling operation (max)

$$O_k(i, j) = \max_{\substack{0 \leq x < m \\ 0 \leq y < n}} I_k(x + i - \frac{m}{2}, y + j - \frac{n}{2})$$

- ▶  $k$  is the channel.
- ▶  $(m, n)$  is the size of the filters.



- ▶ One thing to watch out for is the border handling
- ▶ Other types of pooling exist such as average pooling