Forward And Backpropagation in Convolutional Neural Network.
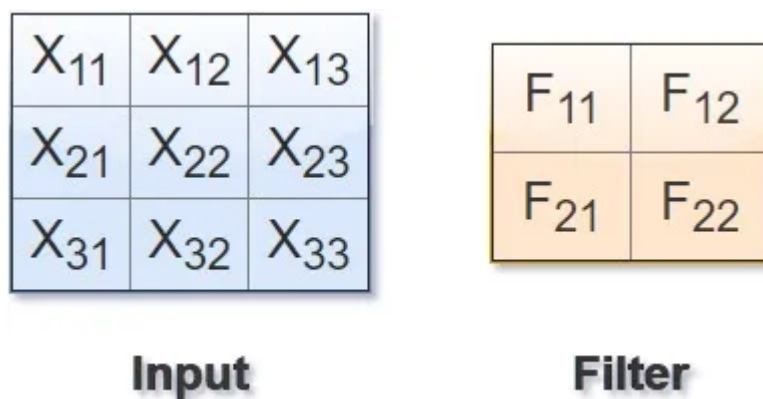
Sujit Rai · Follow

3 min read · Dec 24, 2017

The below post demonstrates the use of convolution operation for carrying out the back propagation in a CNN.

Difference between Convolution Operation and Correlation.

Input   Filter

Let's consider the input and the filter that is going to be used for carrying out the convolution as given above.

Then the correlation of the filter matrix with the input matrix is described in the figure below

$$\begin{array}{|c|c|} \hline O_{11} & O_{12} \\ \hline O_{21} & O_{22} \\ \hline \end{array} = \text{Correlation}\left(\begin{array}{|c|c|c|} \hline X_{11} & X_{12} & X_{13} \\ \hline X_{21} & X_{22} & X_{23} \\ \hline X_{31} & X_{32} & X_{33} \\ \hline \end{array}, \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array}\right)$$
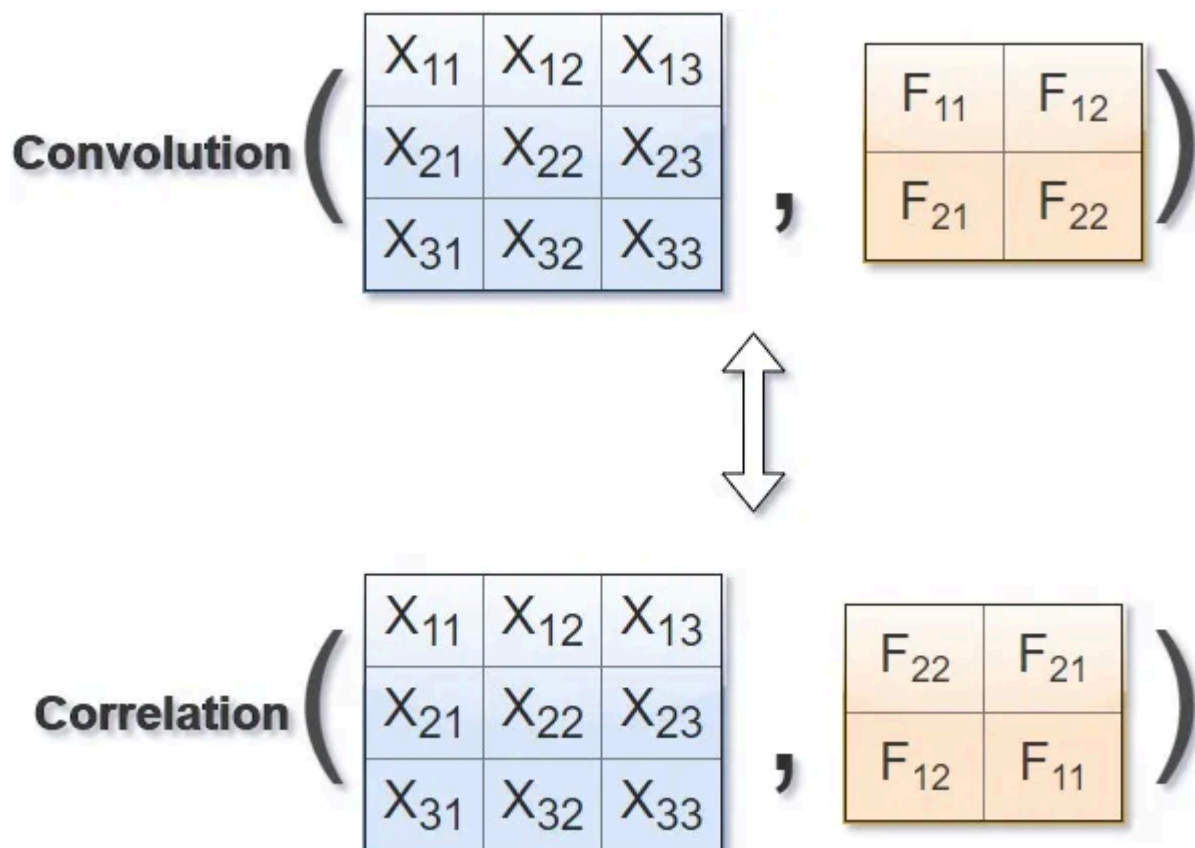
$$O_{11} = F_{11}X_{11} + F_{12}X_{12} + F_{21}X_{21} + F_{22}X_{22}$$

$$O_{12} = F_{11}X_{12} + F_{12}X_{13} + F_{21}X_{22} + F_{22}X_{23}$$

$$O_{21} = F_{11}X_{21} + F_{12}X_{22} + F_{21}X_{31} + F_{22}X_{32}$$

$$O_{22} = F_{11}X_{22} + F_{12}X_{23} + F_{21}X_{32} + F_{22}X_{33}$$

Now, The convolution of the filter matrix with input image is same as rotating the filter by 180 degrees and then carrying out the correlation of the rotated filter matrix with the input matrix.

$$\text{Convolution}\left(\begin{array}{|c|c|c|} \hline X_{11} & X_{12} & X_{13} \\ \hline X_{21} & X_{22} & X_{23} \\ \hline X_{31} & X_{32} & X_{33} \\ \hline \end{array}, \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array}\right)$$

$$\updownarrow$$

$$\text{Correlation}\left(\begin{array}{|c|c|c|} \hline X_{11} & X_{12} & X_{13} \\ \hline X_{21} & X_{22} & X_{23} \\ \hline X_{31} & X_{32} & X_{33} \\ \hline \end{array}, \begin{array}{|c|c|} \hline F_{22} & F_{21} \\ \hline F_{12} & F_{11} \\ \hline \end{array}\right)$$

As can be seen from the above image the convolution operation is same as that of the correlation operation but with rotated filter.

Forward and Backward Propagation using Convolution operation.

Note : To derive the equation of the gradients for the filter values and the input matrix values we will consider that the convolution operation is same as correlation operation, just for simplicity.

Therefore, The convolution operation can be written as described in the figure below.

$$\begin{array}{|c|c|} \hline O_{11} & O_{12} \\ \hline O_{21} & O_{22} \\ \hline \end{array} = \text{Convolution}\left(\begin{array}{|c|c|c|} \hline X_{11} & X_{12} & X_{13} \\ \hline X_{21} & X_{22} & X_{23} \\ \hline X_{31} & X_{32} & X_{33} \\ \hline \end{array} , \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array} \right)$$

$$O_{11} = F_{11}X_{11} + F_{12}X_{12} + F_{21}X_{21} + F_{22}X_{22}$$
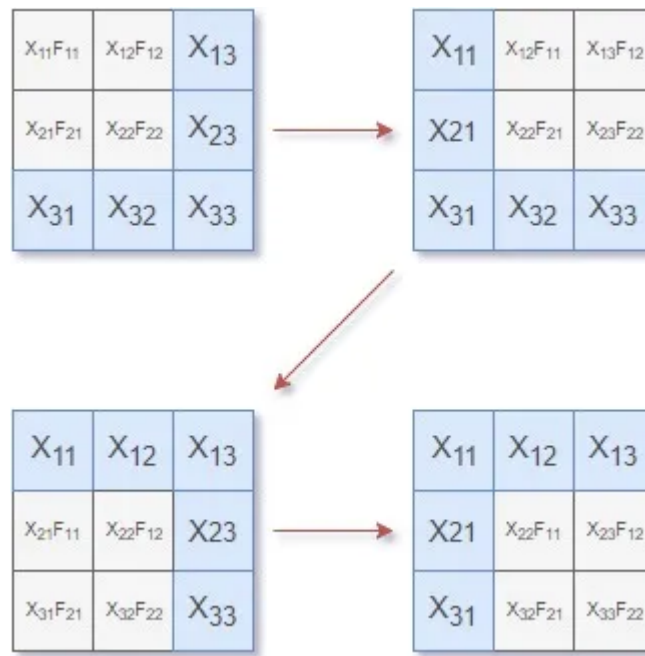$$O_{12} = F_{11}X_{12} + F_{12}X_{13} + F_{21}X_{22} + F_{22}X_{23}$$
$$O_{21} = F_{11}X_{21} + F_{12}X_{22} + F_{21}X_{31} + F_{22}X_{32}$$
$$O_{22} = F_{11}X_{22} + F_{12}X_{23} + F_{21}X_{32} + F_{22}X_{33}$$

Notice that here the filter is not rotated for sake of simplicity, therefore here convolution is same as correlation

It can be visualized in the figure below.

$X_{11}F_{11}$	$X_{12}F_{12}$	X_{13}
$X_{21}F_{21}$	$X_{22}F_{22}$	X_{23}
X_{31}	X_{32}	X_{33}

X_{11}	$X_{12}F_{11}$	$X_{13}F_{12}$
X_{21}	$X_{22}F_{21}$	$X_{23}F_{22}$
X_{31}	X_{32}	X_{33}

X_{11}	X_{12}	X_{13}
$X_{21}F_{11}$	$X_{22}F_{12}$	X_{23}
$X_{31}F_{21}$	$X_{32}F_{22}$	X_{33}

X_{11}	X_{12}	X_{13}
X_{21}	$X_{22}F_{11}$	$X_{23}F_{12}$
X_{31}	$X_{32}F_{21}$	$X_{33}F_{22}$

Now, to calculate the gradients of filter 'F' with respect to the error 'E', following equations needs to solved.

$$\frac{\partial E}{\partial F_{11}} = \frac{\partial E}{\partial O_{11}}\frac{\partial O_{11}}{\partial F_{11}} + \frac{\partial E}{\partial O_{12}}\frac{\partial O_{12}}{\partial F_{11}} + \frac{\partial E}{\partial O_{21}}\frac{\partial O_{21}}{\partial F_{11}} + \frac{\partial E}{\partial O_{22}}\frac{\partial O_{22}}{\partial F_{11}}$$

$$\frac{\partial E}{\partial F_{12}} = \frac{\partial E}{\partial O_{11}}\frac{\partial O_{11}}{\partial F_{12}} + \frac{\partial E}{\partial O_{12}}\frac{\partial O_{12}}{\partial F_{12}} + \frac{\partial E}{\partial O_{21}}\frac{\partial O_{21}}{\partial F_{12}} + \frac{\partial E}{\partial O_{22}}\frac{\partial O_{22}}{\partial F_{12}}$$

$$\frac{\partial E}{\partial F_{21}} = \frac{\partial E}{\partial O_{11}}\frac{\partial O_{11}}{\partial F_{21}} + \frac{\partial E}{\partial O_{12}}\frac{\partial O_{12}}{\partial F_{21}} + \frac{\partial E}{\partial O_{21}}\frac{\partial O_{21}}{\partial F_{21}} + \frac{\partial E}{\partial O_{22}}\frac{\partial O_{22}}{\partial F_{21}}$$

$$\frac{\partial E}{\partial F_{22}} = \frac{\partial E}{\partial O_{11}}\frac{\partial O_{11}}{\partial F_{22}} + \frac{\partial E}{\partial O_{12}}\frac{\partial O_{12}}{\partial F_{22}} + \frac{\partial E}{\partial O_{21}}\frac{\partial O_{21}}{\partial F_{22}} + \frac{\partial E}{\partial O_{22}}\frac{\partial O_{22}}{\partial F_{22}}$$

Here, E is the error obtained.

which evaluates to

$$\frac{\partial E}{\partial F_{11}} = \frac{\partial E}{\partial O_{11}} X_{11} + \frac{\partial E}{\partial O_{12}} X_{12} + \frac{\partial E}{\partial O_{21}} X_{21} + \frac{\partial E}{\partial O_{22}} X_{22}$$

$$\frac{\partial E}{\partial F_{12}} = \frac{\partial E}{\partial O_{11}} X_{12} + \frac{\partial E}{\partial O_{12}} X_{13} + \frac{\partial E}{\partial O_{21}} X_{22} + \frac{\partial E}{\partial O_{22}} X_{23}$$

$$\frac{\partial E}{\partial F_{21}} = \frac{\partial E}{\partial O_{11}} X_{21} + \frac{\partial E}{\partial O_{12}} X_{22} + \frac{\partial E}{\partial O_{21}} X_{31} + \frac{\partial E}{\partial O_{22}} X_{32}$$

$$\frac{\partial E}{\partial F_{22}} = \frac{\partial E}{\partial O_{11}} X_{22} + \frac{\partial E}{\partial O_{12}} X_{23} + \frac{\partial E}{\partial O_{21}} X_{32} + \frac{\partial E}{\partial O_{22}} X_{33}$$

If we look closely this above equation can be written in form of our convolution operation.

$\partial E/\partial F_{11}$	$\partial E/\partial F_{12}$
$\partial E/\partial F_{21}$	$\partial E/\partial F_{22}$

= Convolution

X_{11}	X_{12}	X_{13}
X_{21}	X_{22}	X_{23}
X_{31}	X_{32}	X_{33}

,

$\partial E/\partial O_{11}$	$\partial E/\partial O_{12}$
$\partial E/\partial O_{21}$	$\partial E/\partial O_{22}$

Similarly we can find the gradients of the input matrix 'X' with respect to the error 'E'.

$$\frac{\partial E}{\partial X_{11}} = \frac{\partial E}{\partial O_{11}}F_{11} + \frac{\partial E}{\partial O_{12}}0 + \frac{\partial E}{\partial O_{21}}0 + \frac{\partial E}{\partial O_{22}}0$$

$$\frac{\partial E}{\partial X_{12}} = \frac{\partial E}{\partial O_{11}}F_{12} + \frac{\partial E}{\partial O_{12}}F_{11} + \frac{\partial E}{\partial O_{21}}0 + \frac{\partial E}{\partial O_{22}}0$$

$$\frac{\partial E}{\partial X_{13}} = \frac{\partial E}{\partial O_{11}}0 + \frac{\partial E}{\partial O_{12}}F_{12} + \frac{\partial E}{\partial O_{21}}0 + \frac{\partial E}{\partial O_{22}}0$$

$$\frac{\partial E}{\partial X_{21}} = \frac{\partial E}{\partial O_{11}}F_{21} + \frac{\partial E}{\partial O_{12}}0 + \frac{\partial E}{\partial O_{21}}F_{11} + \frac{\partial E}{\partial O_{22}}0$$

$$\frac{\partial E}{\partial X_{22}} = \frac{\partial E}{\partial O_{11}}F_{22} + \frac{\partial E}{\partial O_{12}}F_{21} + \frac{\partial E}{\partial O_{21}}f_{12} + \frac{\partial E}{\partial O_{22}}F_{11}$$

$$\frac{\partial E}{\partial X_{23}} = \frac{\partial E}{\partial O_{11}}0 + \frac{\partial E}{\partial O_{12}}F_{22} + \frac{\partial E}{\partial O_{21}}0 + \frac{\partial E}{\partial O_{22}}F_{11}$$

$$\frac{\partial E}{\partial X_{31}} = \frac{\partial E}{\partial O_{11}}0 + \frac{\partial E}{\partial O_{12}}0 + \frac{\partial E}{\partial O_{21}}F_{21} + \frac{\partial E}{\partial O_{22}}0$$

$$\frac{\partial E}{\partial X_{32}} = \frac{\partial E}{\partial O_{11}}0 + \frac{\partial E}{\partial O_{12}}0 + \frac{\partial E}{\partial O_{21}}F_{22} + \frac{\partial E}{\partial O_{22}}F_{21}$$

$$\frac{\partial E}{\partial X_{33}} = \frac{\partial E}{\partial O_{11}}0 + \frac{\partial E}{\partial O_{12}}0 + \frac{\partial E}{\partial O_{21}}0 + \frac{\partial E}{\partial O_{22}}F_{22}$$

Now, the above computation can be obtained by a different type of convolution operation known as full convolution. In order to obtain the gradients of the input matrix we need to rotate the filter by 180 degree and calculate the full convolution of the rotated filter by the gradients of the output with respect to error, As represented in the image below.

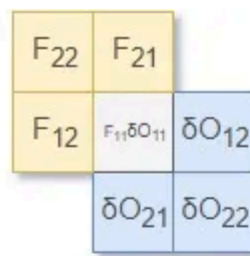| $\partial E/\partial X_{11}$	$\partial E/\partial X_{12}$	$\partial E/\partial X_{13}$
$\partial E/\partial X_{21}$	$\partial E/\partial X_{22}$	$\partial E/\partial X_{23}$
$\partial E/\partial X_{31}$	$\partial E/\partial X_{32}$	$\partial E/\partial X_{33}$

= Full_Convolution

$\left(\begin{array}{|c|c|} \partial E/\partial O_{11} & \partial E/\partial O_{12} \\ \partial E/\partial O_{21} & \partial E/\partial O_{22} \end{array} , \begin{array}{|c|c|} F_{22} & F_{21} \\ F_{12} & F_{11} \end{array} \right)$
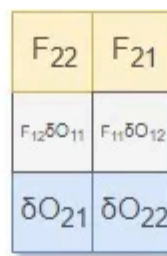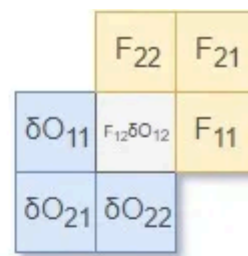
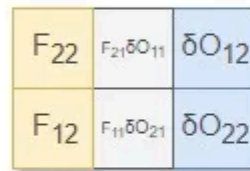The full convolution can be visualized as carrying out the procedure as represented in the figure below.

$\delta X_{11} = F_{11}\delta O_{11}$

$\delta X_{12} = F_{12}\delta O_{11} + F_{11}\delta O_{12}$
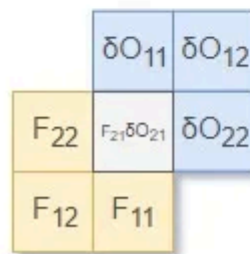
$\delta X_{13} = F_{12}\delta O_{12}$
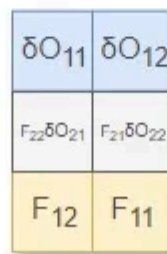
$\delta X_{21} = F_{21}\delta O_{11} + F_{11}\delta O_{21}$

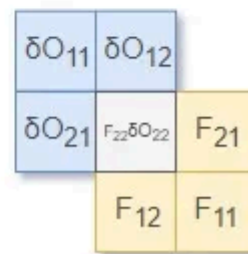$\delta X_{22} = F_{22}\delta O_{11} + F_{21}\delta O_{12} + F_{12}\delta O_{21} + F_{11}\delta O_{22}$

$\delta X_{23} = F_{22}\delta O_{12} + F_{12}\delta O_{22}$

$\delta X_{31} = F_{21}\delta O_{21}$

$\delta X_{32} = F_{22}\delta O_{21} + F_{21}\delta O_{22}$

$\delta X_{33} = F_{22}\delta O_{22}$

Here 'δX' represents the gradients of error with respect to X

Hence both the forward and backward propagation can be performed using the convolution operation.

For calculating the gradients of the pooling and Relu layers the gradients can be calculated by following the same procedure of using chain rule of derivatives.

Machine Learning   Cnn   Backpropagation Cnn

Written by Sujit Rai

Follow

125 Followers · 18 Following