6.819 PSET 4
10/05/2017

ISAAC KONTOMAH

Collaborators:Devin Morgan , Suman Nepal , Abishkar Chhetri , Khamoya Ikhofua,Afika Nyati

1 Convolution

a.*Forward Propagation*
$x_{out}[i] = f_n(x_{in})$

$$x_{out} = \sum_{j=0}^{k-1} x_{in}[i - \frac{k-1}{2} + j]W[j]$$

Where i is specific index of x_{in} where convolution is being done and j ranges from 0 to $k-1$

b. *Gradients of Input and Update Rules*
$W^{(i+1) = W^i + \eta \frac{\partial C}{\partial W}}$
$\frac{\partial C}{\partial X_{in1}} = \frac{\partial C}{\partial X_{out1}} \frac{\partial X_{out1}}{\partial X_{in1}} + ... + \frac{\partial C}{\partial X_{outN}} \frac{\partial X_{outN}}{\partial X_{in1}}$
$\frac{\partial X_{outi}}{\partial X_{inj}} = \frac{\partial}{\partial x_{in}[i]}(x_{in}[i - \frac{k-1}{2} + 1]W[1] + x_{in}[i - \frac{k-1}{2} + 2]W[2] + ... + x_{in}[i - \frac{k-1}{2} + g]w[g]), g = [0, k-1]$
$= \frac{\partial}{\partial x_{in}[i]}(x_{in}[i - \frac{k-1}{2} + g]W[\frac{k-1}{2}]), i = i + g - \frac{k-1}{2}$, hence $g = \frac{k-1}{2}$
$= \frac{\partial}{\partial x_{in}[i]}(x_{in}[i - \frac{k-1}{2} - (i-j)]W[\frac{k-1}{2}])$
$= \frac{\partial}{\partial x_{in}[i]}(x_{in}[i - \frac{k-1}{2} + j]W[\frac{k-1}{2}]) = W[\frac{k-1}{2}]$

$$\frac{\partial C}{\partial x_{in}[j]} = \sum_{i=0}^{N-1} \frac{\partial C}{\partial x_{out}[i]} \frac{\partial x_{out}[i]}{\partial x_{in}[j]} = \sum_{i=0}^{N-1} \frac{\partial C}{\partial x_{out}[i - \frac{k-1}{2} + j]} W[\frac{k-1}{2}]$$

Hence

$$\frac{\partial C}{\partial x_{in}[j]} = \sum_{i=0}^{N-1} \frac{\partial C}{\partial x_{out}[i - \frac{k-1}{2} + j]} W[\frac{k-1}{2}]$$

$$\frac{\partial C}{\partial W[j]} = \sum_{i=0}^{k-1} \frac{\partial C}{\partial x_{out}[i]} \frac{\partial X_{out}[i]}{\partial W[j]}, j$$

specific point where we compute gradient
$= \frac{\partial}{\partial W[j]}(x_{in}[i - \frac{k-1}{2} + 1]W[1] + x_{in}[i - \frac{k-1}{2} + 2]W[2] + ... + x_{in}[i - \frac{k-1}{2} + g]w[g], g = [0, k])$

$$= \sum_{i=0}^{k-1} \frac{\partial}{\partial w[j]}(x_{in}[i - \frac{k-1}{2} + j]W[j]) = x_{in}[i - \frac{k-1}{2} + j]$$

$$\frac{\partial C}{\partial W[j]} = \sum_{i=0}^{k-1} \frac{\partial C}{\partial x_{out}[i]} x_{in}[i - \frac{k-1}{2} + j]$$

Hence

$$W^{(i+1)} = W^{(i)} + \eta \sum_{i=0}^{k-1} \frac{\partial C}{\partial x_{out}[i]} x_{in}[i - \frac{k-1}{2} + j]$$

c. *Handling Boundaries*

Since the kernel is being slid across the image, it is likely to go out of bounds.Hence, zero pad the kernel when out of range to take care of out of bounds error for $j \notin [1 - \frac{k-1}{2}, i + \frac{k-1}{2}]$