atrous_convolution

Works by increasing the size of the filter by appending zeroes to fill the gap between the parameters

Atrous Convolution

Increase size by Dilation Rate

Dilation Rate = 1

3 x 3 Filter

Dilation Rate = 2

5 x 5 Filter

Dilation Rate = 3

7 x 7 Filter

Last pooling layer has stride = 1 instead of 2

Downsample 8x only

Downsampling by 8x to compare each pixel

Series of Atrous layers applied to capture larger context

Use Bilinear upsampling to produce result of same size

Training

Upsampling

Result Feature Map

spatial_pyramidal_pooling

Before SPP input images of multiple resolutions are supplied to FCN. Feature maps are then joined together to extract multi-scale information (Expensive computation and time)

Problem

Capture multiscale information from feature map using single input image

Input image of any size can be provided

Spatial Pyramidal Pooling

Advantage

Single Input Image

CNN Layers

Feature Maps

Spatial Pyramidal Pool

1x1

2x2

4x4

Concatenated by converting to 1D vector, Gets multi-scale information

concatenation

atrous_spp

Atrous Spatial Pyramidal Pooling

Fusing information from different scale and applying to atrous convolution. Inputs are convolved with different dilation rates and outputs are fused together

spatial_pyramidal_pooling

Concept

Input Feature Map

Convolve with

Filter Size (3x3)

Filter Size (3x3)

Filter Size (3x3)

Filter Size (3x3)

Dilation rate 6

Dilation rate 12

Dilation rate 18

Dilation rate 24

Output1

Output2

Output3

Output4

Outputs concatenated due to same size

Concatenation

Stride to keep downsample 8x

Fused Output

Convolve with 1x1 map

Upsampling

To keep Global Contextual Information

Output1

Get the required number of channels

Pass through 1x1 Convolution

Improved Results

Multitude techniques for results improvement by Google

Deeplab Neural Network

Logic

Three main improvements

Theory

Problem: Excessive downsizing due to consecutive pooling. (Downsampled by 32x)

Loss of Information

Upsampling 32x will cause

Compute and Memory expensive operation

Solution to understand large context using same number of parameters

Atrous convolution

altrous_convolution

Altrous spatial pyramidal pooling

spatial_pyriamidal_pooling

atrous_spp

Conditional Random Field