后厂理工学院
houchangtech.com

Lane Segmentation Week 5

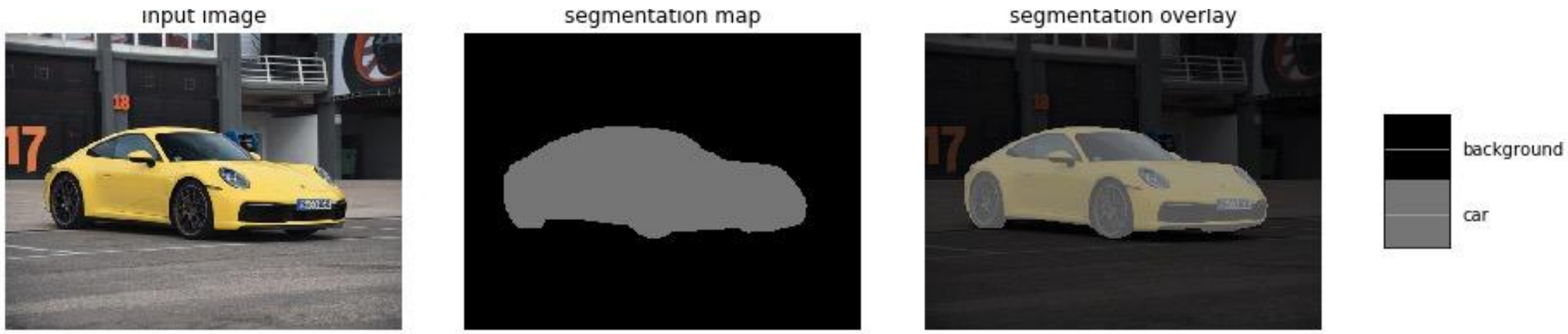
HCT CV Class

学习目标

- 掌握Deeplab v3+的原理和实现

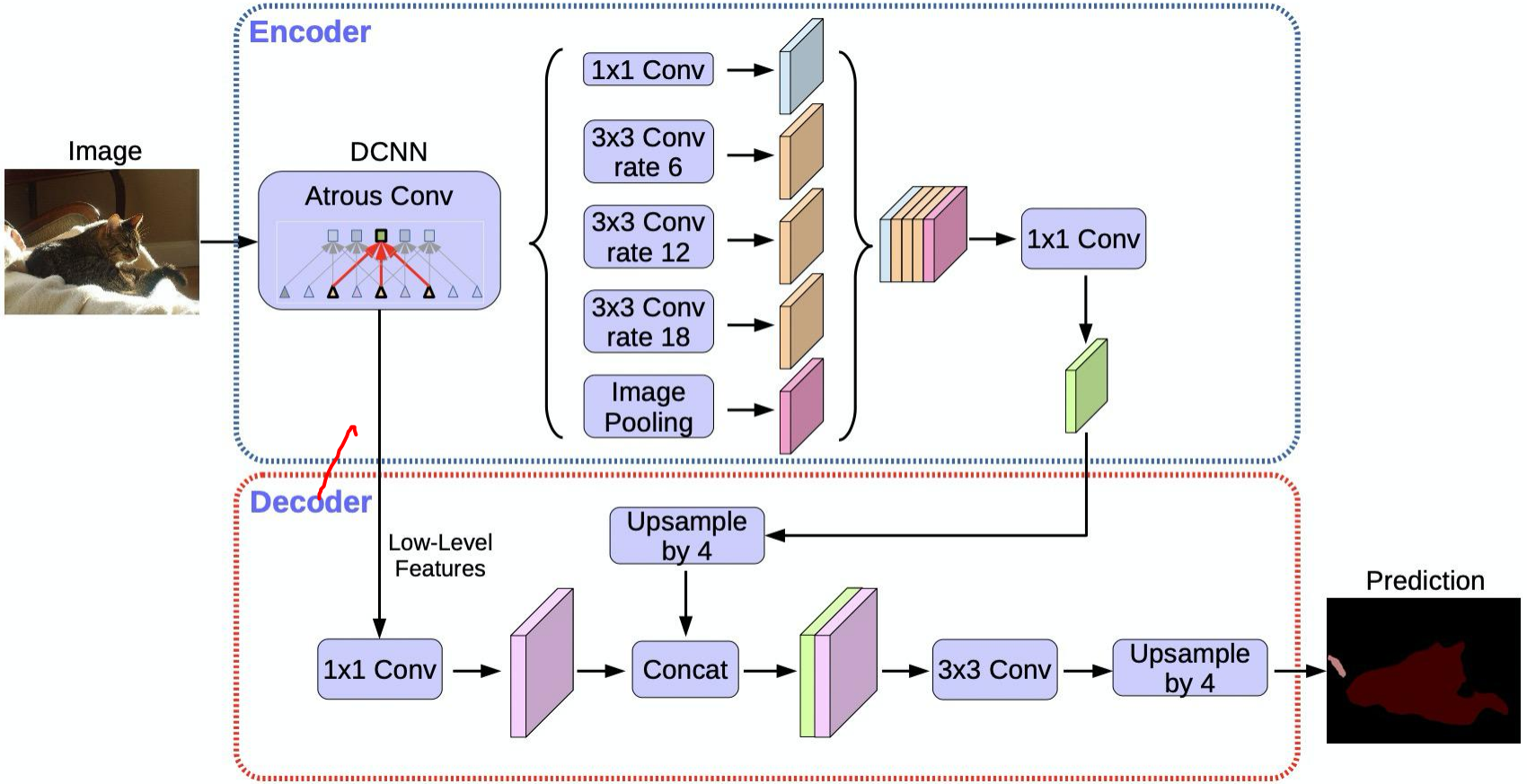
Bodypix

deeplab v3+ demo

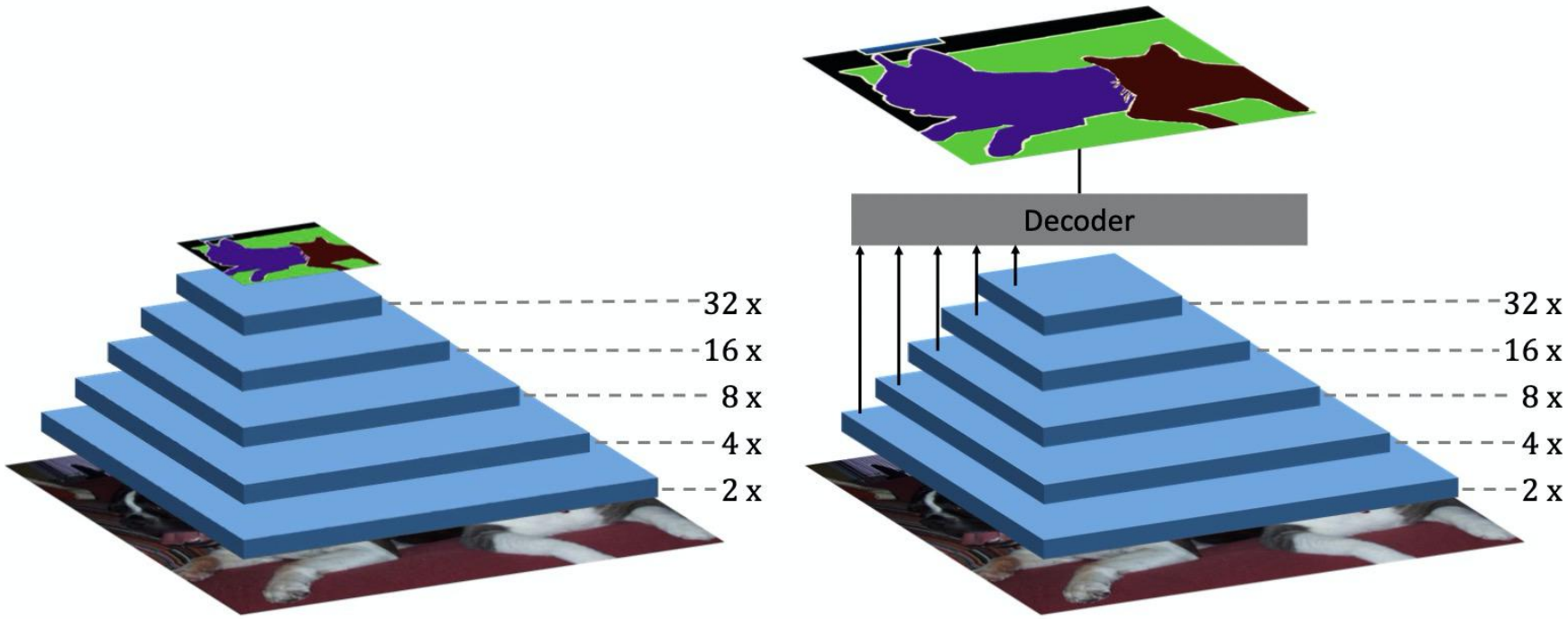


https://ai.googleblog.com/2018/03/semantic-image-segmentation-with.html

deeplab v3+



Semantic Segmentation

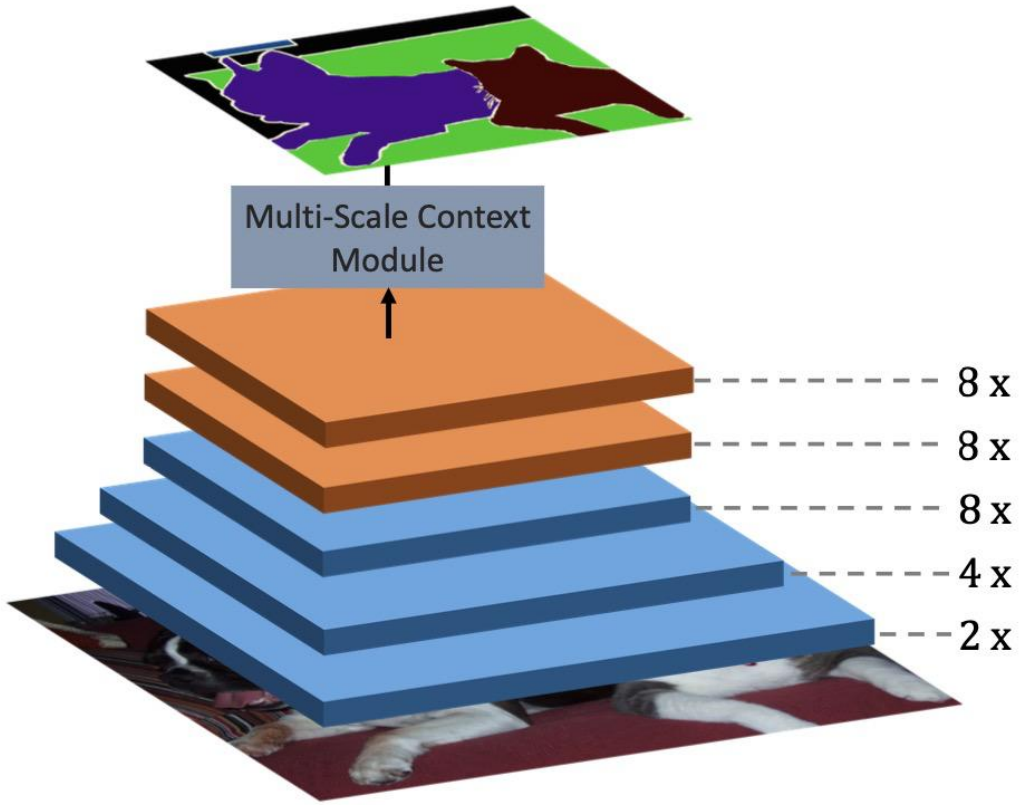


(a) FCN [22]

(b) EncoderDecoder

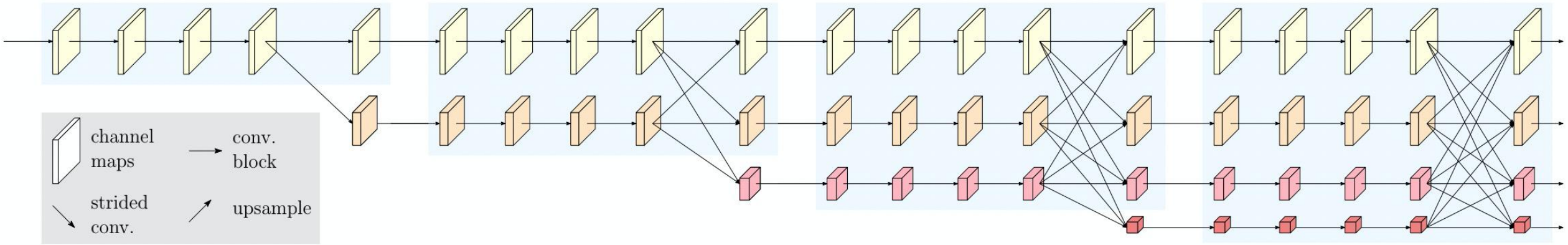
Note

DilatedFCN



Note

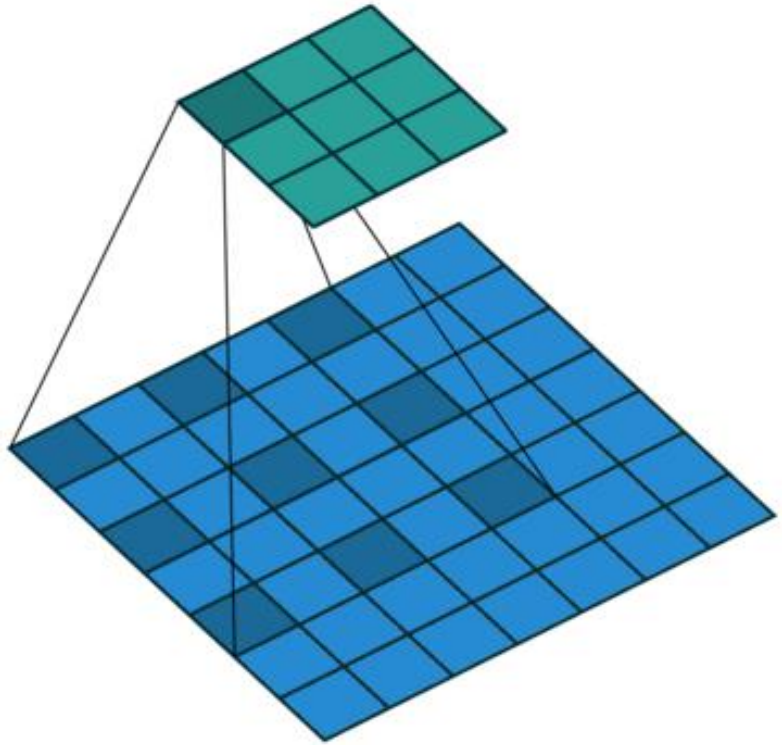
HRNet



Note

Note

Dilated Convolution



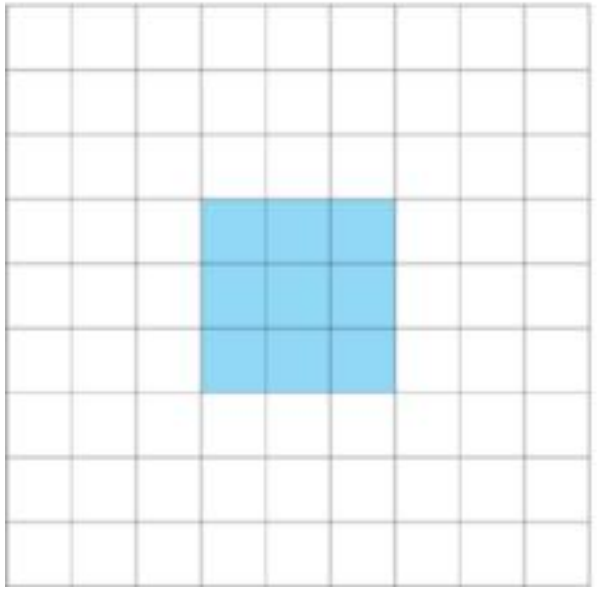
Dilated Convolution

- 在语义分隔、目标检测、生成算法中都有广泛采用

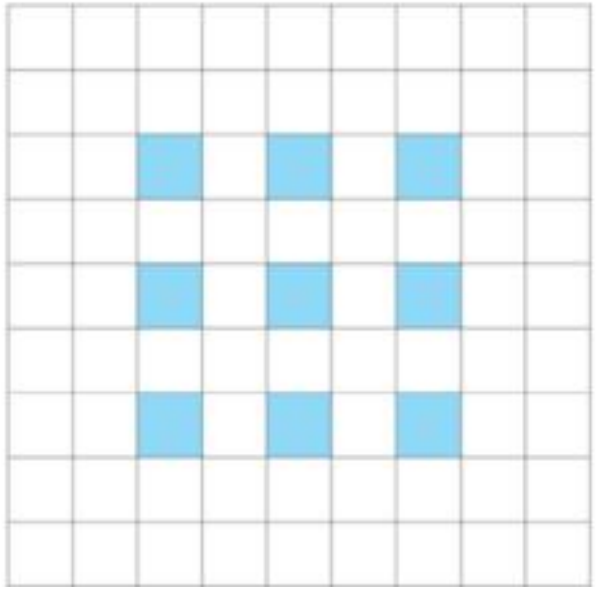
Dilated Convolution

$$y[i] = \sum_k x[i + r \cdot k] w[k]$$

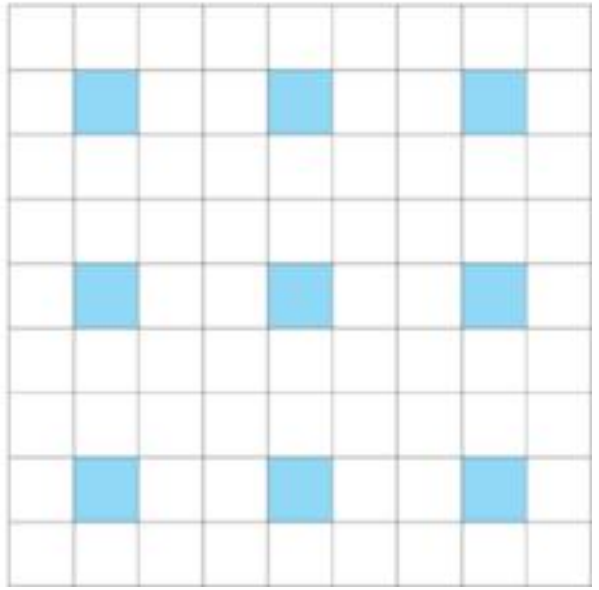
rate = 1



rate = 2



rate = 3



Dilated Convolution

- Conv2d(..., dilation=1, ...)
- 正常卷积是膨胀卷积的特殊情况: dilation=1

- Input: $(N, C_{in}, H_{in}, W_{in})$
- Output: $(N, C_{out}, H_{out}, W_{out})$ where

$$H_{out} = \left\lfloor \frac{H_{in} + 2 \times \text{padding}[0] - \text{dilation}[0] \times (\text{kernel\_size}[0] - 1) - 1}{\text{stride}[0]} + 1 \right\rfloor$$

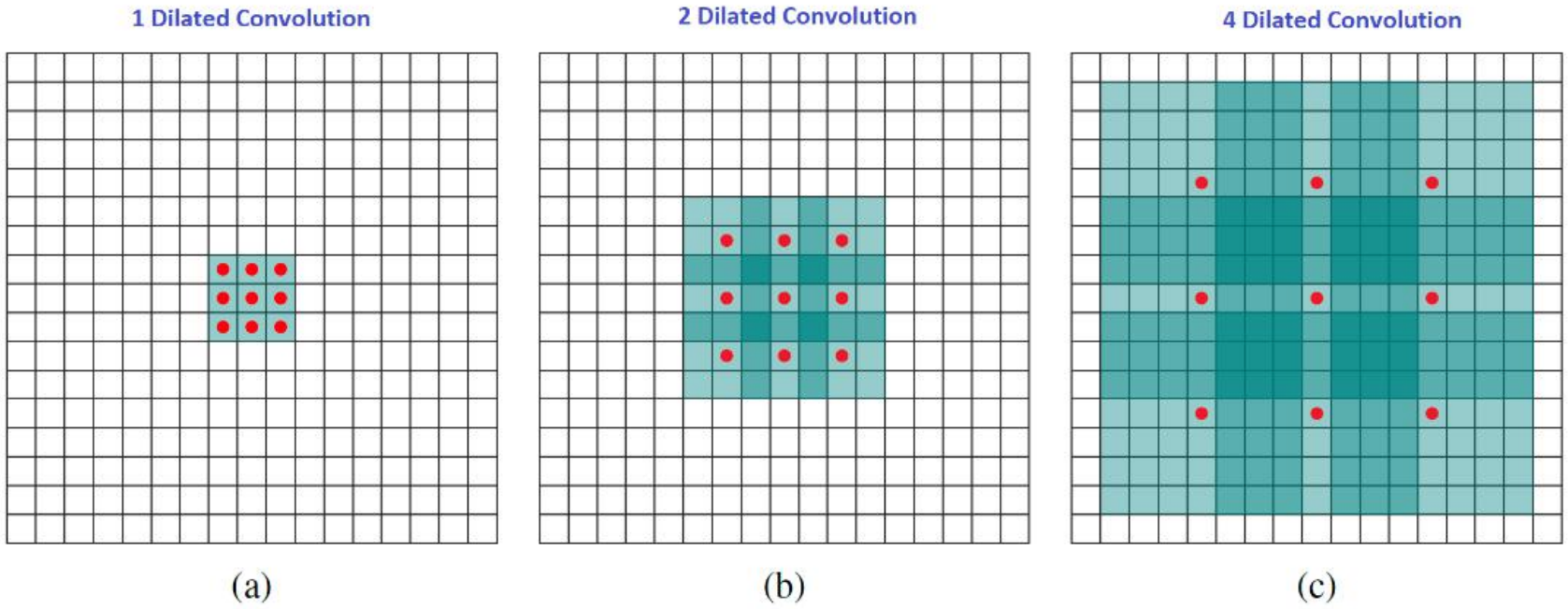
$$W_{out} = \left\lfloor \frac{W_{in} + 2 \times \text{padding}[1] - \text{dilation}[1] \times (\text{kernel\_size}[1] - 1) - 1}{\text{stride}[1]} + 1 \right\rfloor$$

Dilated Convolution

- 保持Output == Input (Spatial)

$$pad = \frac{dilation \times (kerne - 1)}{2}$$

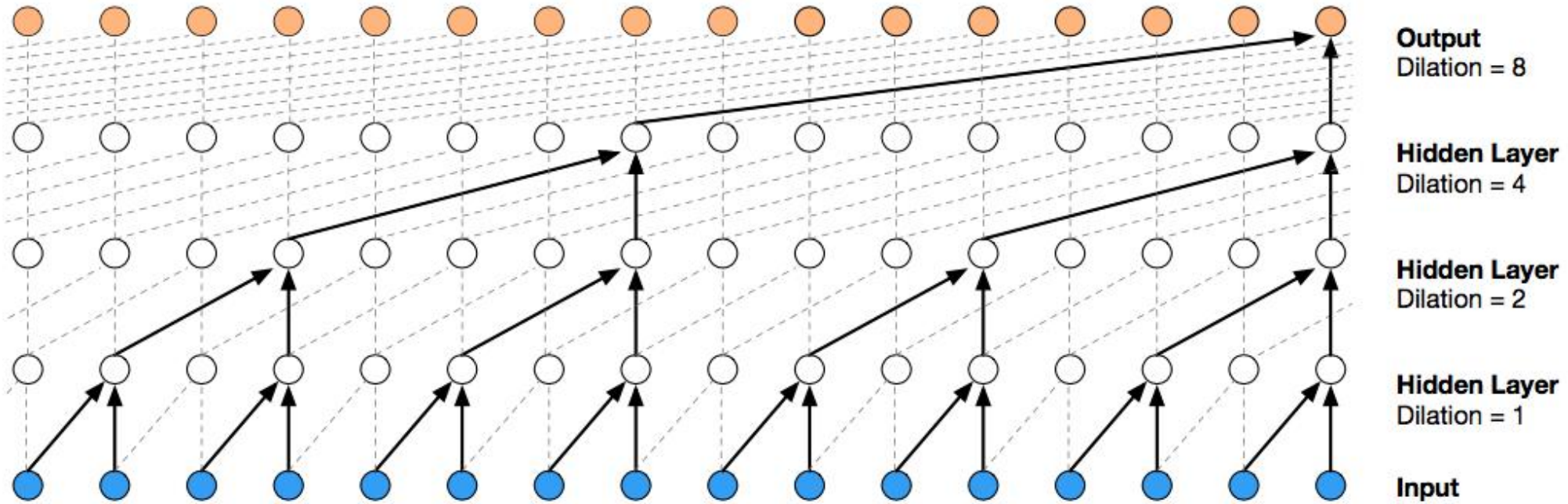
receptive field



Dilated Convolution

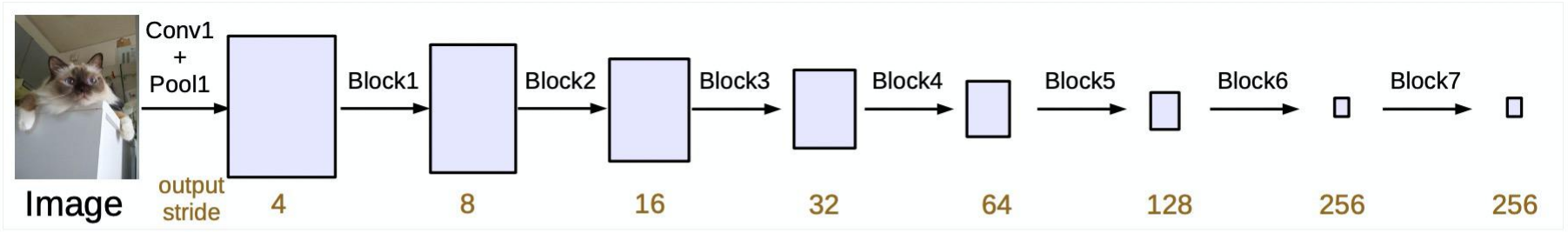
- 在filter中填充0
- 输入等间隔采样

Dilated Convolution

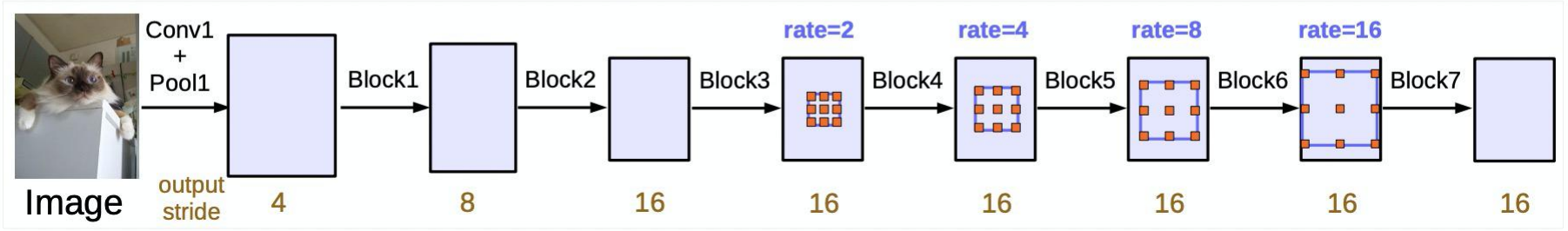


Note

DilatedFCN

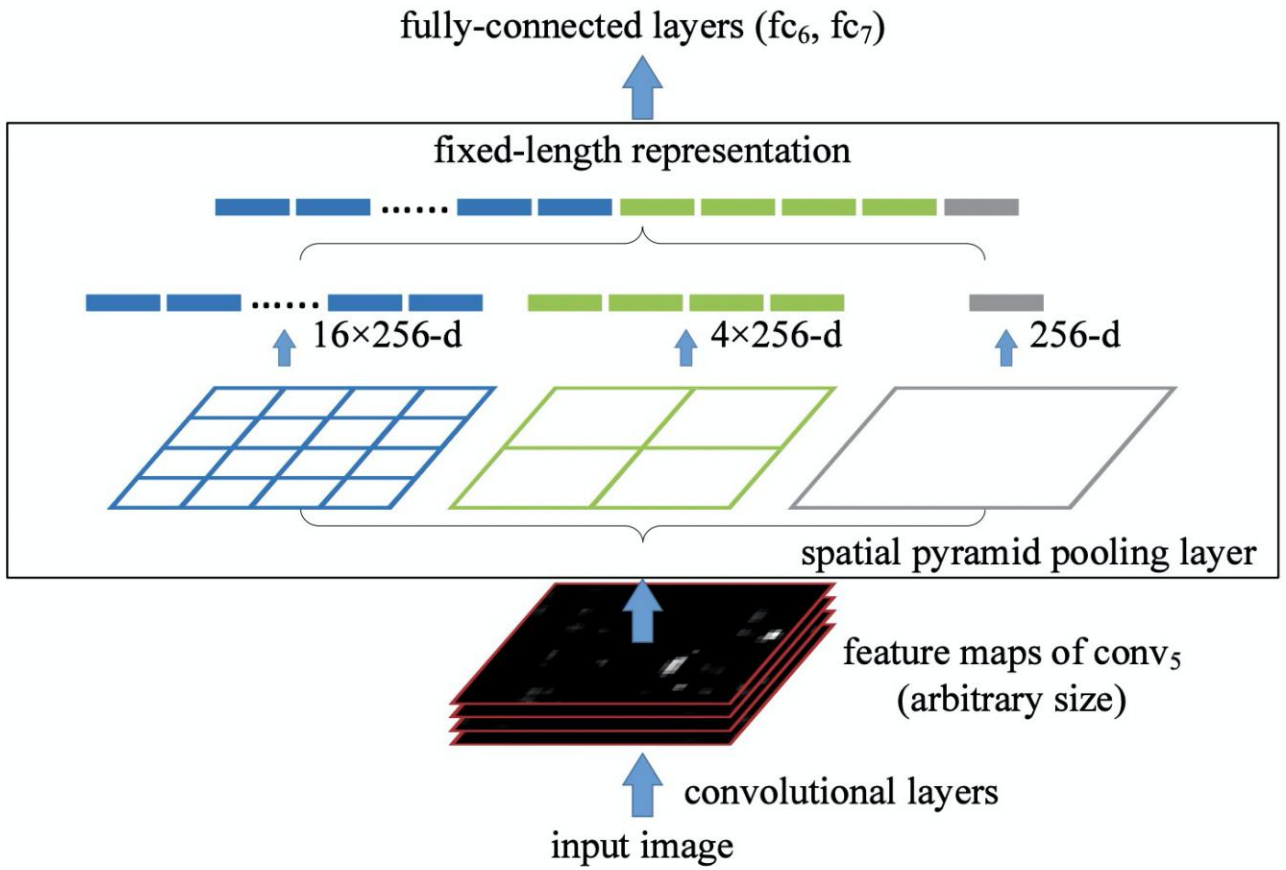


(a) Going deeper without atrous convolution.

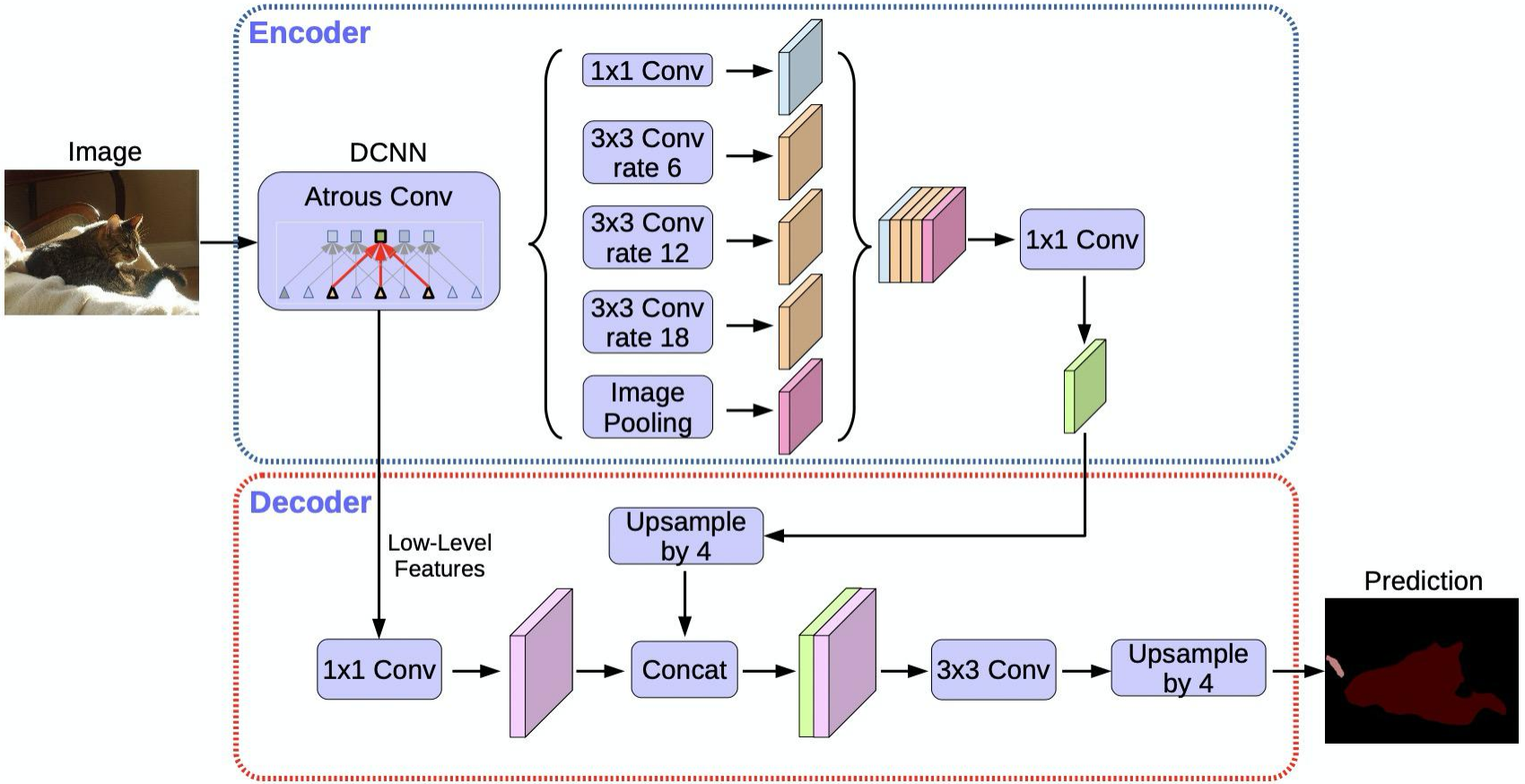


Note

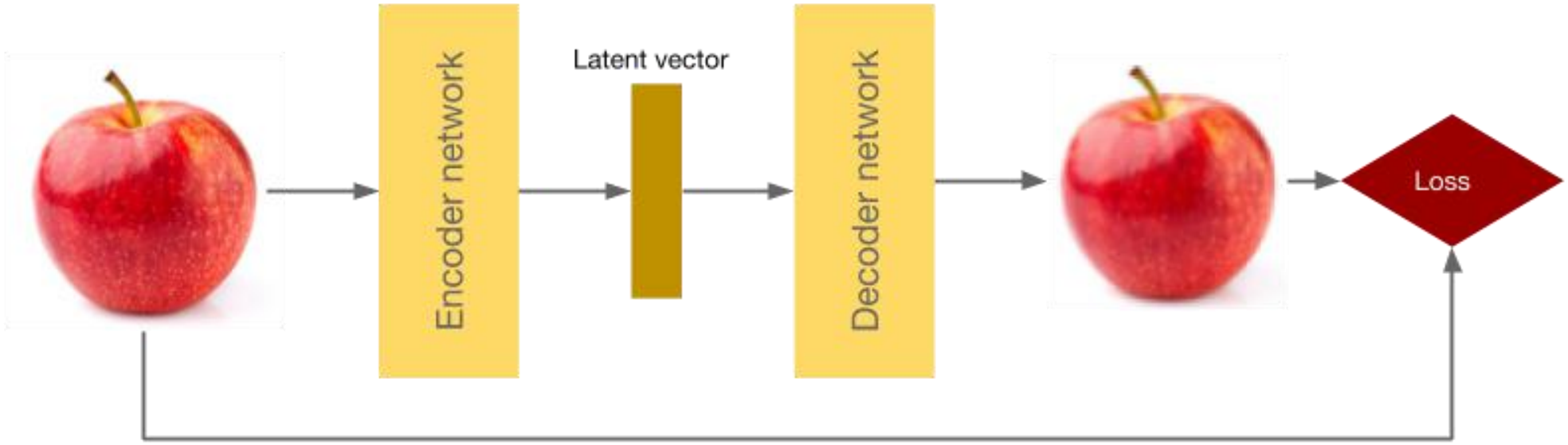
Spatial Pyramid Pooling



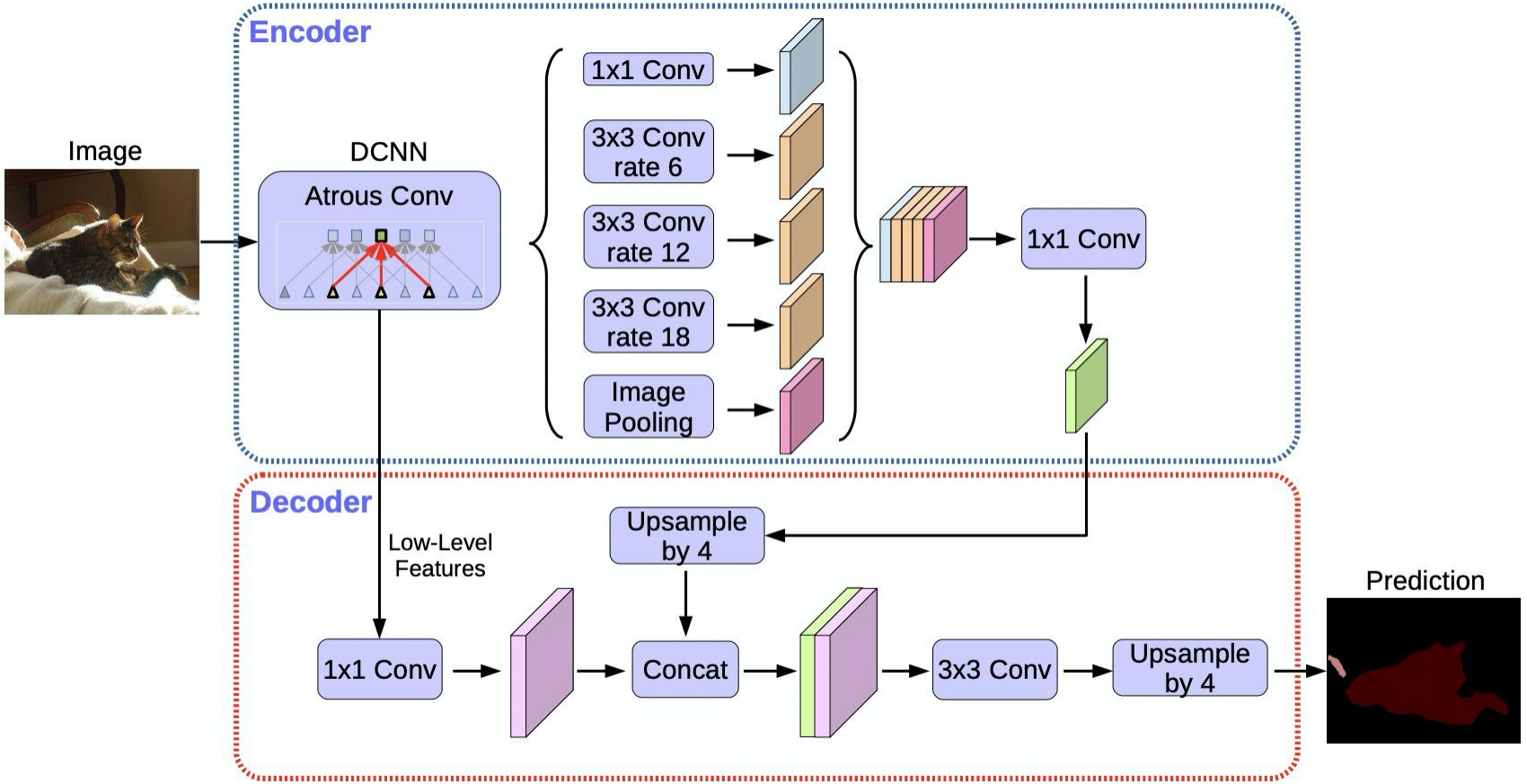
Atrous Spatial Pyramid Pooling



Encoder-Decoder

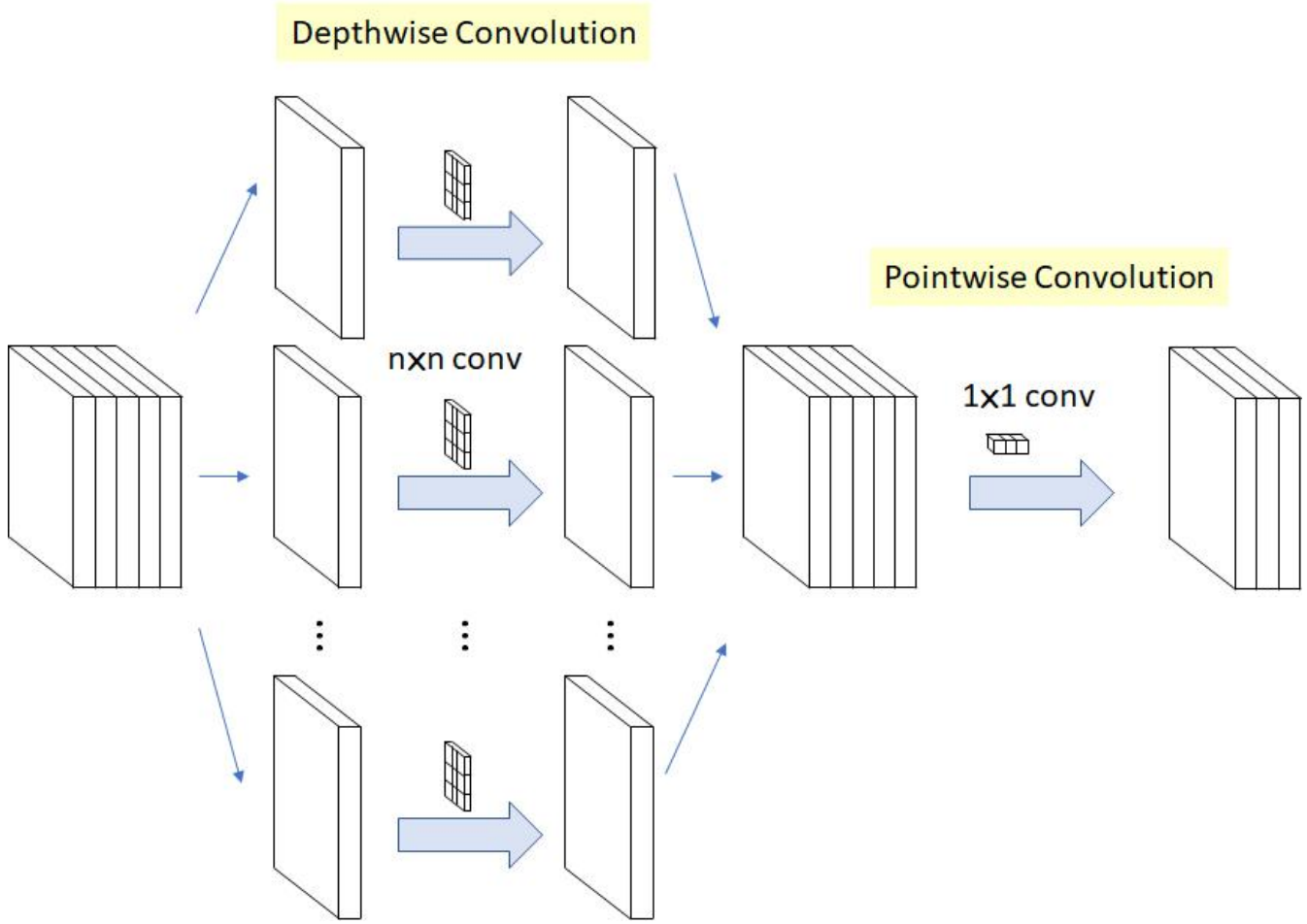


Encoder-Decoder



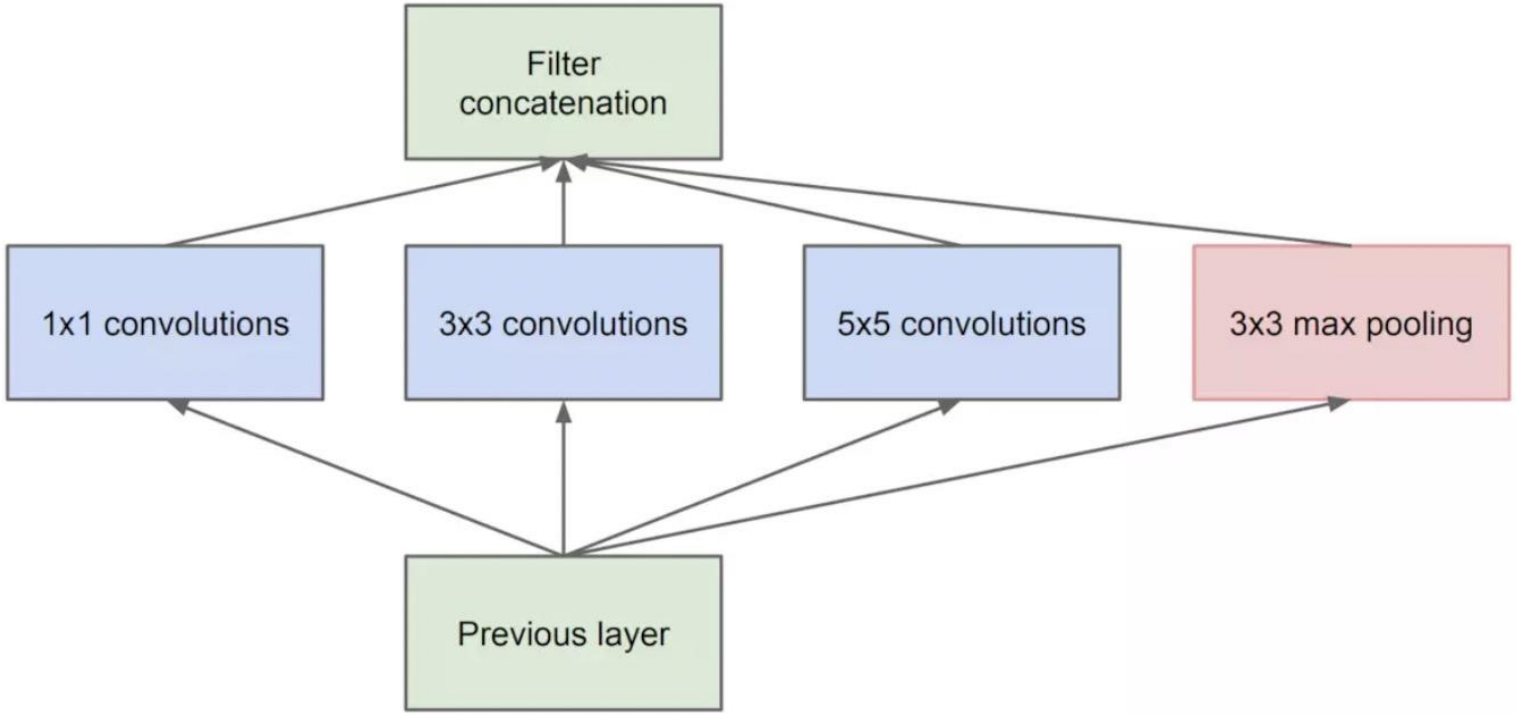
Note

Depthwise Separable Convolutions

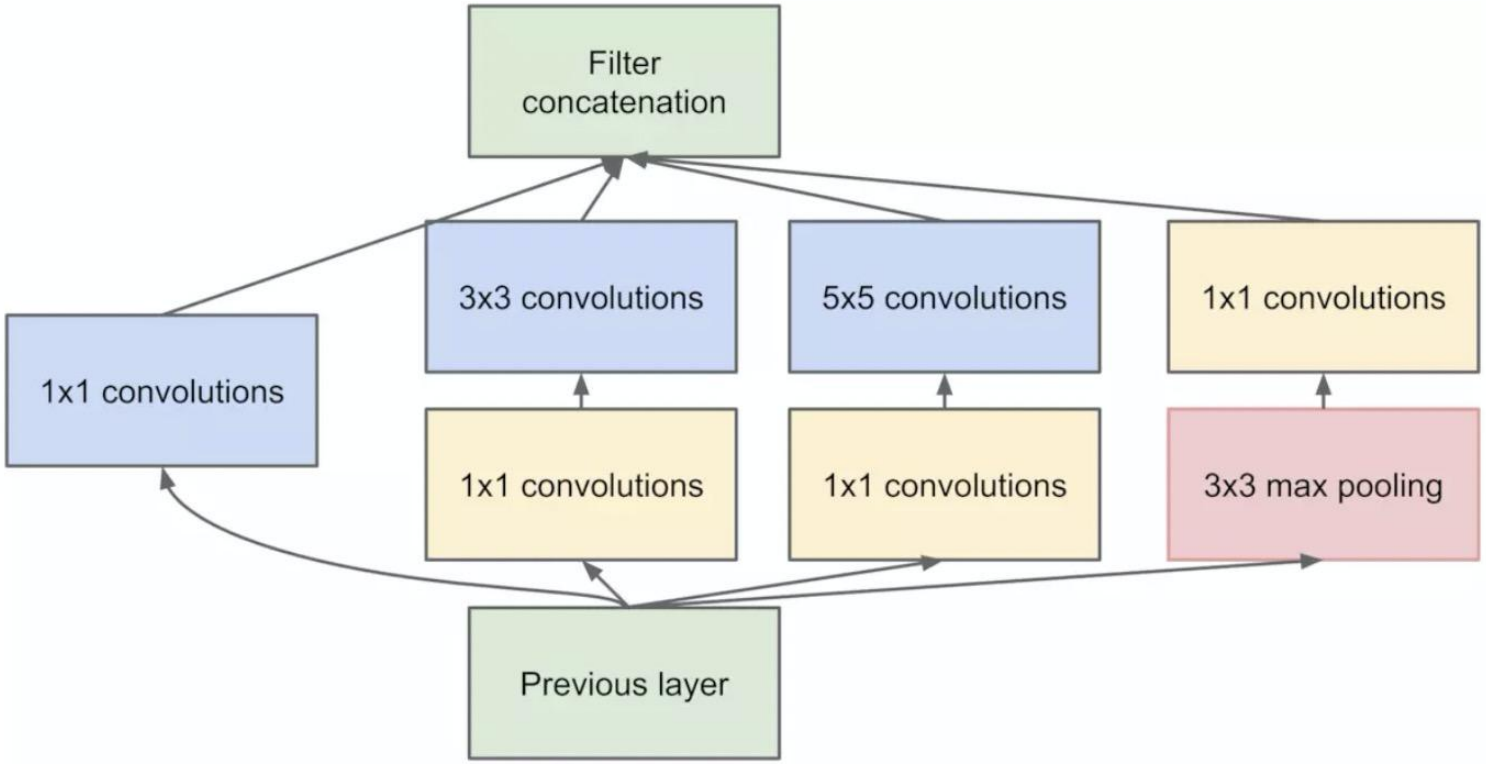


Note

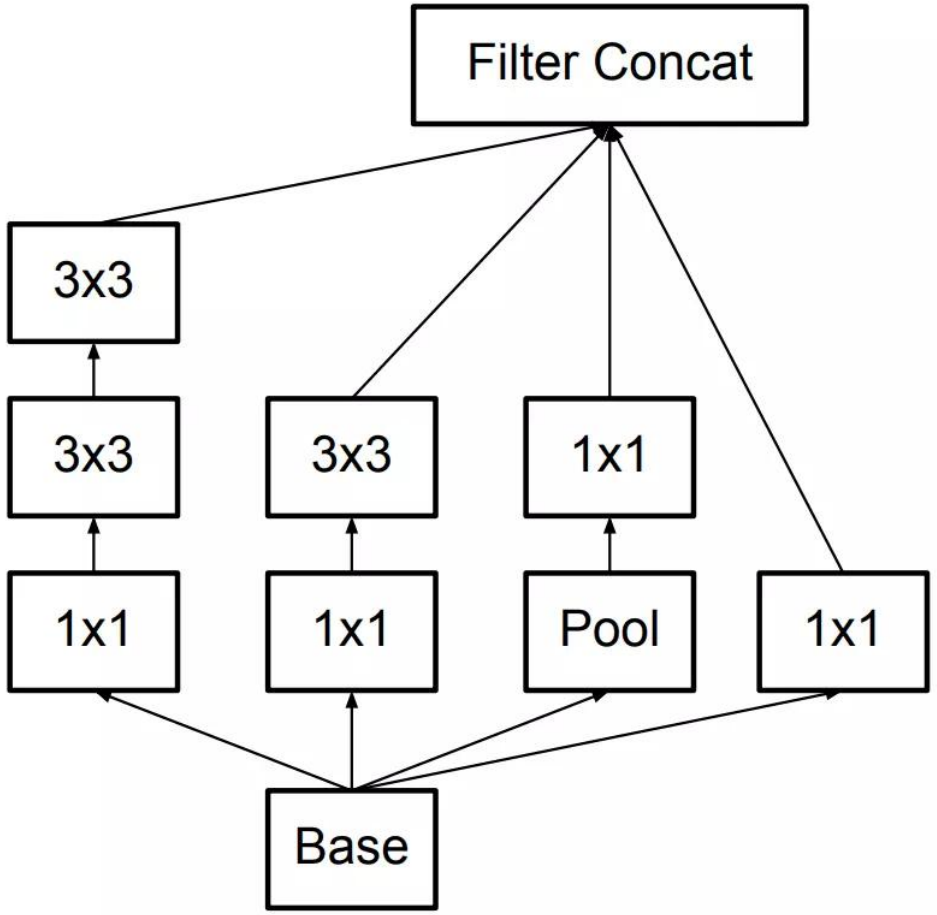
Inception



Inception



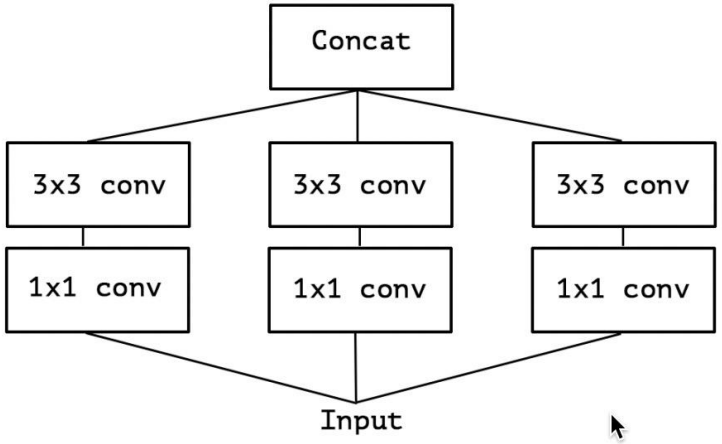
Inception

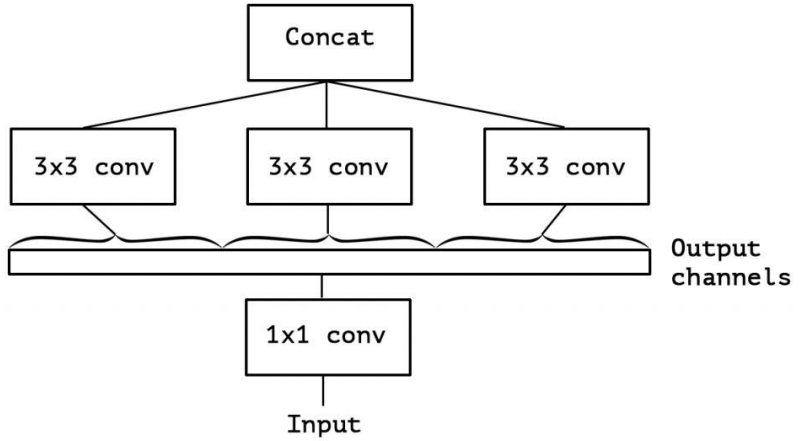


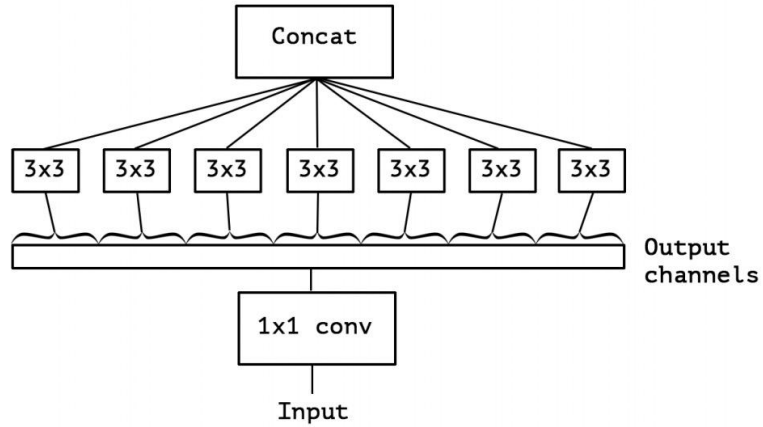
Xception(Extream Inception)

- that the map ping of cross-channels correlations and spatial correlations in the feature maps of convolutional neural networks can be entirely decoupled.

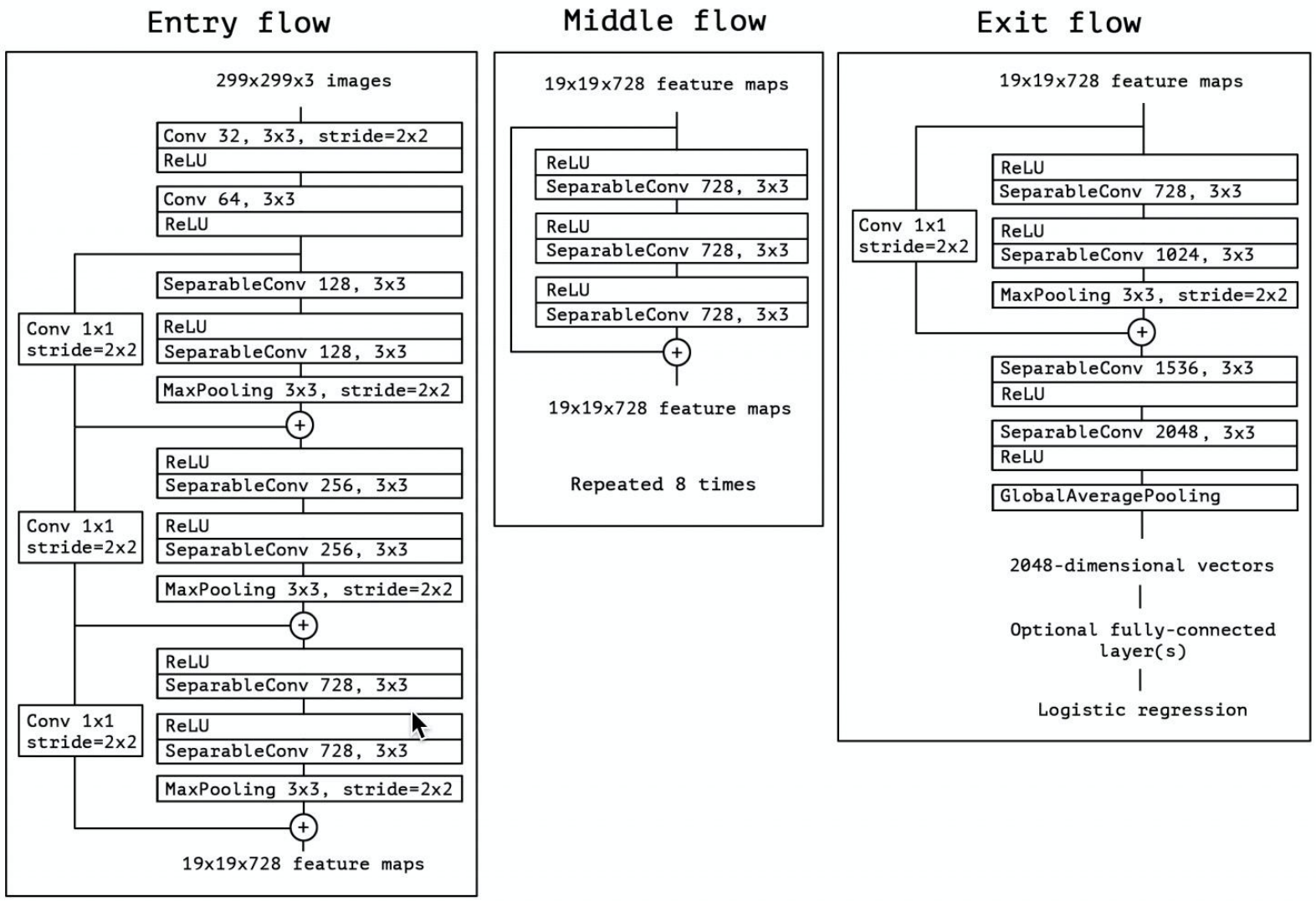
Xception(Extream Inception)



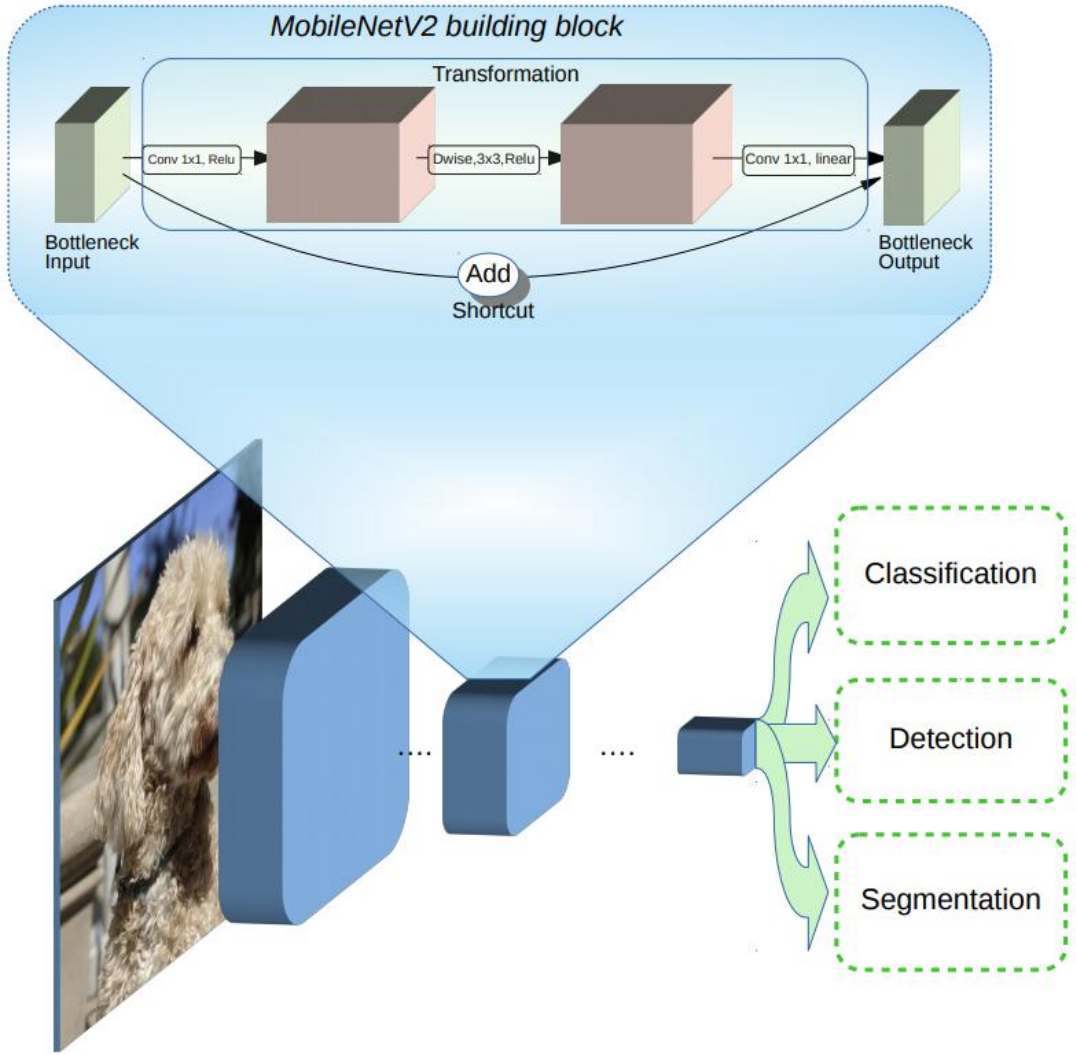




Xception Architecture

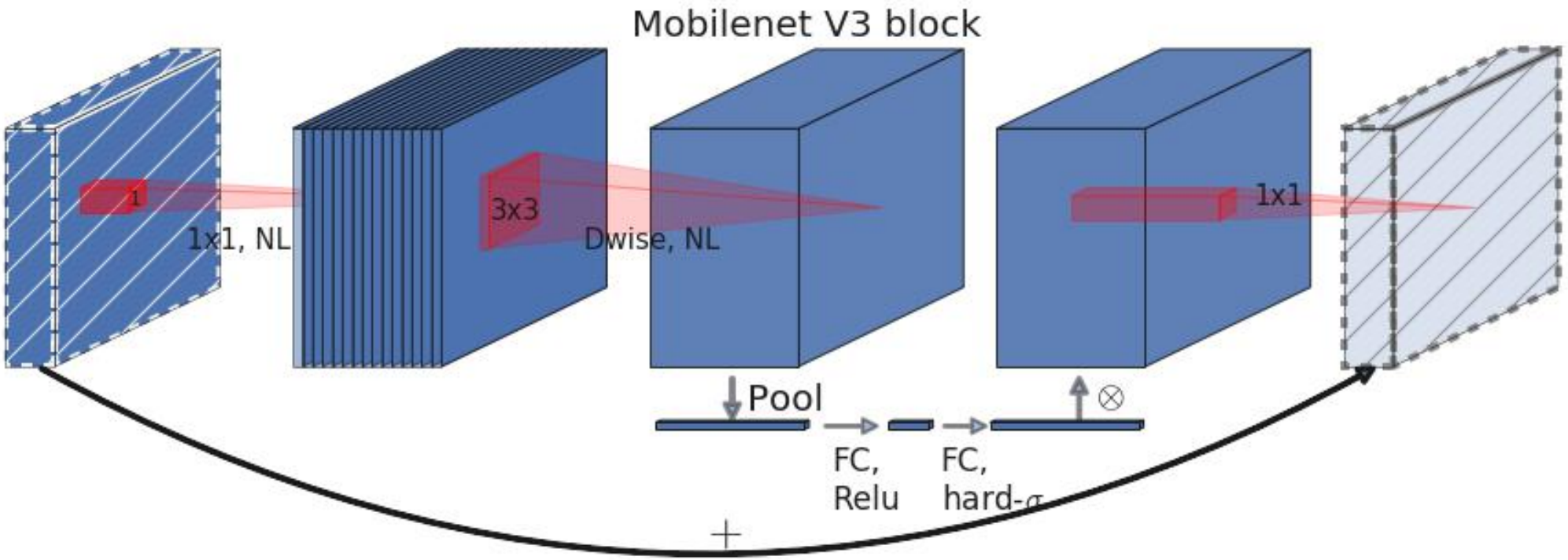


Backbone

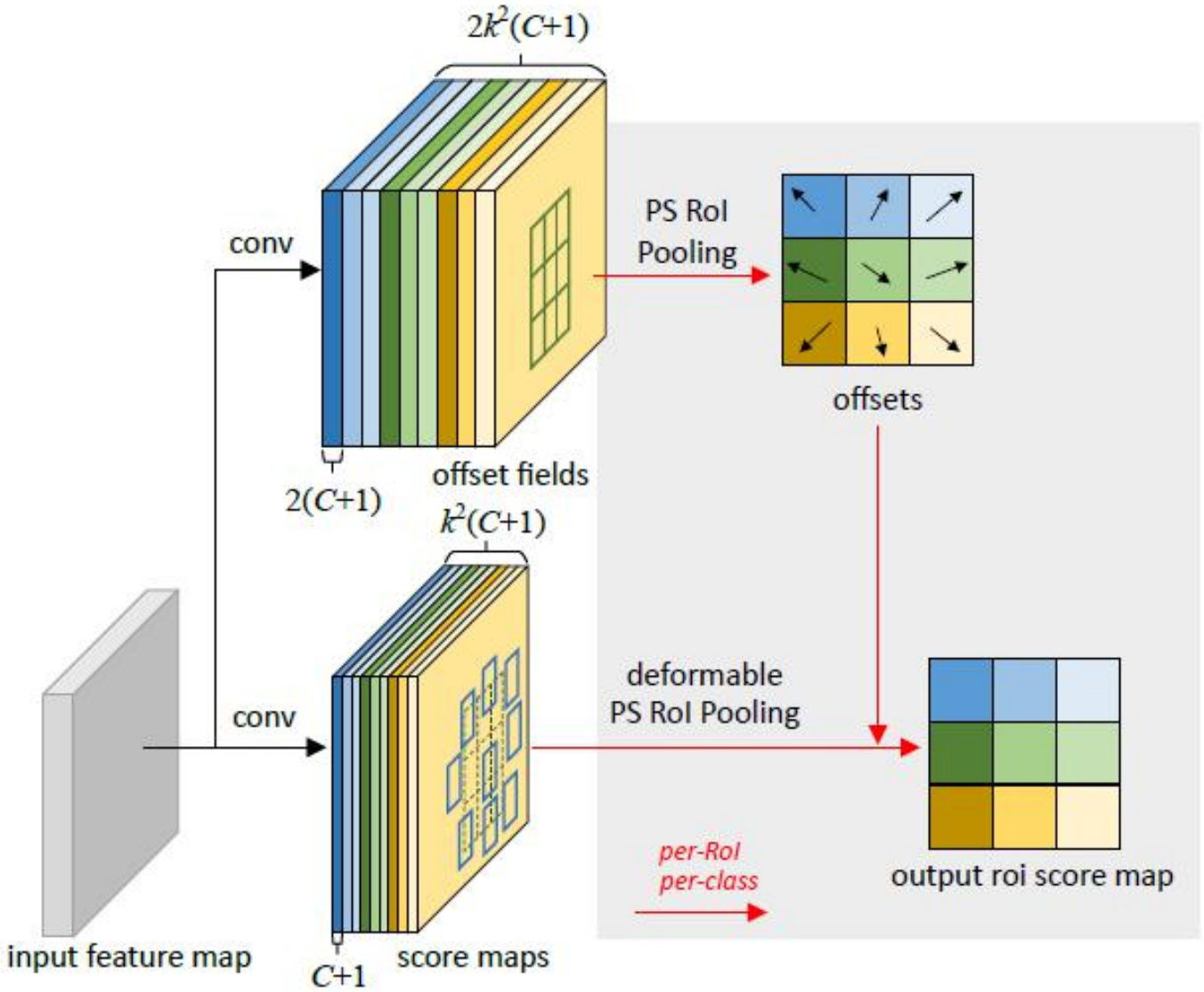


Note

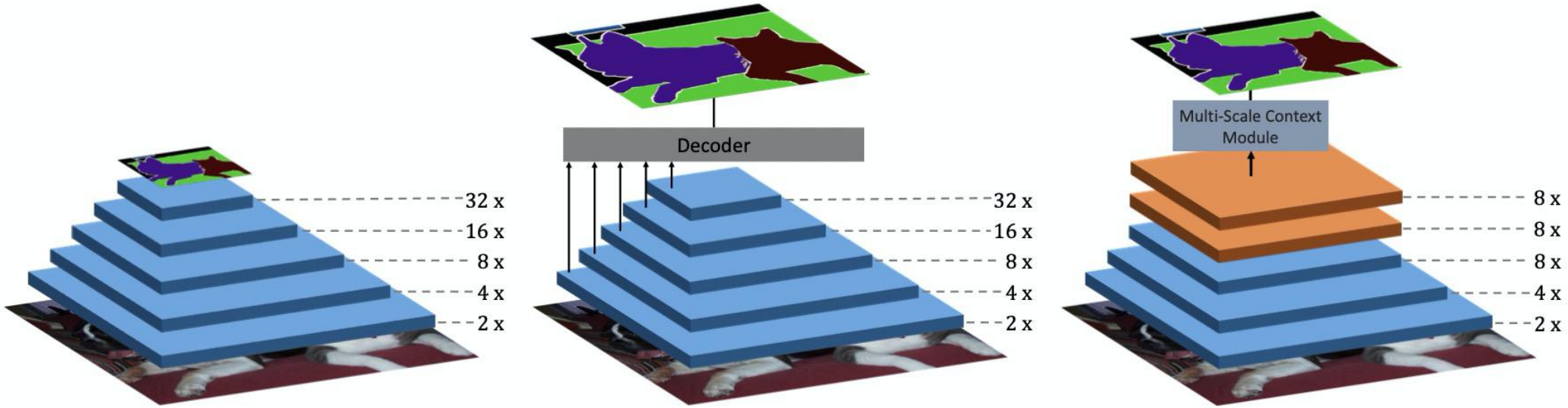
Backbone



Backbone

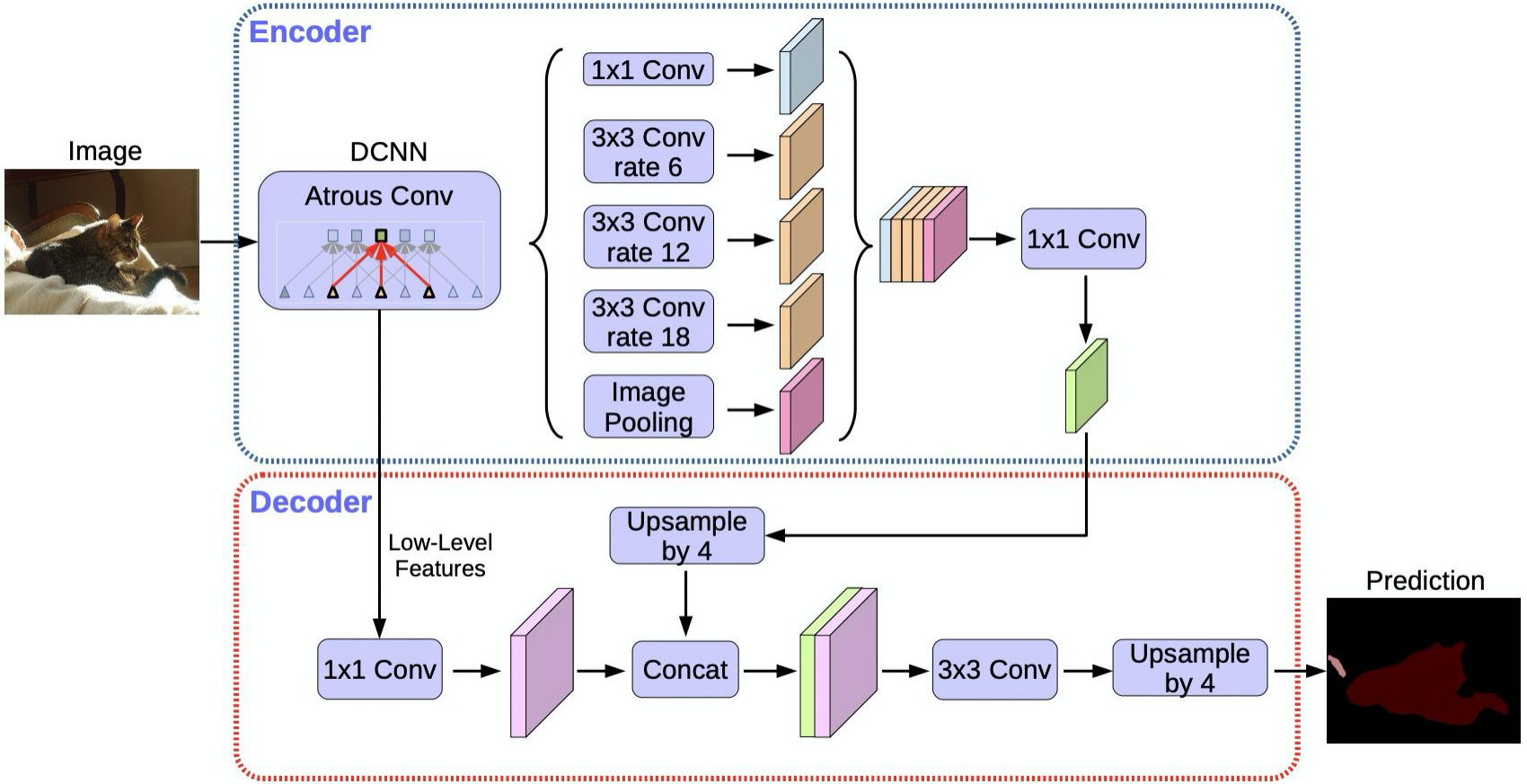

$2k^2(C+1)$
conv
PS RoI Pooling
offsets
$2(C+1)$
offset fields
$k^2(C+1)$
conv
deformable PS RoI Pooling
per-RoI per-class
output roi score map
input feature map
$C+1$
score maps

模型对比



(a) FCN [22]　(b) EncoderDecoder　(c) DilatedFCN

模型对比


Encoder
Image
DCNN
Atrous Conv
1x1 Conv
3x3 Conv rate 6
3x3 Conv rate 12
3x3 Conv rate 18
Image Pooling
1x1 Conv
Decoder
Low-Level Features
Upsample by 4
1x1 Conv
Concat
3x3 Conv
Upsample by 4
Prediction

课程总结

- Dilated Convolution
- ASPP
- Encoder-Decoder
- Xception

重难点

- Dilated Convolution
- ASPP
- Xception

课程作业

- 在Pytorch deeplab v3的基础上实现deeplab v3+

参考资料

- Deeplab v3+ Tensorflow

  https://github.com/tensorflow/models/tree/master/research/deeplab
- Deeplab v3+ PaddlePaddle

  https://github.com/gujingxiao/Lane-Segmentation-Solution-For-BaiduAI-Autonomous-Driving-Competition/blob/master/models/deeplabv3p.py
- Deeplab v3 Pytorch

  https://github.com/pytorch/vision/blob/master/torchvision/models/segmentation/deeplabv3.py

后厂理工学院

houchangtech.com

一所专注前沿互联网技术领域的创新实战大学