DEEPACTION

长按识别上方二维码关注

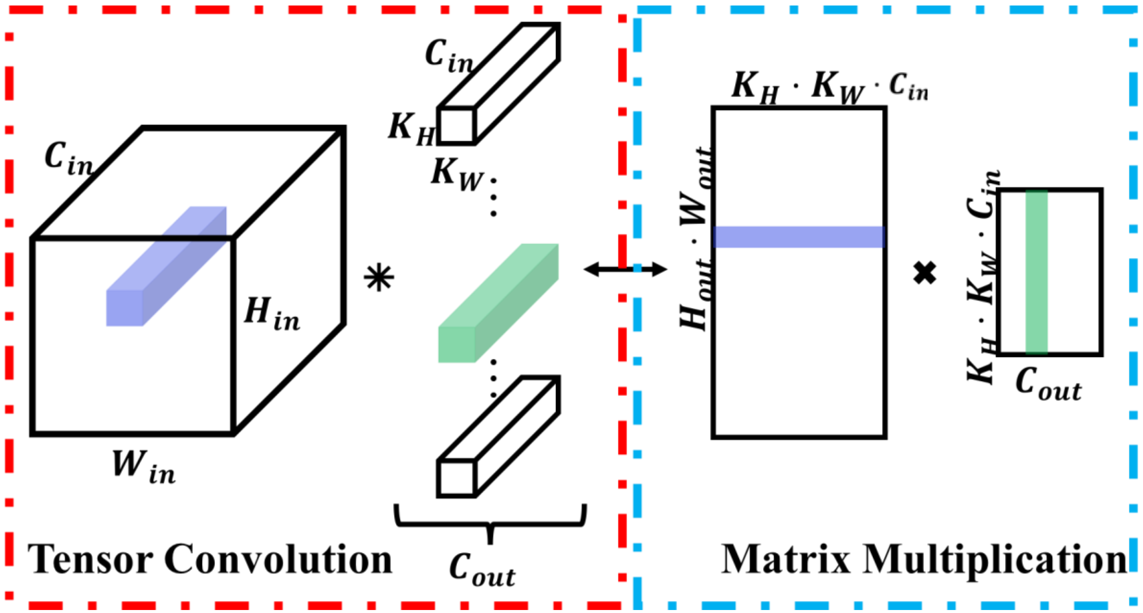
CNN基本结构

CNN基本单元

- 卷积层
- 池化层
- 激活层
- BN层
- 全联接层

卷积

- 卷积的基本参数
  - 卷积核　　滑动步长　padding
  - 输入通道　输出通道
- 卷积计算过程
  - 直接计算
  - 将卷积转化为矩阵乘法　（im2col）
- 卷积的好处
  - 权重共享



im2col

预先确定的基本概念：
Filter
Kernel

Im2col Example

Kernel
$K_H = 2$
$K_W = 2$

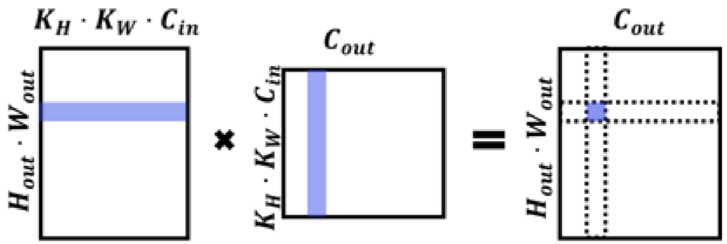
W=1 C=1 H=1

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

17	18	19	20
21	22	23	24
25	26	27	28
29	30	31	32

C=2

Tensor$(N * H * W * C)$



Im2col

- 假设:
  Tensor 1*4*4*2
  Kernel 2*2
  stride 2
- 存储Tensor时，先存储最后一维，然后存储倒数第二维，依此类推

Matrix$(H_{out}W_{out} * K_H K_W C)$

1	17	2	18	5	21	6	22
3	19	4	20	7	23	8	24
9	25	10	26	13	29	14	30
11	27	12	28	15	31	16	32

新型卷积

- Group conv
- Depthwise conv

- Deconv
- Dilated conv

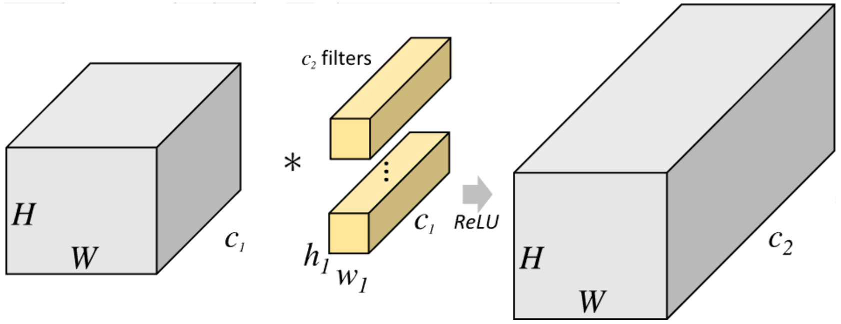
- Deformable conv

[卷积操作动画 https://github.com/vdumoulin/conv_arithmetic](https://github.com/vdumoulin/conv_arithmetic)

Group Convolution

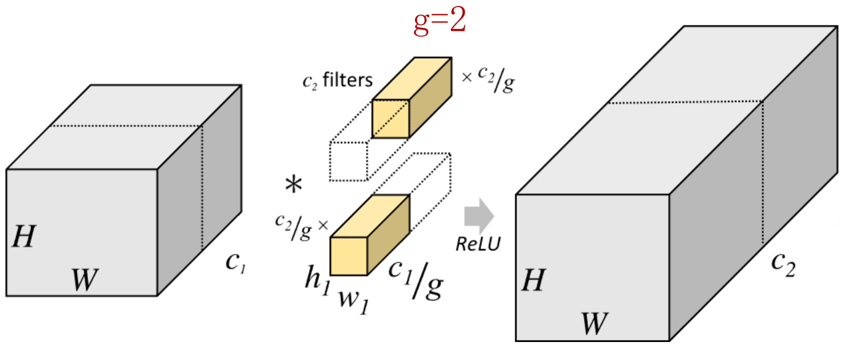
1. 输入通道c_1可以分成 g 组，每组有 c_1/g 个输入特征图
2. 对于每组有 c_2/g 个filter， 可以得到 c_2/g 个输出特征图
3. 最后将每组的输出特征图串联起来，得到 c_2 个输出特征图

➢ Original Convolution



$$P_{conv} = c_1 \cdot c_2 \cdot h_l \cdot w_l$$
$$\#_{conv} = H \cdot W \cdot c_2 \cdot h_l \cdot w_l \cdot c_1$$

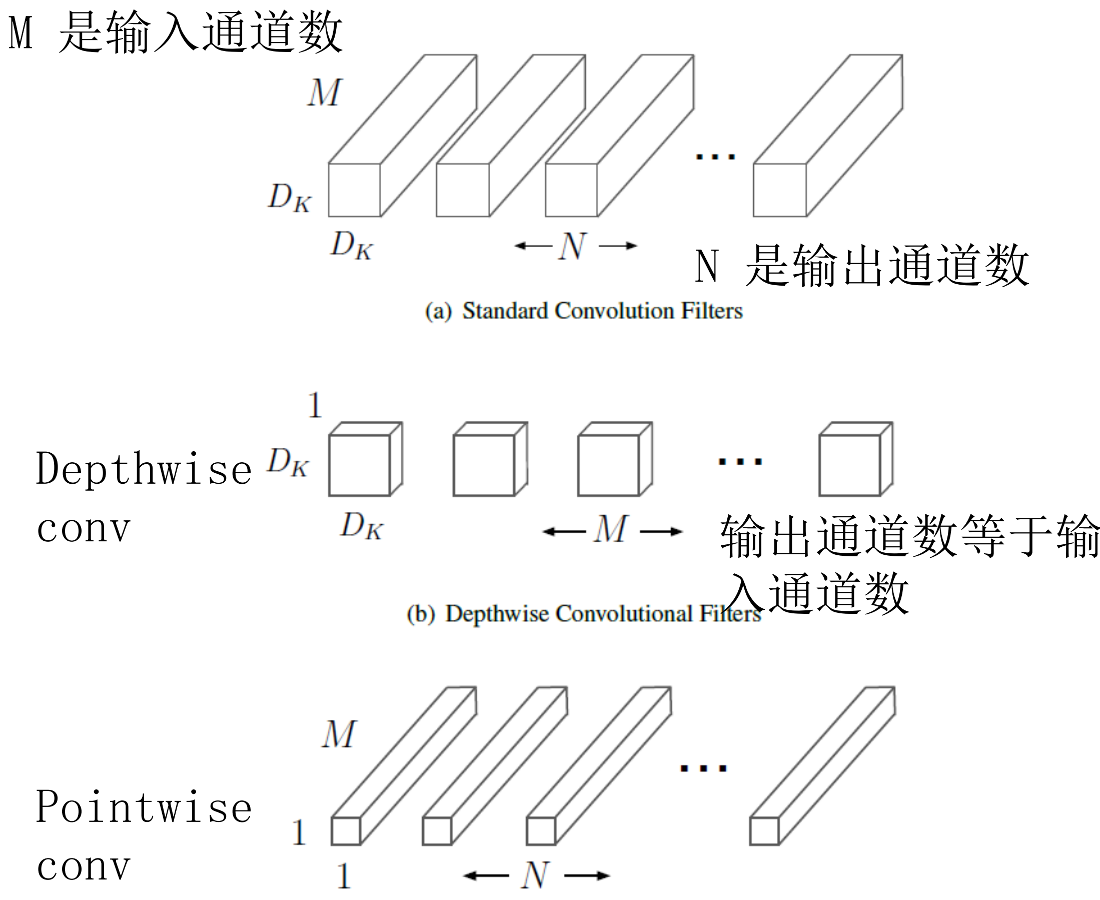
➢ Group Convolution



$$P_{conv} = (c_1/g) \cdot (c_2/g) \cdot K_H \cdot K_W \cdot g$$
$$\#_{conv} = H \cdot W \cdot (c_2/g) \cdot h_l \cdot w_l \cdot (c_1/g) \cdot g$$

Depthwise Conv

- Group conv的极限版
- MobileNet的基本结构
- 专职做空间维度的信息交流，与Pointwise Conv配套使用



(a) Standard Convolution Filters

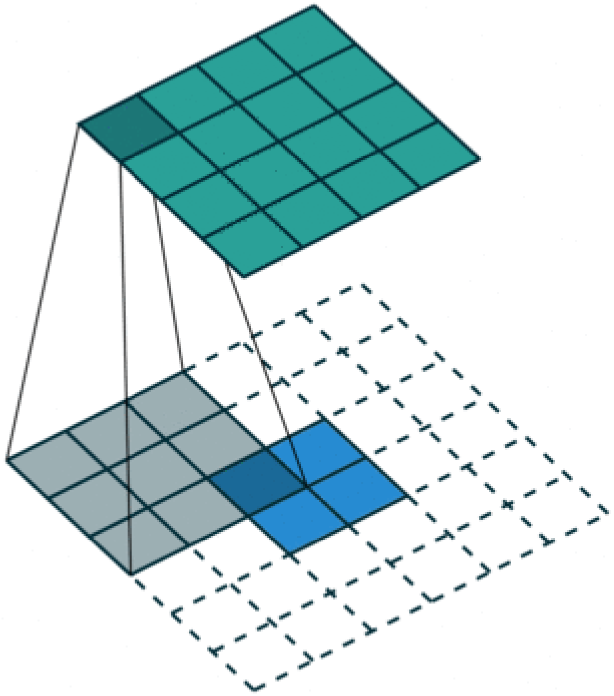
(b) Depthwise Convolutional Filters

(c) 1×1 Convolutional Filters called Pointwise Convolution in the context of Depthwise Separable Convolution

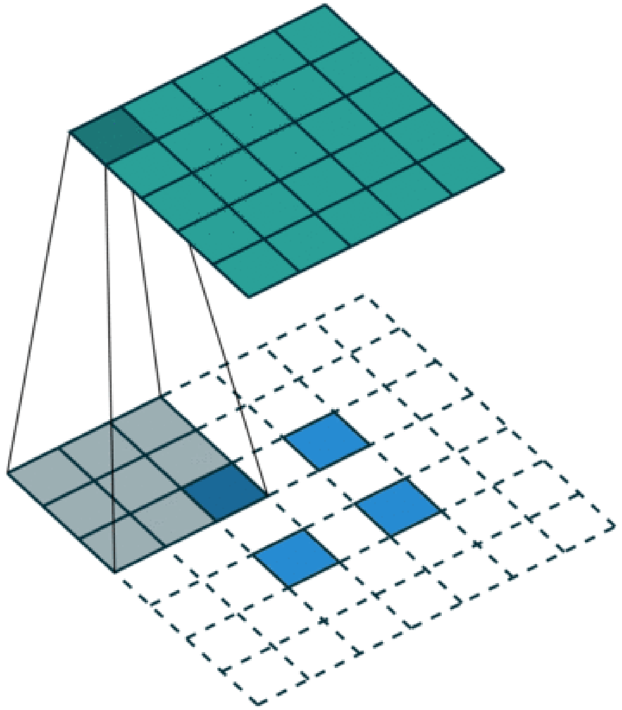
Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv:1704.04861, 2017.

Deconv

- 目的：上采样



Stride=1

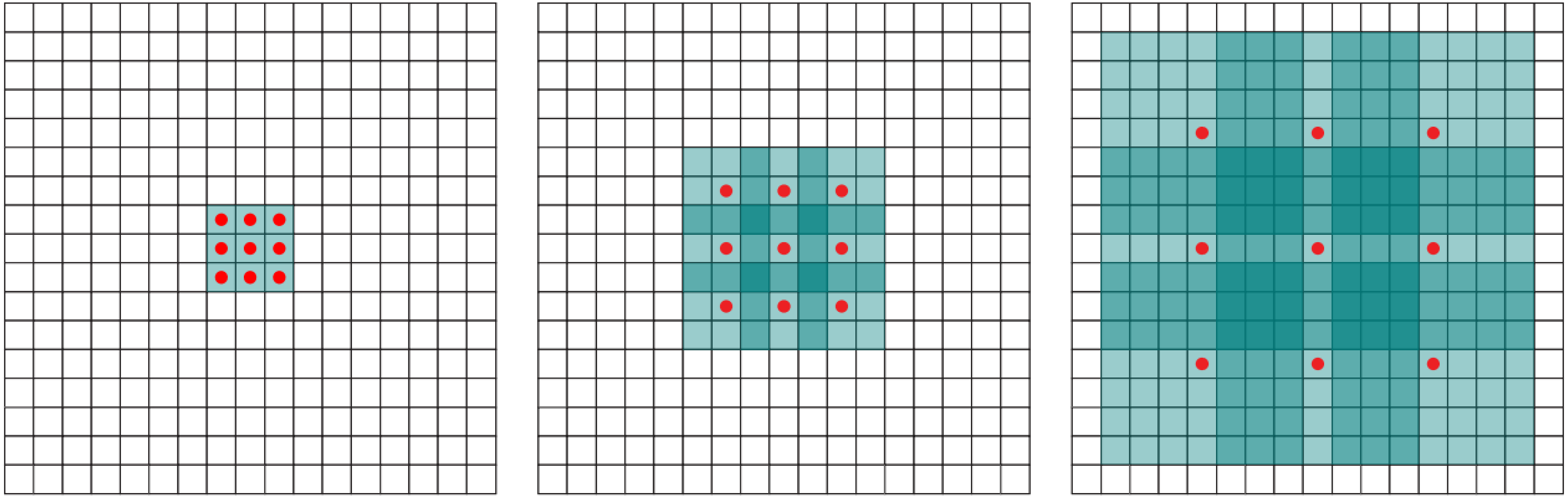


Stride=2

https://github.com/vdumoulin/conv_arithmeti

Dilated Conv

- 目的：增大感受野
- 用途：图像分割



(a) (b) (c)

Deformable Conv

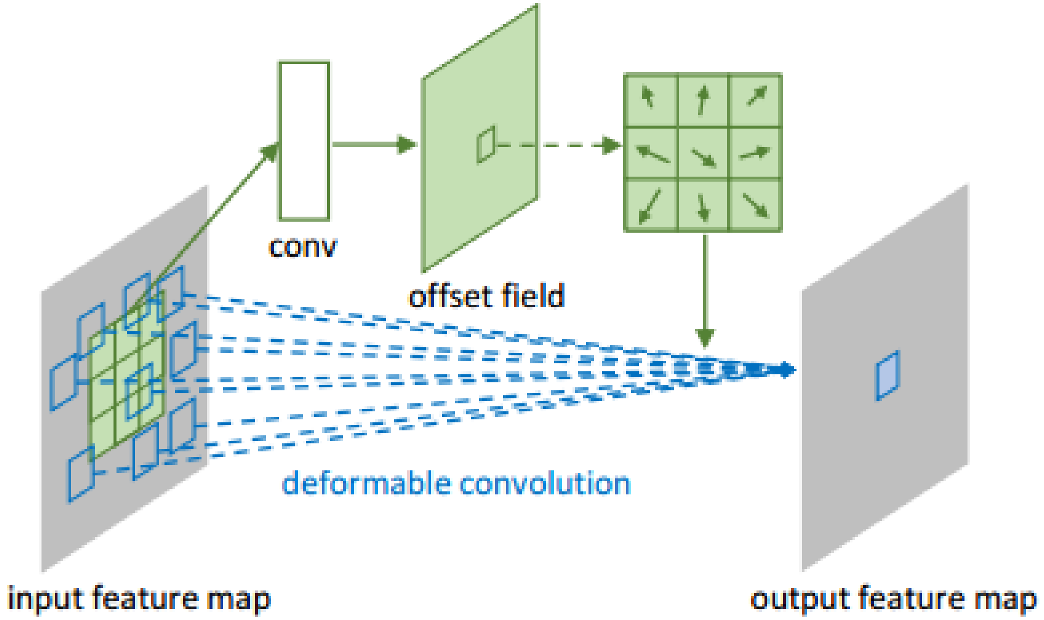
普通卷积的位置偏移

$$\mathcal{R} = \{(-1,-1),(-1,0),\ldots,(0,1),(1,1)\}$$

$$\mathbf{y}(\mathbf{p}_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} \mathbf{w}(\mathbf{p}_n) \cdot \mathbf{x}(\mathbf{p}_0 + \mathbf{p}_n).$$

可变卷积的位置偏移

$$\mathbf{y}(\mathbf{p}_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} \mathbf{w}(\mathbf{p}_n) \cdot \mathbf{x}(\mathbf{p}_0 + \mathbf{p}_n + \Delta \mathbf{p}_n).$$



Batch Normalization

- 作用：归一化

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{Var[x^{(k)}]}}$$

$$y^{(k)} = \gamma^{(k)}\hat{x}^{(k)} + \beta^{(k)}$$

激活函数

- Sigmoid $$f(z)=\frac{1}{1+e^{-z}}$$
- Tanh $$f(z)=\frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$$
- Relu $$f(z)=\begin{cases}0, & z<0\\ z, & z\geq 0\end{cases}$$
    - Leaky Relu $$f(z)=\begin{cases}az, & z<0\\ z, & z\geq 0\end{cases}$$
    - ELU $$f(z)=\begin{cases}a(e^{z}-1), & z<0\\ z, & z\geq 0\end{cases}$$

CNN基本单元

- 卷积层
- 池化层
- 激活层
- BN层
- 全联接层

CNN基本模块