

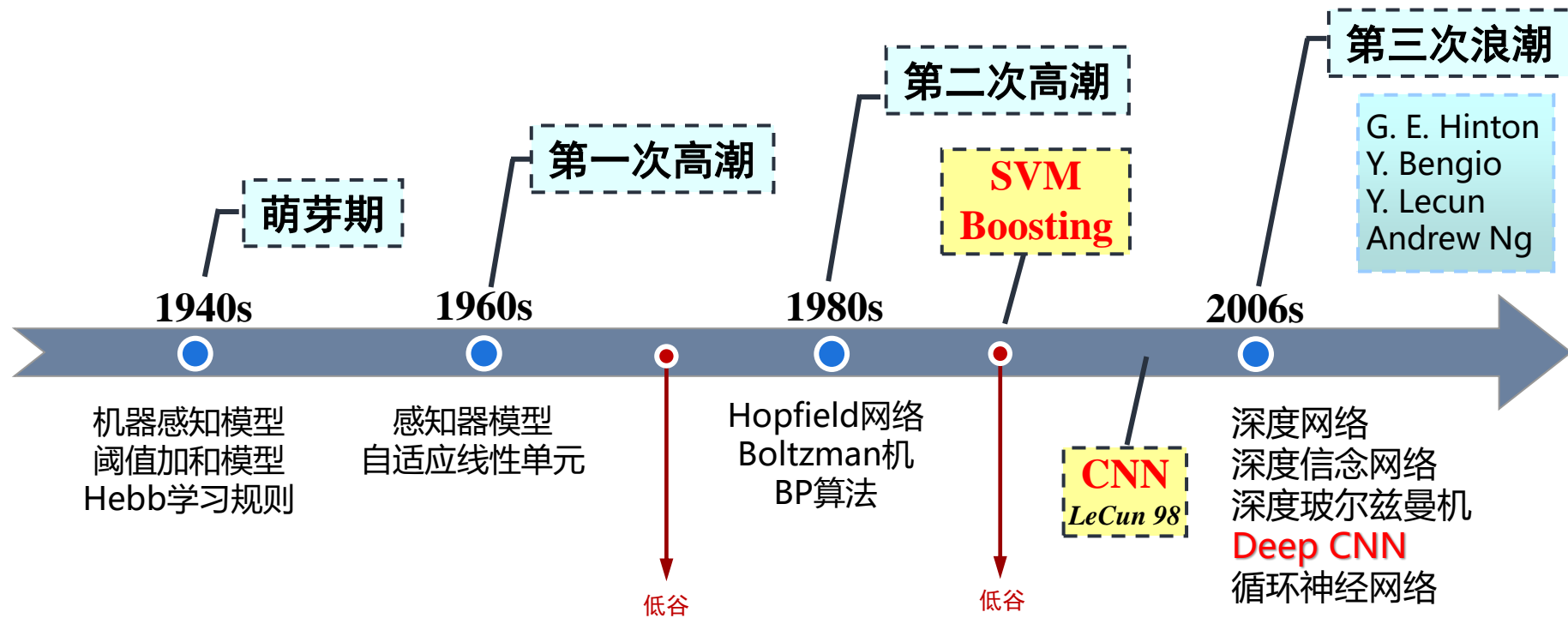
# 卷积神经网络：从欧氏空间到非欧氏空间

---

报告人：方 深 博士在读  
中科院自动化研究所

- 人工神经网络发展浪潮
- 卷积计算与卷积神经网络结构
- 卷积神经网络拓展至非欧空间
- 本课程内容设置

# 人工神经网络发展浪潮



## 第三次浪潮——卷积神经网络

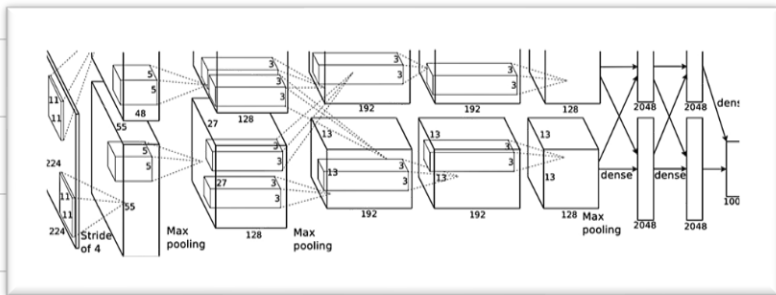
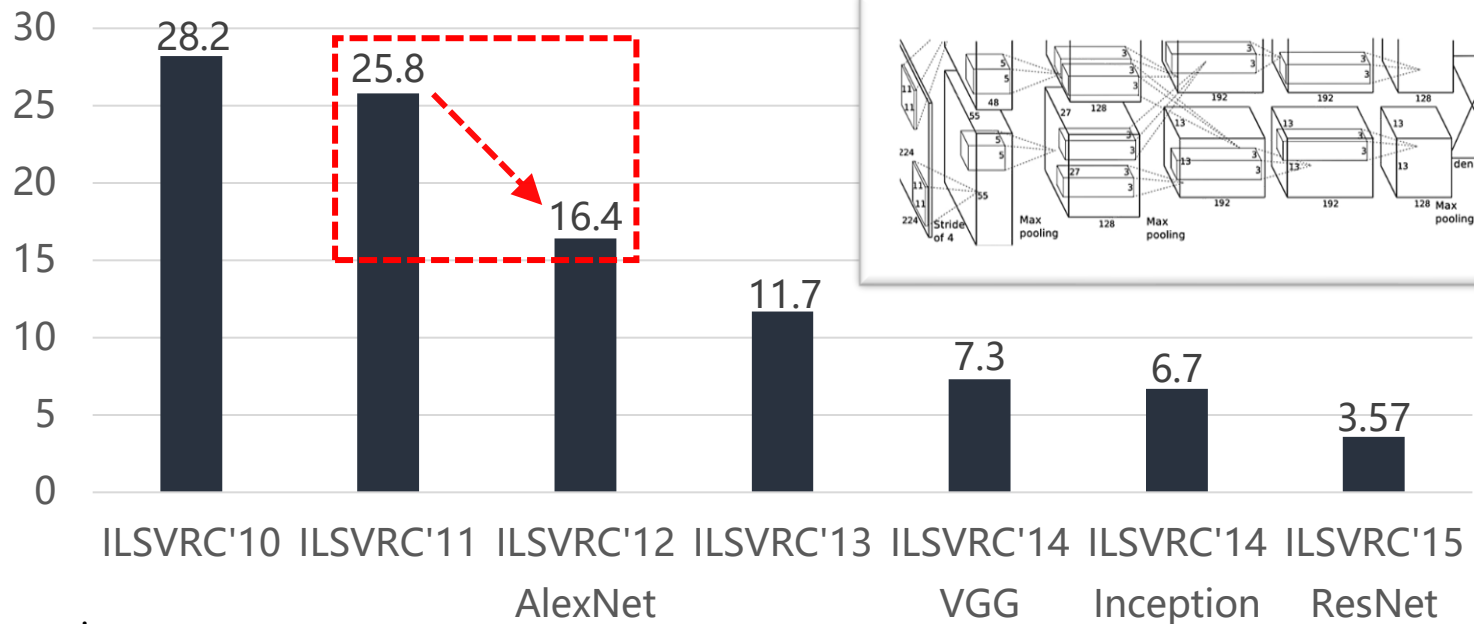
- 加拿大多伦多大学教授，机器学习领域泰斗Geoffery Hinton及其学生在《科学》上发表了一篇论文<sup>[1]</sup>，开启了深度学习在学术界和工业界的新浪潮！



[1] Hinton, G. E . Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786):504-507.

## 第三次浪潮——卷积神经网络

ImageNet 历年竞赛冠军性能列表<sup>[1]</sup>

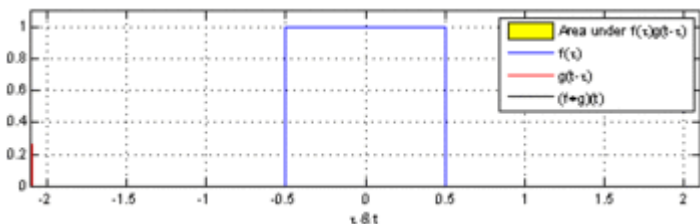


[1] <http://www.image-net.org>

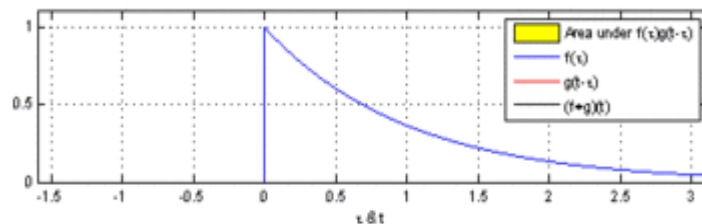
## 卷积定义

- 卷积是分析数学中的一种重要的运算
- 设  $f(x)$  和  $g(x)$  是  $\mathbb{R}$  上的两个可积函数，连续形式卷积定义如下

$$\int_{-\infty}^{\infty} f(\tau)g(x-\tau)d\tau$$



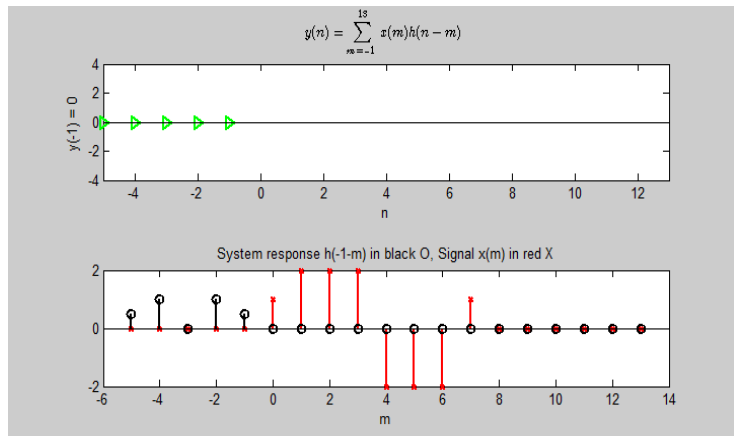
常值函数卷积结果



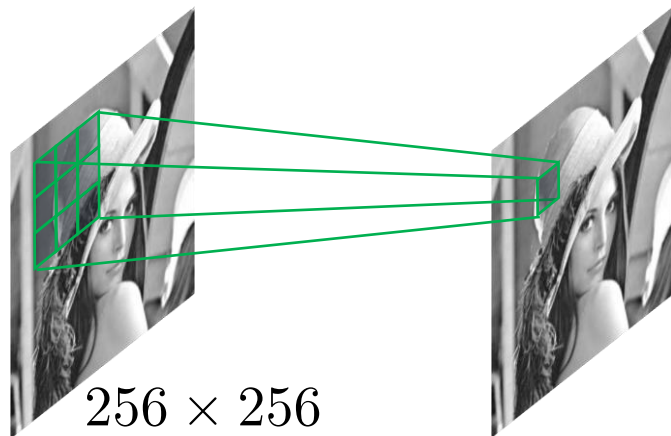
反比例函数卷积结果

## 离散空间卷积

$$y_n = \mathbf{x} * \mathbf{w} = \sum_{k=1}^K w_k x_{n-k}$$



一维离散卷积

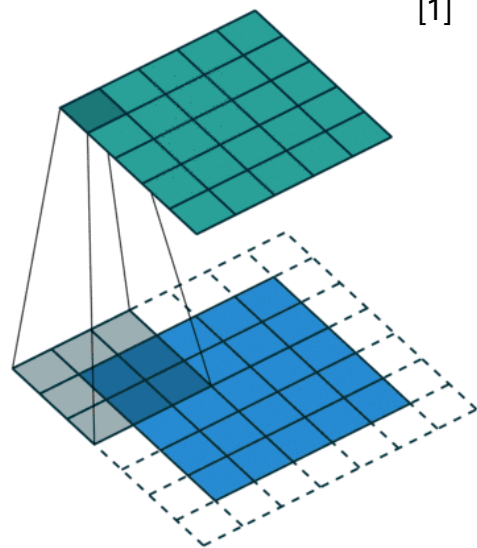


256 × 256

二维离散卷积

## 基本概念——卷积

- 卷积核大小 (Kernel Size): 卷积操作感受野, 在二维卷积中, 通常设置为3, 即卷积核大小为 $3 \times 3$
- 步长 (Stride): 卷积核遍历图像时的步幅大小, 默认值通常设置为1
- 边界扩充 (Padding): 样本边界的处理方式
- 输入与输出通道 (Channels): 构建卷积层时需定义输入通道数量 $I$ , 和输出通道数量 $O$ , 每个网络层的参数量为 $I \times O \times K$  ( $K$ 为卷积核的参数个数)



[1] [http://www.sohu.com/a/160696860\\_610300](http://www.sohu.com/a/160696860_610300)



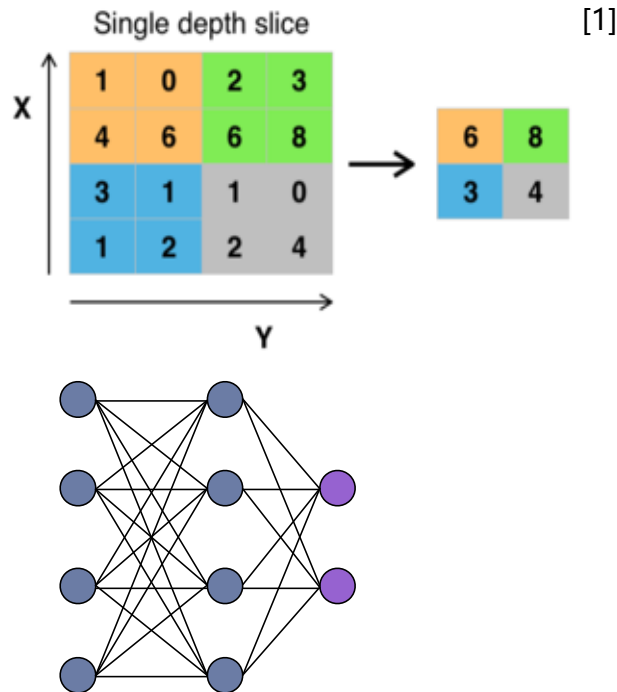
## 基本概念——池化、全连接

### ➤ 池化层

- 特殊形式卷积
- 降维、减少数据计算量，减缓过拟合，特征不变性（平移、尺度）

### ➤ 全连接层

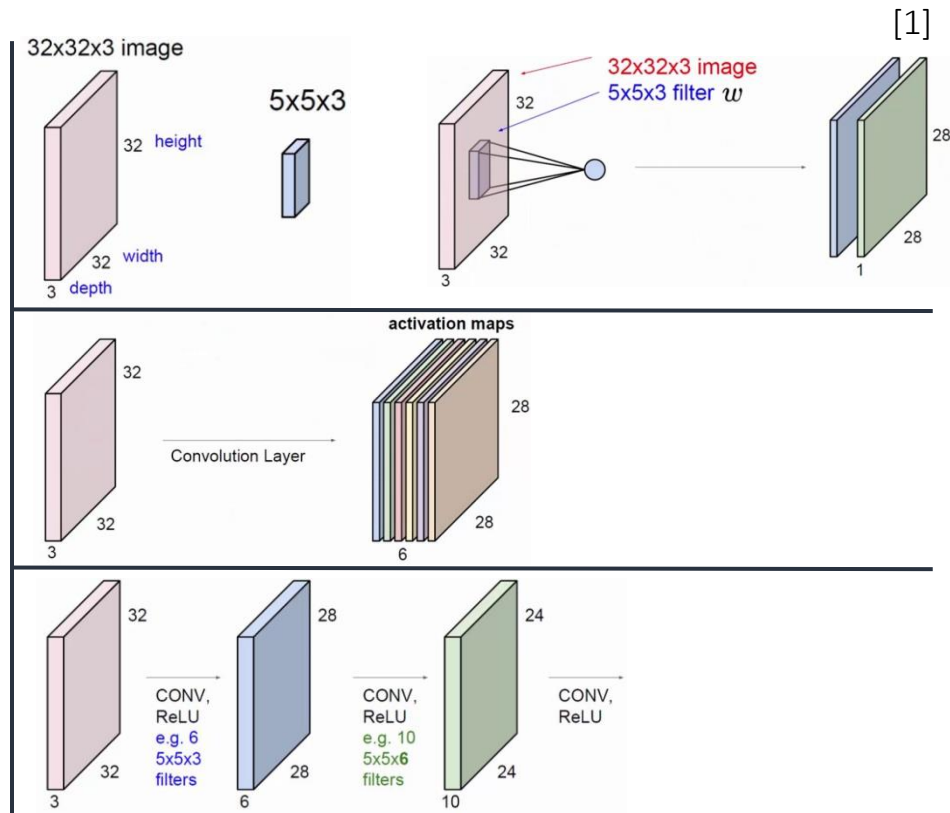
- 模型输出层
- 分类、回归



[1] <https://zh.wikipedia.org/wiki/%E5%8D%B7%E7%A7%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C>

## 多层卷积神经网络示例

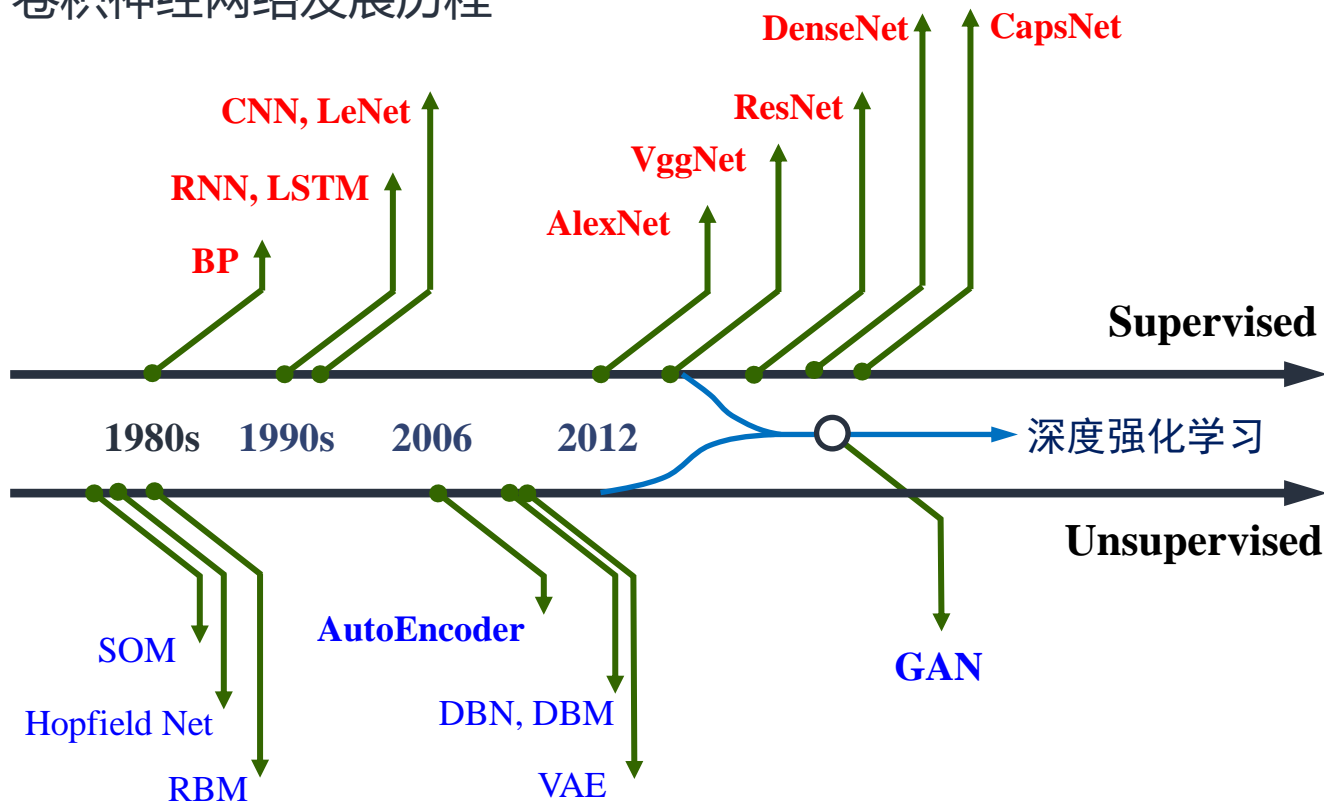
- 卷积核大小为 $5 \times 5$ ，步长为1，不扩充边界，输入通道为3，输出通道为2
- 输出通道数为6，其余参数不变。
- 第一层的输出通道数为3，第二层输出通道数为6。



[1] [https://blog.csdn.net/weixin\\_42451919/article/details/81381294](https://blog.csdn.net/weixin_42451919/article/details/81381294)

# 卷积计算与神经网络结构

## 卷积神经网络发展历程

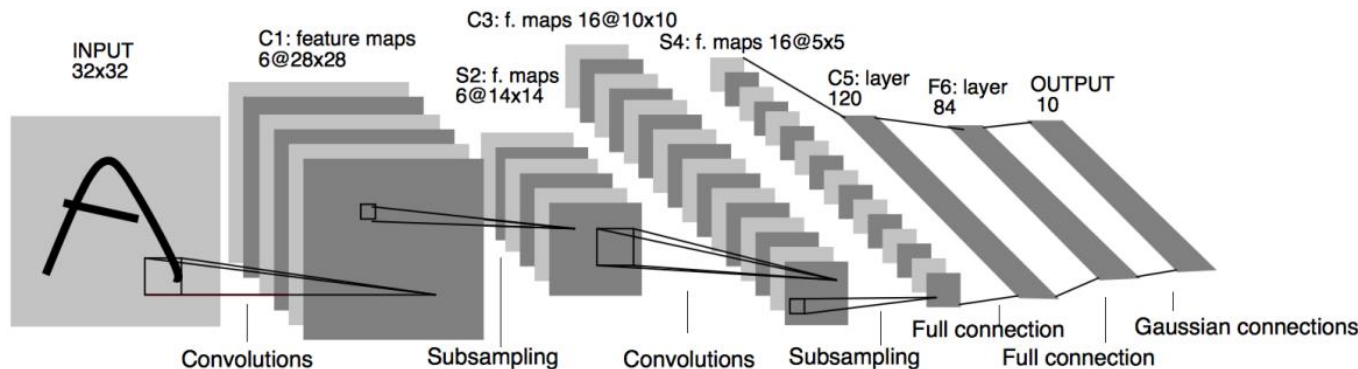


为什么能得到  
飞速发展:

- ✓ 数据爆炸: 图像数据、文本数据、语音数据、社交网络数据、科学计算等
- ✓ 计算性能大幅提高

## 现代卷积神经网络结构基础——LeNet

- LeNet诞生于1994年，由Yann LeCun提出，用于手写字符识别与分类
- 6层网络结构：两个**卷积层**，两个**下采样层**和两个**全连接层**
- 卷积层包括两部分：**卷积计算**和**sigmoid非线性激活函数**

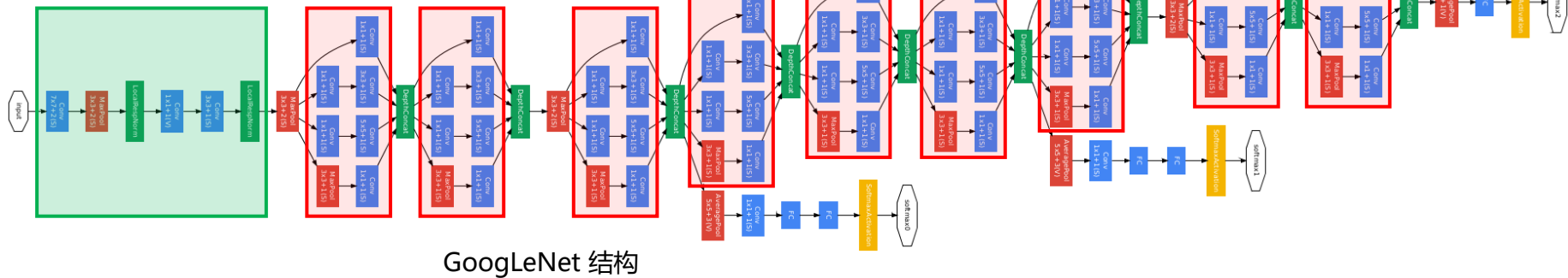
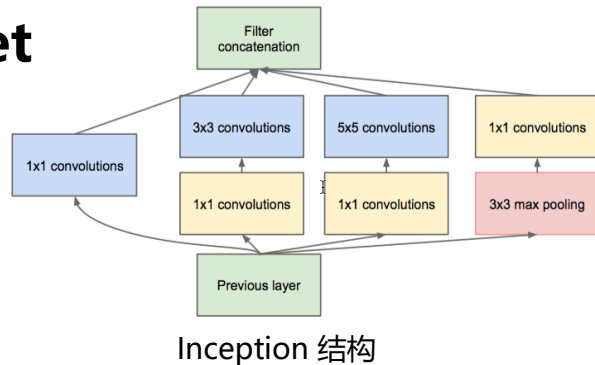


## 研究重心转向卷积神经网络——AlexNet、VGGNet

- **网络更深**：AlexNet一共8层，VGGNet一共16层或19层
- **数据增广**：为增强模型泛化能力，对 $256 \times 256$ 原始图像进行随机裁剪，得到尺寸为 $224 \times 224$ 图像，输入网络进行训练
- **ReLU非线性激活函数**：减少计算量，缓解梯度消失，缓解过拟合。ReLU激活函数现已成为神经网络中最通用的激活函数
- **Dropout**：全连接层神经元以一定概率失活，失活神经元不再参与训练。Dropout的引用，有效缓解了模型的过拟合
- **Pre-Training**：先训练一部分小网络，确保稳定之后，在此基础上网络逐渐加深。

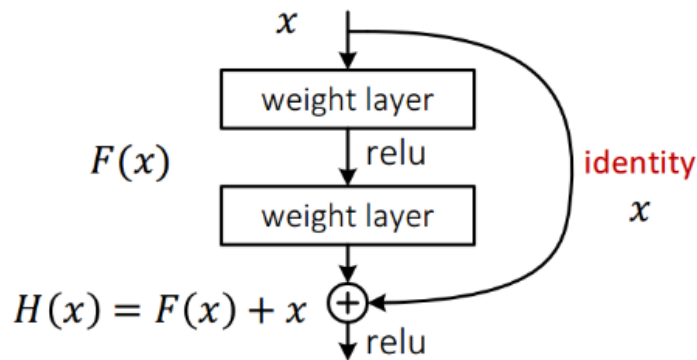
## 卷积神经网络深度、宽度扩展——GoogLeNet

- 网络更深：GoogLeNet一共22层
- 多分辨率结构：引入Inception结构替代传统卷积+激活
- 计算量降低：采用 $1 \times 1$ 卷积核来实现数据降维

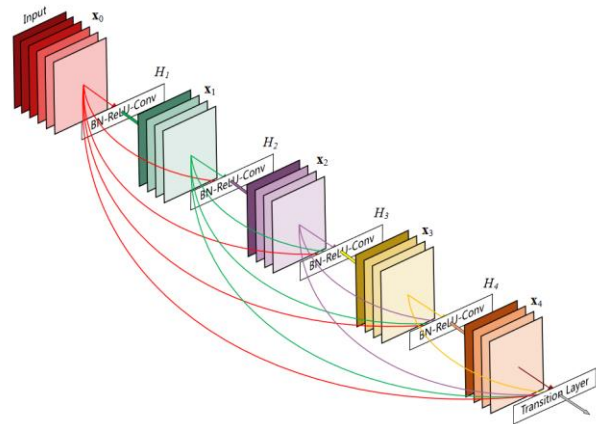


## 卷积神经网络深度、宽度再扩展——ResNet、DenseNet

- **网络更深**：ResNet已超过一百层（ResNet-101）
- **残差连接**：特征经两条路线传递，常规路线与捷径
- **跳跃连接**：底层特征与高层特征相融合



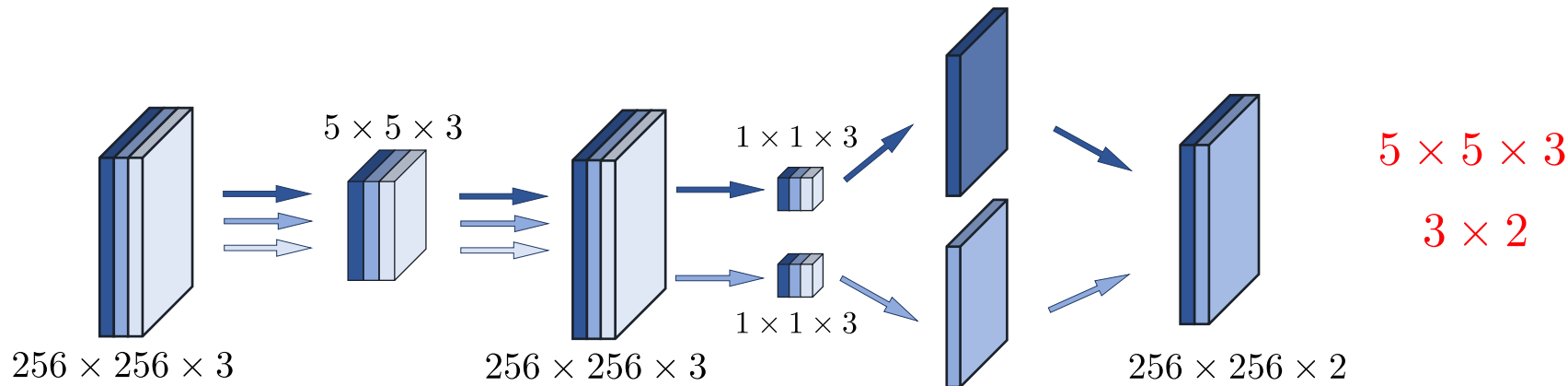
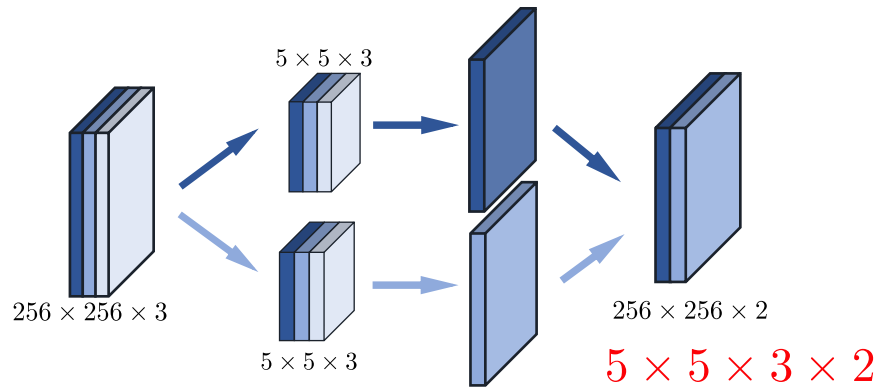
ResNet残差连接示意图



DenseNet结构示意图

## 卷积神经网络通用性扩展

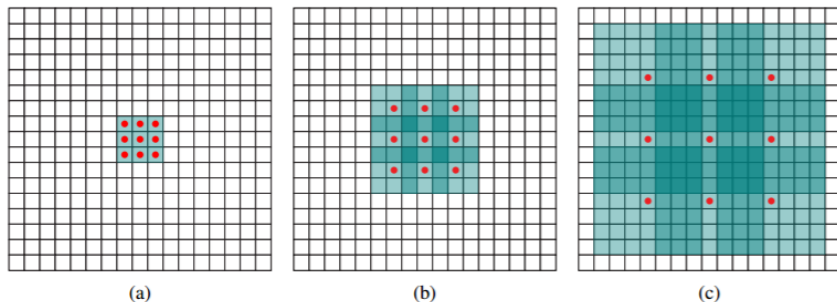
- 深度可分离卷积
  - $5 \times 5$ 分通道卷积
  - $1 \times 1$ 卷积融合各通道特征



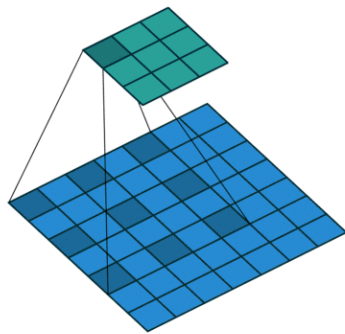


## 卷积神经网络通用性扩展

- 空洞卷积（膨胀卷积）
  - 局部输入不变
  - 感受野变大



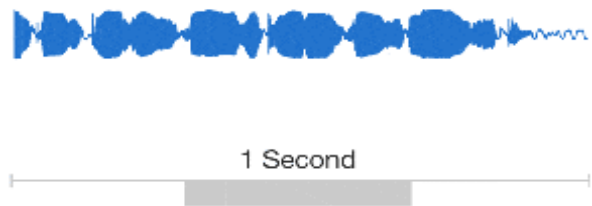
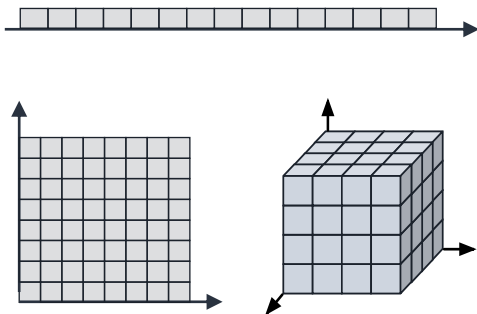
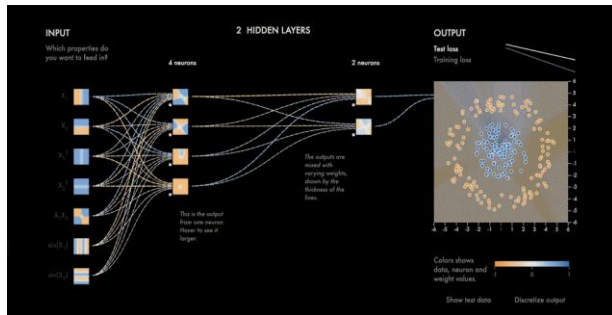
(a) 空洞卷积感受野



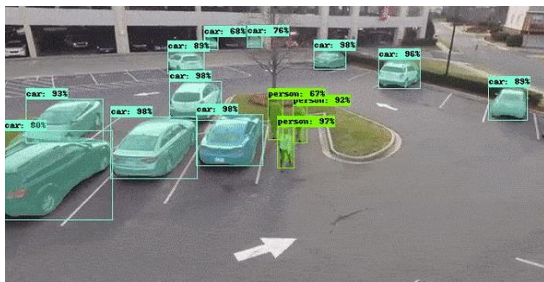
(b) 空洞卷积计算过程

# 卷积神经网络计算范式

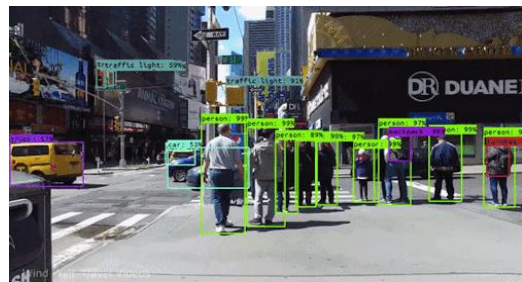
- 多维欧式空间
- 局部空间响应
- 卷积参数共享



语音



图像

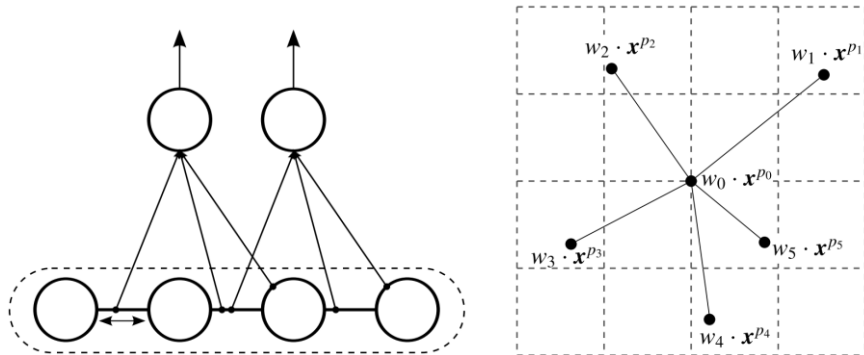


## 视频

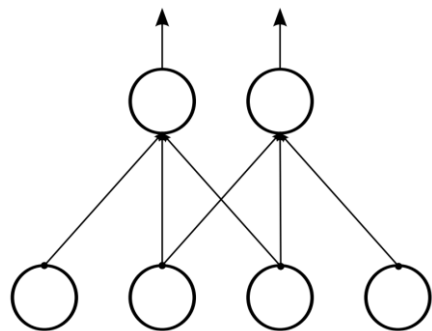
## 欧式空间非规则化连接—活性卷积

- 活性卷积 (CVPR 2017)
- 卷积核形状可变

### 活性卷积



### 基本卷积



1 <sub>x1</sub>	1 <sub>x0</sub>	1 <sub>x1</sub>	0	0
0 <sub>x0</sub>	1 <sub>x1</sub>	1 <sub>x0</sub>	1	0
0 <sub>x1</sub>	0 <sub>x0</sub>	1 <sub>x1</sub>	1	1
0	0	1	1	0
0	1	1	0	0

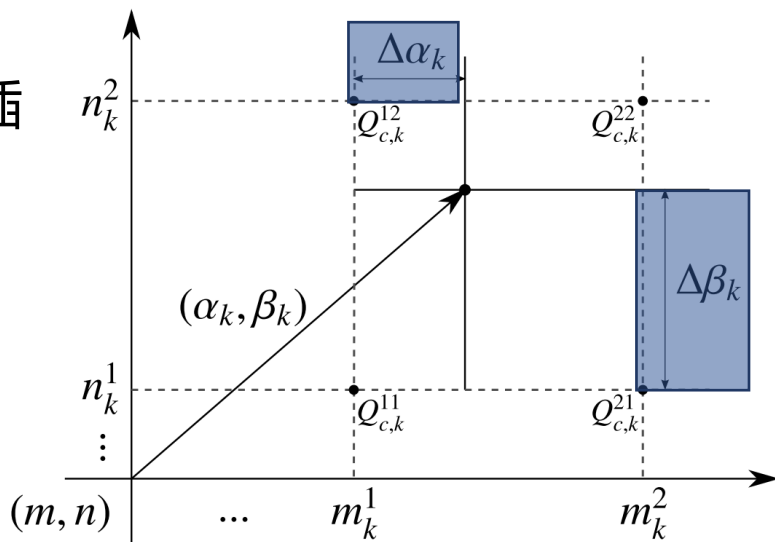
Image

4		

Convolved  
Feature

## 欧式空间非规则化连接—活性卷积

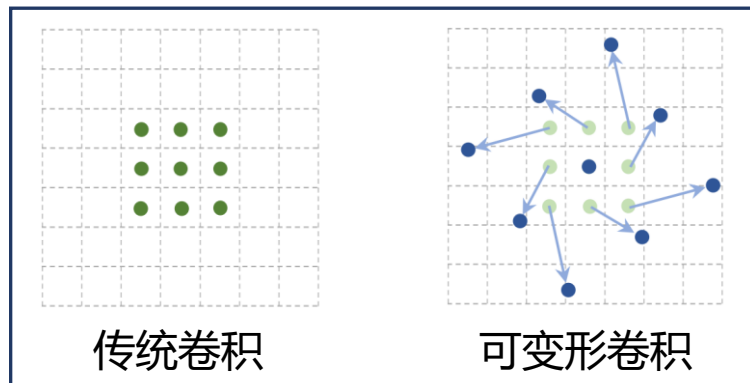
- 活性卷积 (CVPR 2017)
  - 双线性插值：离散坐标下，可以通过插值方法计算得到连续位置的像素值
  - 可学习参数  $\Delta\alpha_k, \Delta\beta_k$
  - 可变卷积核形状固定



Active Convolution: Learning the Shape of Convolution for Image Classification

## 欧式空间非规则化连接—可变形卷积

- 可变形卷积 (ICCV 2017)
  - 卷积核位置参数化
  - 双线性插值连续化
  - 传统BP算法训练



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

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

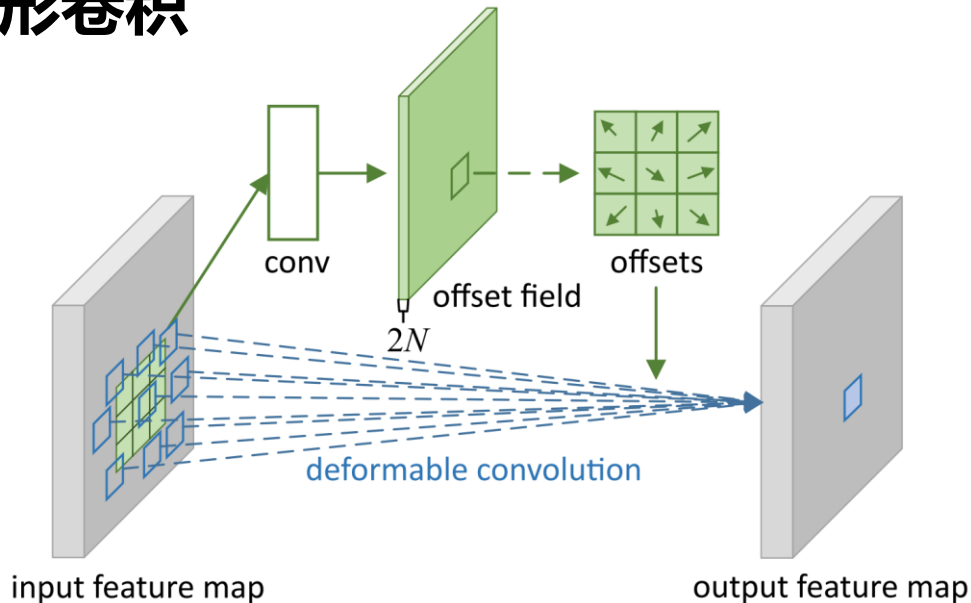


$$y(\mathbf{p}_0) = \sum_{\mathbf{p}_n \in \mathcal{R}} w(\mathbf{p}_n) \cdot \mathbf{x}(\mathbf{p}_0 + \mathbf{p}_n + \Delta \mathbf{p}_n)$$

$$\mathbf{x}(\mathbf{p}) = \sum_{\mathbf{q}} G(\mathbf{q}, \mathbf{p}) \cdot \mathbf{x}(\mathbf{q})$$

## 欧式空间非规则化连接—可变形卷积

- 可变形卷积 (ICCV 2017)
  - $3 \times 3$  可变形卷积 ( $N=9$ )
  - 每个位置对应一个偏置
  - 偏置通过额外卷积学习
  - 每个偏置为二维向量



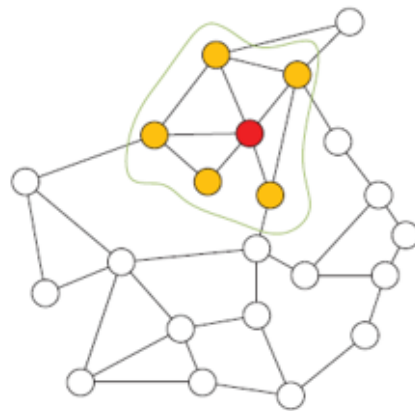
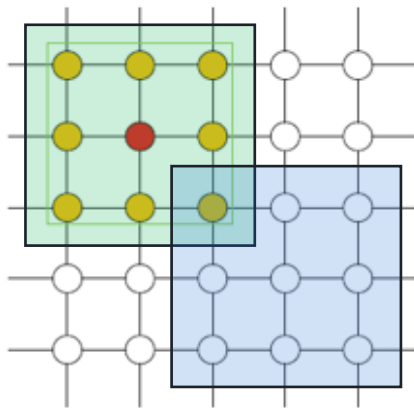
Deformable Convolutional Networks

## 欧式空间卷积神经网络

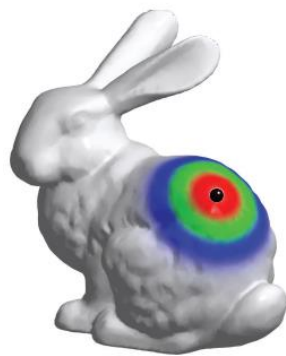
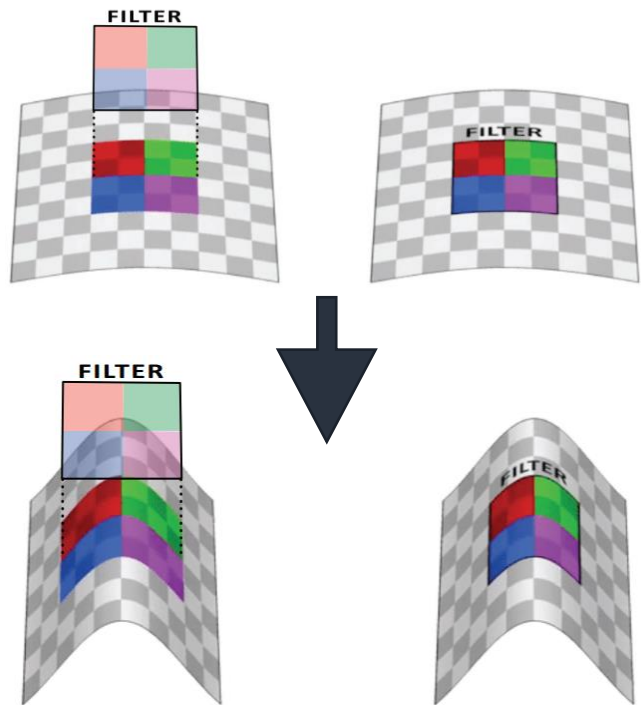
- 处理固定输入维度数据、局部输入数据必须有序
- 语音、图像、视频（规则结构）满足以上两点要求

## 非欧式空间结构数据

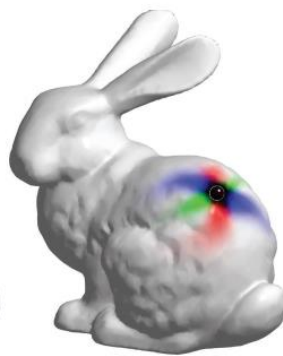
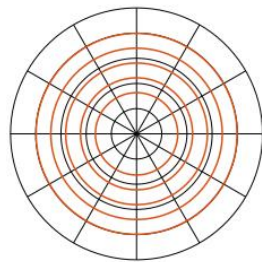
- 局部输入维度可变
- 局部输入排列无序



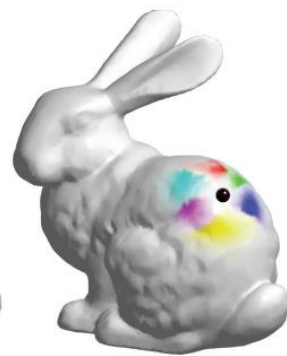
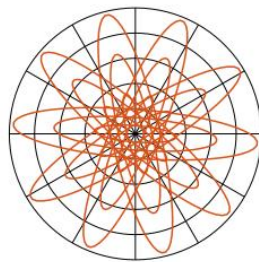
## 非欧式空间卷积神经网络



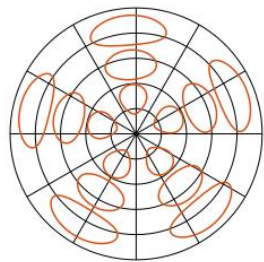
Diffusion distance



Anisotropic  
heat kernel



Geodesic polar  
coordinates





- 绪论（本次课程内容回顾）
- 谱域图卷积介绍
- 空域图卷积介绍（一）
- 空域图卷积介绍（二）
- 图卷积的实践应用
- 基于PyTorch的图卷积代码实现



# 谢谢大家

---

END