深度学习典型网络

1. LeNet

**1990-Handwritten Digit Recognition with a Back-Propagation Network**

**Introduction：**

1. Large-scale BP能够工作于真实的图像识别而不需要大型的预训练。
2. 输入不同于先前的研究中以图像特征向量作为输入，取而代之的是直接以原始图像作为输入，表明BP network可以挖掘处理大量的低层次信息。
3. 已有的研究（1989）中的模型泛化能力主要依赖于先验知识的网络结构设计。
4. 基本的约束原则是在保证网络的问题表达能力的基础上，减少模型需要学习的参数。

**Zipcode Recognition：**

1. Problem： 2D image space -> category space。
2. 字符数据库：Train Set: 7291 handwriting + 2549 printed digits. Test Set: 2007 + 700

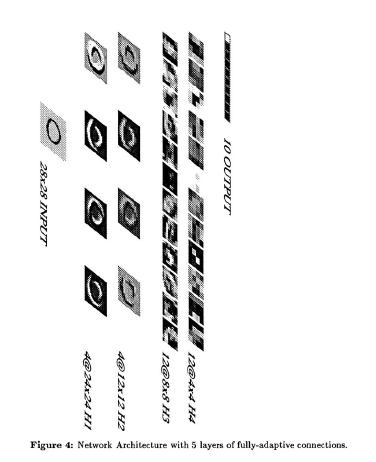
**Preprocession**

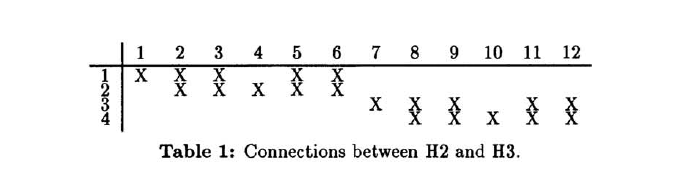
1. Size: 40\*60(around) -> normalized in 16\*16 pixels.
2. Intensity Scale Range: (-1, 1).

**The Network**

1. 网络一共6层，其中第一层输入层，为原始字符图像，最后一层输出层为全连接层，共十个可能输出，代表十个数字 。中间包括4个隐藏层，H1为卷积层，H2为降采样（pooling）层，H3为卷积层，H4为降采样层
2. 采用权值共享，减少训练参数。
3. 采用降采样，进一步减少训练参数。

（图像的平稳特性， 平移不变性）





1. AlexNet

**2012. ImageNet Classification with Deep Convolutional Neural Networks》**

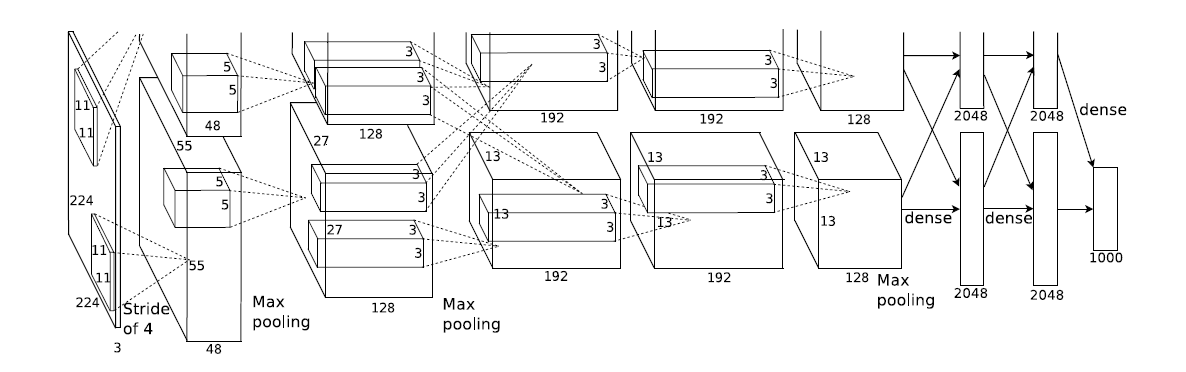
**DataSet:**

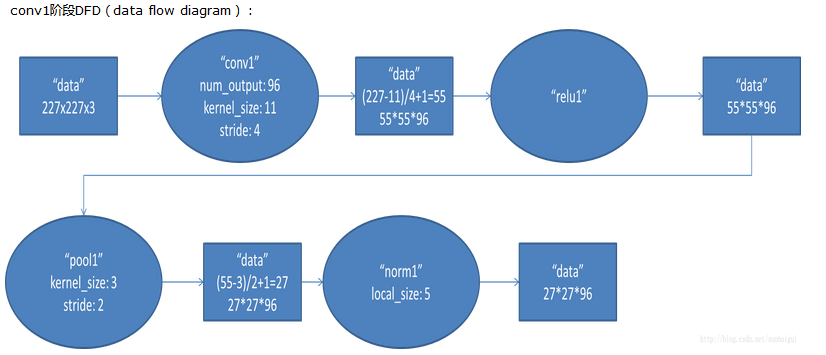
1. ImageNet: 15million, labeled, 22000 classes.
2. ILSVRC:ImageNet Large-Scale Visual Recogonition Challenge.
3. Top-1 and top-5: where the top-5 error rate is the fraction of test images for which the correct label is not among the five labels considered most probable by the model.

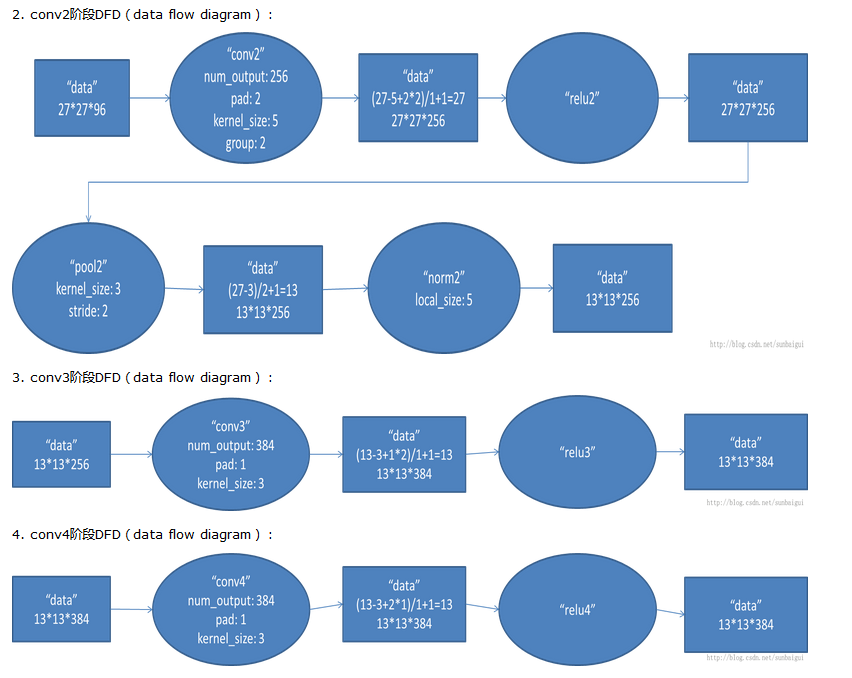
mAP？

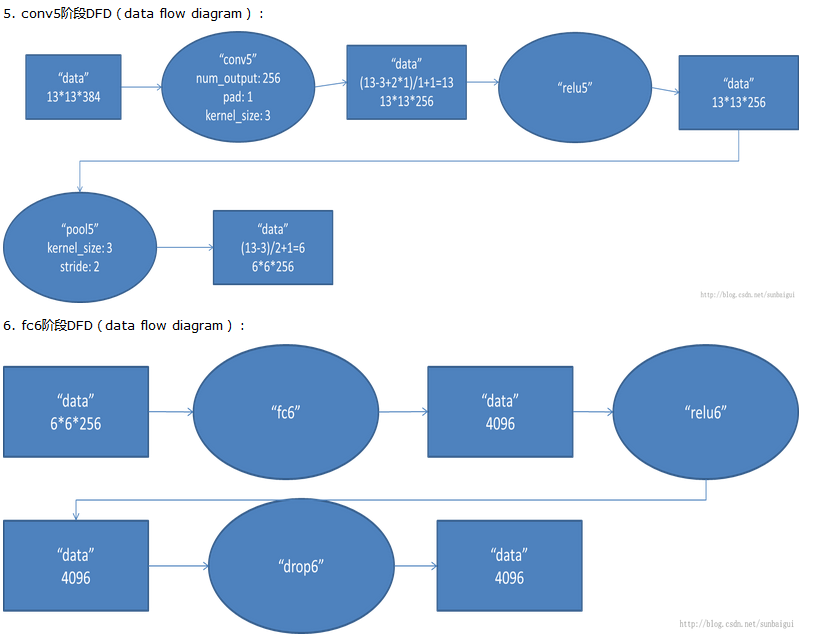
1. Data Restriction: AlexNet need a fixed input size of image(256X256).

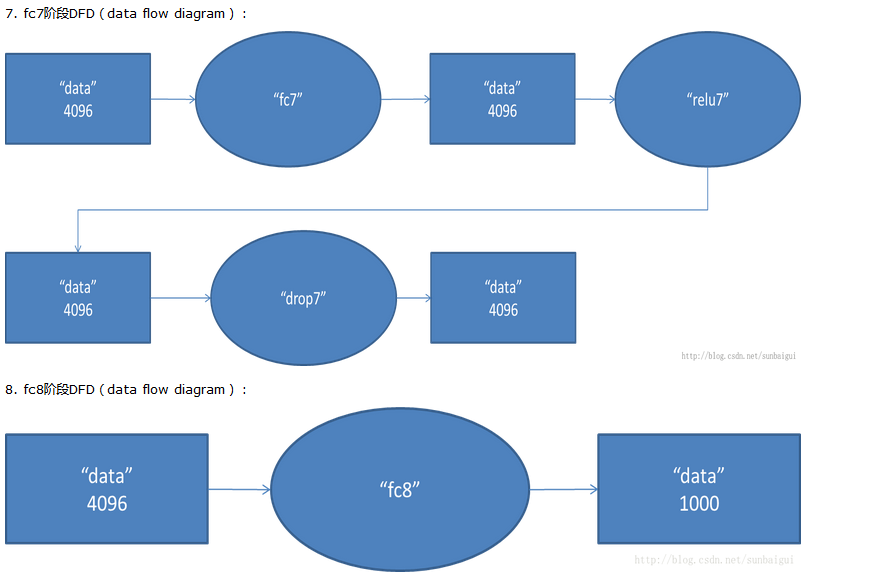
**The Architecture**

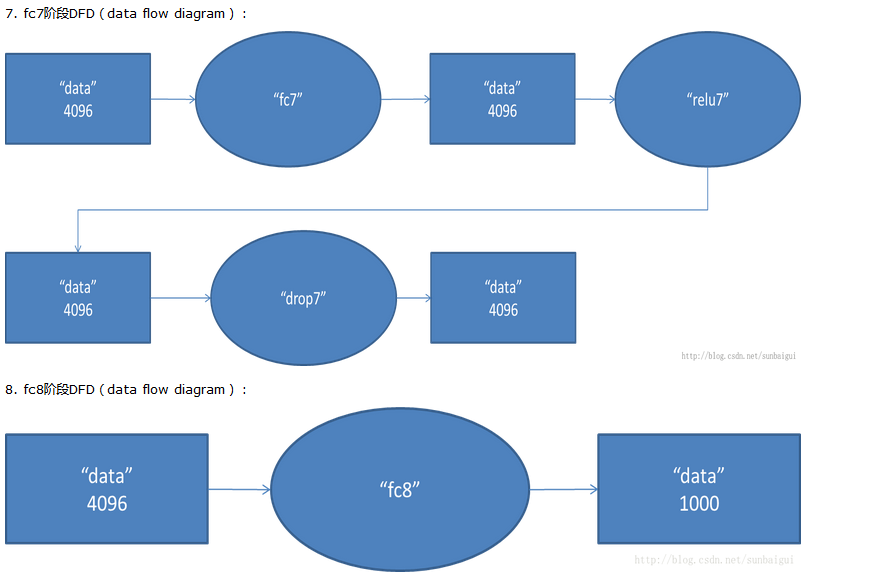


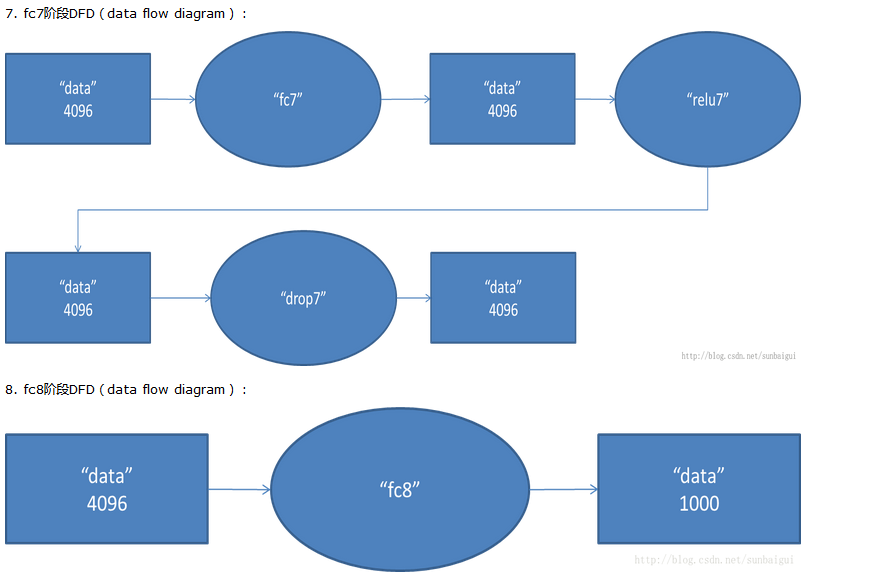






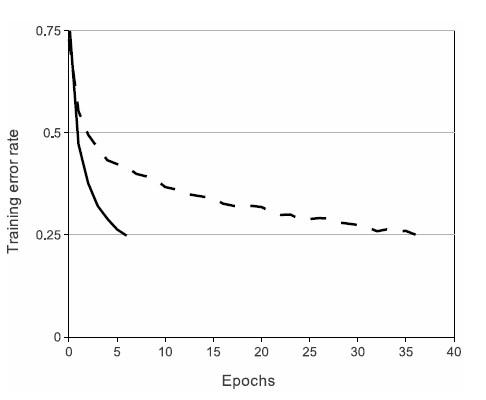






**ReLU**

1. Relu function: f(x) = Max(0,x).



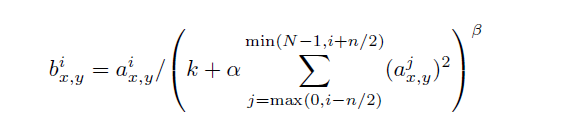
作用：

1. 对输入不需要限制
2. 加速迭代收敛
3. 计算简单反向传播速度快
4. 输入为正时才被激活，否则抑制，有助于控制过拟合
5. 常用的激活函数





**Local Response Normalization（局部响应归一化）**



计算邻近的feature map的归一化值，原文说的是kernel map,应该是一个意思

作用：防止过拟合，抑制过分激活的神经元？

原文：This sort of response normalization implements a form of lateral inhibition inspired by the type found in real neurons, creating competition for big activities amongst neuron outputs computed using different kernels。

**Overlapping Pooling(s<z)**

1. Premote precision by 0.4% and 0.3% in top-1 & top-5 error.
2. We generally observe during training that models with overlapping pooling find it slightly more difficult to overfit
3. Pooling layer Output Size : (InputSize – KernelSize)/Stride+1.

**Reducing Overfitting**

1. Data Augmentation:

The first form of data augmentation consists of generating image translations and horizontal reflections.

The second form of data augmentation consists of altering the intensities of the RGB channels in training images. Specifically, we perform PCA on the set of RGB pixel values throughout the ImageNet training set

1. DropOut
2. 最简单的方式是多个model一起组合判断，然而神经网络里面代价太高。
3. DropOut在每次训练时以0.5的概率将神经元关闭，这样其不参与前向和后向传播。就相当于每次在训练不同的网络结构（不同的模型）
4. DropOut有效防止过拟合而且能提升收敛速度。（只用在了前连个全连接层后面）
5. VGG
6. GoogleNet
7. R-CNN
8. Fast-RCNN
9. Faster-RCNN
10. YOLO
11. SSD