Normalization

为什么需要 Normalization?

我们希望处理的数据是独立同分布，独立同分布的数据可以简化常规机器学习模型的训练、提升机器学习模型的预测能力。

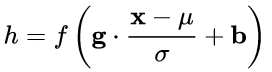
在把数据喂给机器学习模型之前，“白化（whitening）”是一个重要的数据预处理步骤。白化一般包含两个目的：

1.去除特征之间的相关性 —> 独立;

2.使得所有特征具有相同的均值和方差 —> 同分布。

标准白化操作的目的是“独立同分布”.

白化最典型的方法就是PCA.

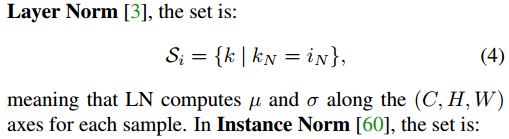


**为了保证模型的表达能力不因为规范化而下降**.

 is the set of pixels in which the mean and std are computed, and m is the size of this set. Many types of feature normalization methods mainly differ in how the set Si is defined (Figure 2), discussed as follows。

1.Layer Normalization

LN 针对单个训练样本进行，不依赖于其他数据，因此可以避免 BN 中受 mini-batch 数据分布影响的问题，可以用于 小mini-batch场景、动态网络场景和 RNN，特别是自然语言处理领域。此外，LN 不需要保存 mini-batch 的均值和方差，节省了额外的存储空间.



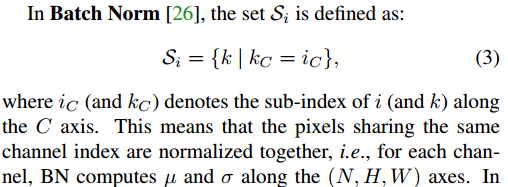
2.Weight Normalization

对网络权值W进行normalization（L2 norm），因此也称为Weight Normalization。

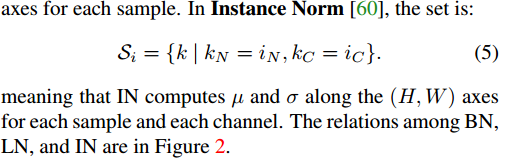
3.Batch Normalization(纵向规范化)

Batch Normalization是对网络某一层输入数据进行normalization。

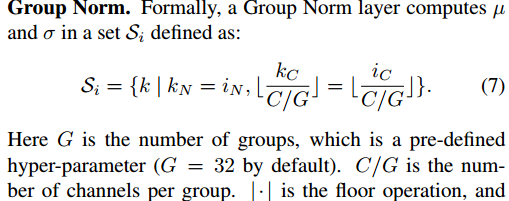
每个 mini-batch 比较大，数据分布比较接近。在进行训练之前，要做好充分的 shuffle. 否则效果会差很多.不适应于动态的网络结构 和 RNN 网络.

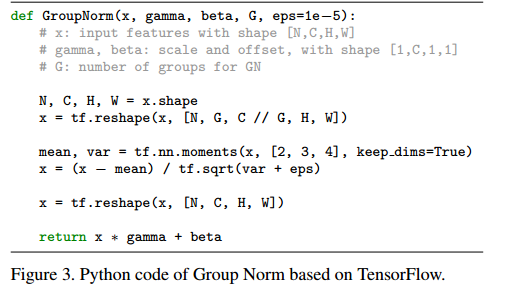


4.Instance Normalization



5.Group Normalization





比较:

