最近的工作表明，如果卷积网络在靠近输入的层和靠近输出的层之间包含较短的连接，则卷积网络可以更深，更准确，并且更有效。在本文中，我们接受了这一观察，并介绍了密集卷积网络（DenseNet），它以前馈的方式将每一层连接到每一层。而具有L层的传统卷积网络具有L个连接 - 每个层与其后续层之间的连接 - 我们的网络具有L（L + 1）2个直接连接。对于每个图层，所有在前图层的要素图用作输入，其自身的要素图用作所有后续图层的输入。 DenseNets具有几个引人注目的优点：它们可以缓解消失梯度问题，加强特征传播，鼓励特征重用，并大幅减少参数数量。我们在四个竞争激烈的对象识别基准测试任务（CIFAR10，CIFAR-100，SVHN和ImageNet）上评估我们提出的架构。 DenseNets在大多数设备上获得了相对于最新技术的显着改进，同时需要更少的内存和计算来实现高性能。代码和型号可在https://github.com/liuzhuang13/ DenseNet上获得。

Introduction

Related Work

DenseNets

Experiments

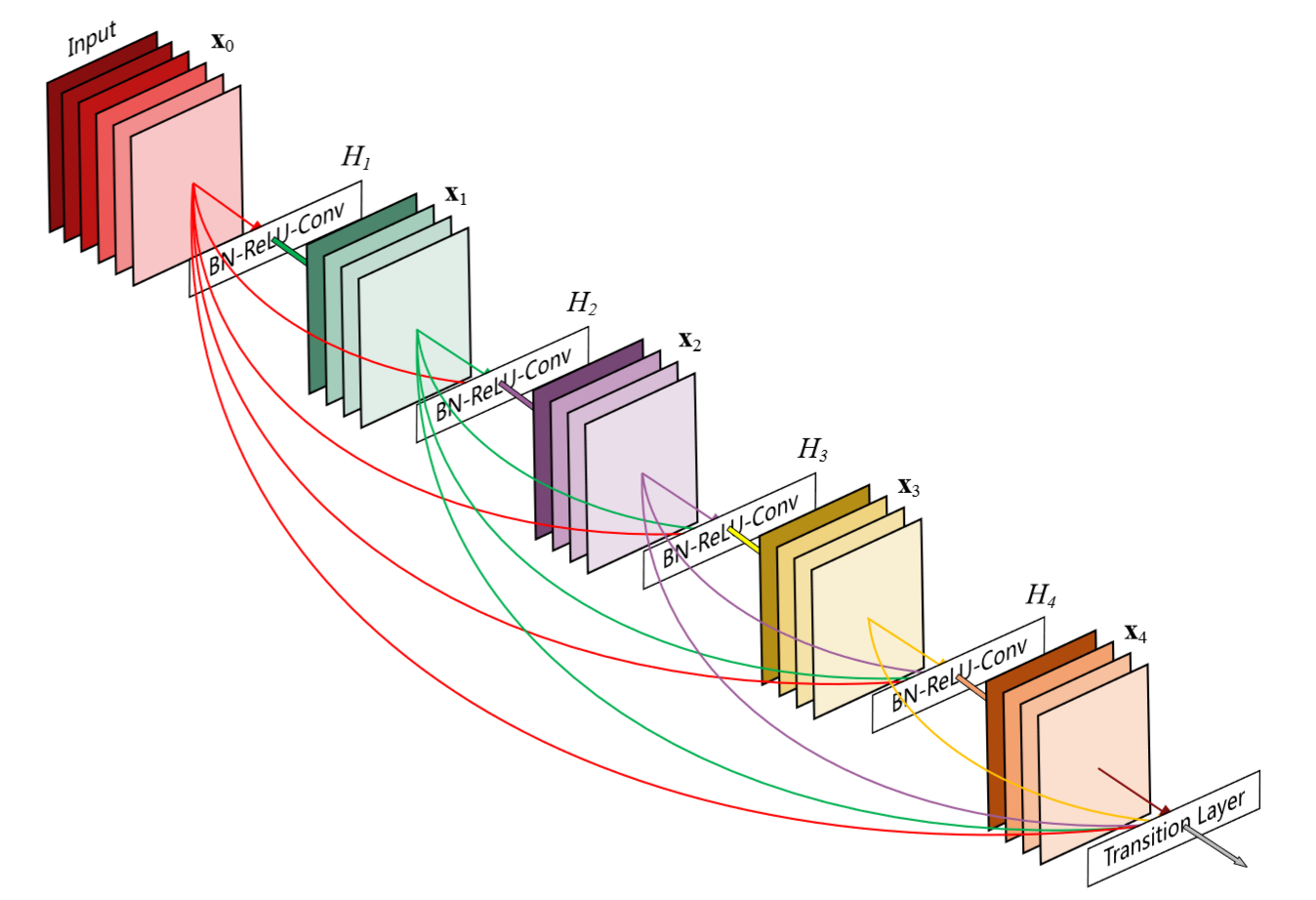
Discussion

Conclusion

Introduction

保证网络中层间最大的信息流，将每个层都直接的连在一起。

每层都获得了额外的来自之前所有层的输入，并且把自己的特征图传给所有的下面的层



与ResNet相比，我们不在特征传到下一层之前就把他们加起来(summation)，我们通过连结（concatenate）来结合特征。第l层有l个输入，包含所有之前的卷积block的特征图，它的自己的特征图传给了所有的后面的L-l层。这样在L层的网络中有L(L+1)/2个连接。因为这种密集连接模式，此网络成为Dense Convolutional Network (DenseNet).

Densenet

x0为输入，Hl(.)代表一种非线性的转换，包含Batch Normalization (BN)，ReLU激活函数，池化，卷积。第l层的输出为xl

ResNet使用跳跃连接，绕过了非线性转换。ResNet的优点是梯度通过恒等函数从从后面层流向前面的层。但是缺点是恒等函数和输出是通过求和(summation)组合在一起的，这会阻碍网络中的信息流。

Dense connectivity

第l层接收前面所有层的特征图作为输入

Composite function

Hl(.)代表Batch Normalization (BN)，ReLU激活函数和3×3的卷积

不同层的连接需要相同的特征图，但是池化层也是必不可少的。引入多重密集连接的dense blocks，在每个block中间的是transition layers（过渡层），transition layer包含batch normalization 层，1×1卷积和池化

Growth rate k

Hl(.)函数输出k个特征图，表明第l层有k(l-1)+k0个输入特征图，k0是输入图像的通道数。超参数k设定的小，为12，为了防止网络过宽，提高参数效率。

Bottleneck layers

和ResNet一样，在3×3卷积之前加入1×1卷积来减少维度，增加计算效率

BN-ReLU-Conv(1× 1)-BN-ReLU-Conv(3×3) 版本的Hl(.),，称为DenseNet-B

Compression压缩

在transition layers中使用。设dense block有m个特征图，经过transition layers产生θm个输出。将θ<1的DenseNet称为DenseNet-C，如果同时使用bottleneck 和θ < 1 ，称为DenseNet-BC.

