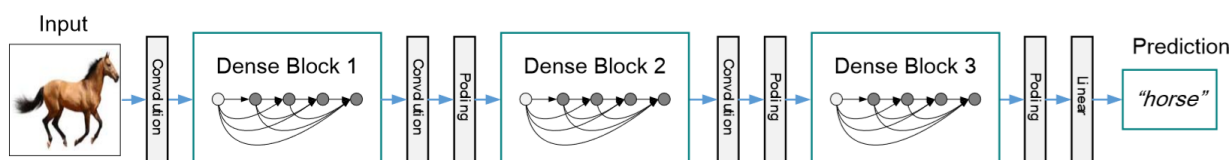


## 简介

作者认为ResNet直接把残差和输出相加，which may impede the information flow in the network.

## Dense连接



前面层的输出作为后面所有层的输入，形成dense连接

$$\mathbf{x}_\ell = H_\ell([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}]), \quad (2)$$

## composition function

dense block的基本结构 $H_l()$ 是三个层连接在一起，分别是BN-ReLU-3x3Conv

## 池化层

池化对于减少参数至关重要但是池化之后feature map尺寸变化无法再进行dense连接。所以作者分了两个模块dense block和transition layers。transition layers为过渡层，BN-1x1Conv-2x2 AvgPool

## 增长率

设每层 $l$ 的输出  $H_l$  的channel为 $k$ ，那么第 $l$ 层的输入channel为  $k_0 + k \times (l - 1)$ 。因为dense layer的增长率很快，所以其channel(layer)可以选比较小，比如 $k=12$ 。将 $k$ 命名为网络增长率。实验证明小的增长率已经可以有很好的效果。这是因为dense结构的知识是共享的，因此每个层不需要很大的channel。

## Bottleneck层

虽然每层只输出 $k$ 个channel但是后面层的输入channel是很大的，实验发现1x1的bottleneck卷积对网络的作用很大。作者将 BN-ReLU-Conv(1x1)-BN-ReLU-Conv(3x3) 的  $H_l$  结构称作 DenseNet-B 结构，在实验当中利用1x1卷积产生4k个feature map。

## 进一步压缩

为了使得模型更加紧凑，作者在transition layers也减了channel。是得其channel的输出为  $\lfloor \theta m \rfloor$  实验中  $\theta = 0.5$ 。这个变体成为DenseNet-C，与B变体结合后成为DenseNet-BC

## 实现细节

除了ImageNet数据集其他数据集都采用了三个dense block的结构，每个block层数相同。在进入第一个dense block之前，用了16 channel的卷积。最后一个dense block之后接AvgPool。三个Dense block feature-map size为32x32, 16x16, 8x8。

对于原始的DenseNet采用了  $\{L=40, k=12\}$  ,  $\{L=100, k=12\}$  和  $\{L=100, k=24\}$  三种结构。而对于变体结构DenseNet-BC，采用的设定为  $\{L=100, k=12\}$  ,  $\{L=250, k=24\}$  和  $\{L=190, k=40\}$