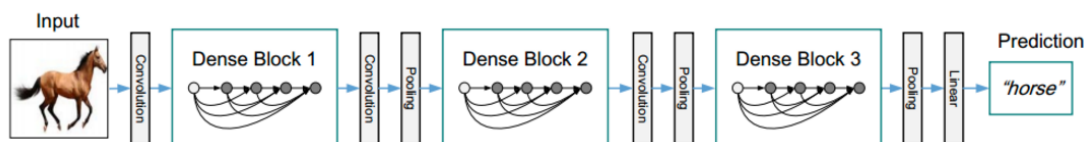


## 1.1 亮点

- (1) 作者设计出一个网络，使得让每一个层都接受它前面所有层的输出，这样的话，对于普通的神经网络，L 层一共有  $L$  个 connection.但是对于 densely convolutional network，L 层一共有  $L(L-1)/2$  个 connection。
- (2) 有效缓解梯度消失问题
- (3) 重复利用特征，提高特征质量。

## 1.2 原理



对于每一个 block 内部的卷积层采用 densely connected 的方式，即每一个 unit 的输入都是前面所有 unit 的输出的 concat。上图中每一个 unit 表征的操作为 BN-ReLU-Conv(1x1)-BN-ReLU-Conv(3x3)，引入 Conv(1x1)的目的是为了减少计算量。同一个 block 内部每一个 unit 输出 feature map 的 channel 都固定为  $4k$ ， $k$  表示 growth rate，显然  $k$  越大，网络的参数越多，capacity 越大，学习能力越强。

在 Dense Block 内，作者引入一个参数  $k$  (growth rate)，它表示“Dense Block”内每一层输出 feature map 的个数。如果“Dense Block”的层数较多、 $k$  值较大，那么“Dense Block”最后一层的输入个数是很大的，因此  $k$  取值不应该太大。

为什么 Dense Block 是分离的呢？我认为 pooling layer 是一个很重要的因素，作者想用 pooling layer。在 Dense Block 内，第  $L$  层来自它前面所有层的输入是以“concatenate”的形式组成一个 single tensor 的，这样就要求它前面所有层具有相同的 feature map sizes。这和 pooling layer 改变 feature map sizes 的特点相矛盾。作者做了一个折中，以 pooling layer 为界分割 Dense Block，即论文中的 transition layer。

## 1.3 实验

Method	Depth	Params	C10	C10+	C100	C100+	SVHN
Network in Network [22]	-	-	10.41	8.81	35.68	-	2.35
All-CNN [31]	-	-	9.08	7.25	-	33.71	-
Deeply Supervised Net [20]	-	-	9.69	7.97	-	34.57	1.92
Highway Network [33]	-	-	-	7.72	-	32.39	-
FractalNet [17]	21	38.6M	10.18	5.22	35.34	23.30	2.01
with Dropout/Drop-path	21	38.6M	7.33	4.60	28.20	23.73	1.87
ResNet [11]	110	1.7M	-	6.61	-	-	-
ResNet (reported by [13])	110	1.7M	13.63	6.41	44.74	27.22	2.01
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23	37.80	24.58	1.75
	1202	10.2M	-	4.91	-	-	-
Wide ResNet [41]	16	11.0M	-	4.81	-	22.07	-
	28	36.5M	-	4.17	-	20.50	-
with Dropout	16	2.7M	-	-	-	-	1.64
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46	35.58*	24.33	-
	1001	10.2M	10.56*	4.62	33.47*	22.71	-
DenseNet ( $k = 12$ )	40	1.0M	<b>7.00</b>	5.24	<b>27.55</b>	24.42	1.79
DenseNet ( $k = 12$ )	100	7.0M	<b>5.77</b>	<b>4.10</b>	<b>23.79</b>	<b>20.20</b>	1.67
DenseNet ( $k = 24$ )	100	27.2M	<b>5.83</b>	<b>3.74</b>	<b>23.42</b>	<b>19.25</b>	<b>1.59</b>
DenseNet-BC ( $k = 12$ )	100	0.8M	<b>5.92</b>	4.51	<b>24.15</b>	22.27	1.76
DenseNet-BC ( $k = 24$ )	250	15.3M	<b>5.19</b>	<b>3.62</b>	<b>19.64</b>	<b>17.60</b>	1.74
DenseNet-BC ( $k = 40$ )	190	25.6M	-	<b>3.46</b>	-	<b>17.18</b>	-