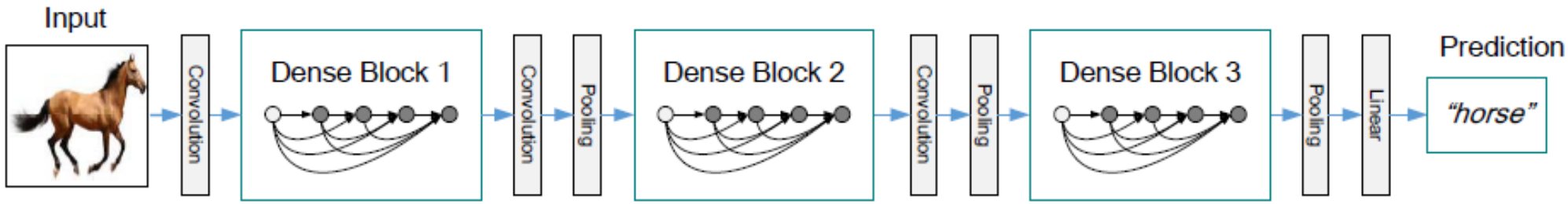
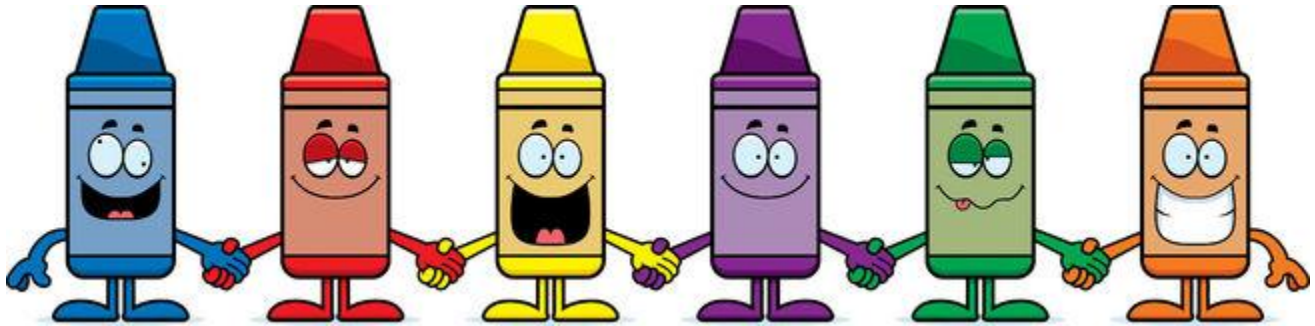
Densely Connected Convolutional Networks

Ziqiang Zheng

which connects each layer to every other layers in a feed-forward fashion



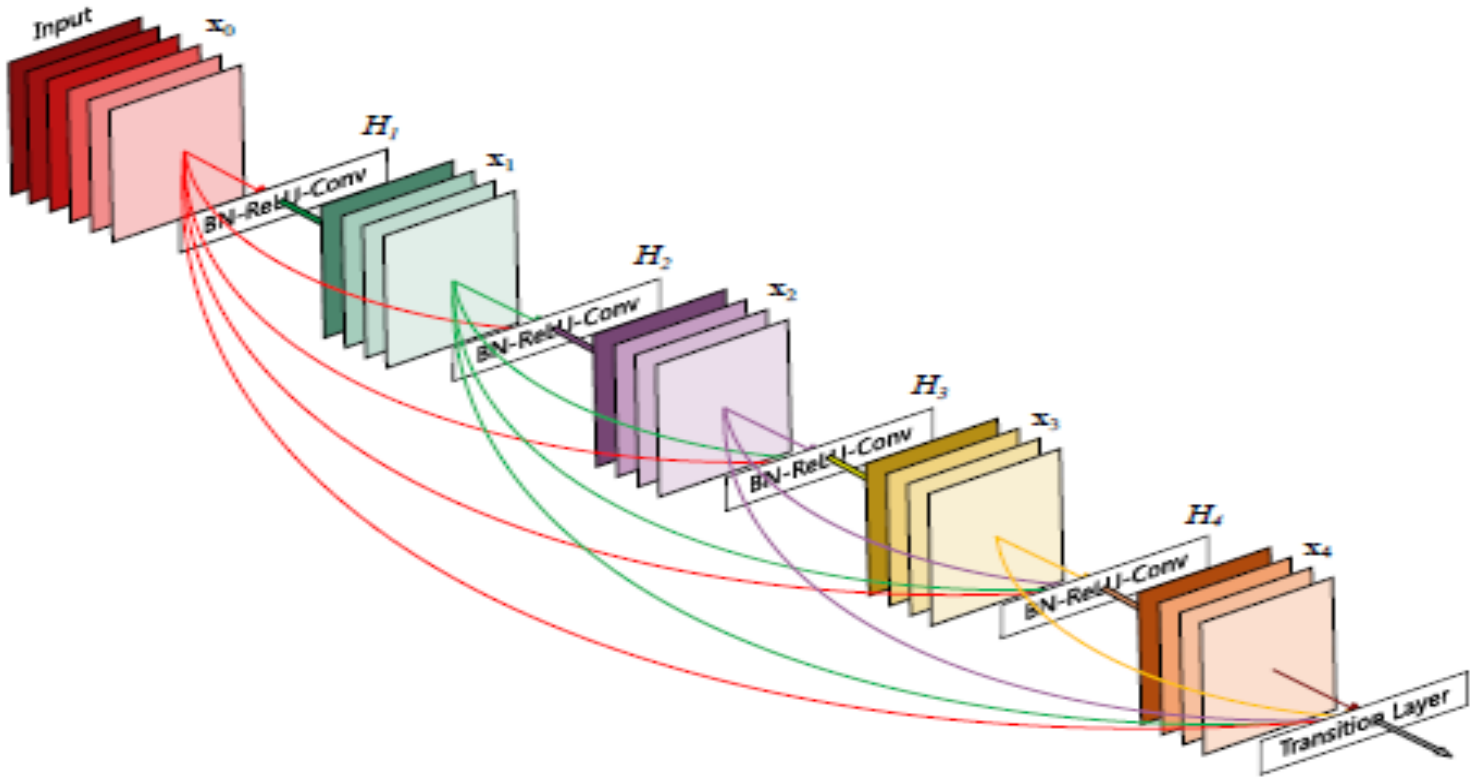
$$L * \frac{L+1}{2} \; connections$$



Why dense connection?

- Direct access to the gradients from the loss function
- Improved flow of information and gradients throughout the network
- Utilizing multi-level features in CNNs
- Skip connection

Feature reuse



Compared to ResNet

- ResNet

$$\mathbf{x}_\ell = H_\ell(\mathbf{x}_{\ell-1}) + \mathbf{x}_{\ell-1}.$$

- Densnet

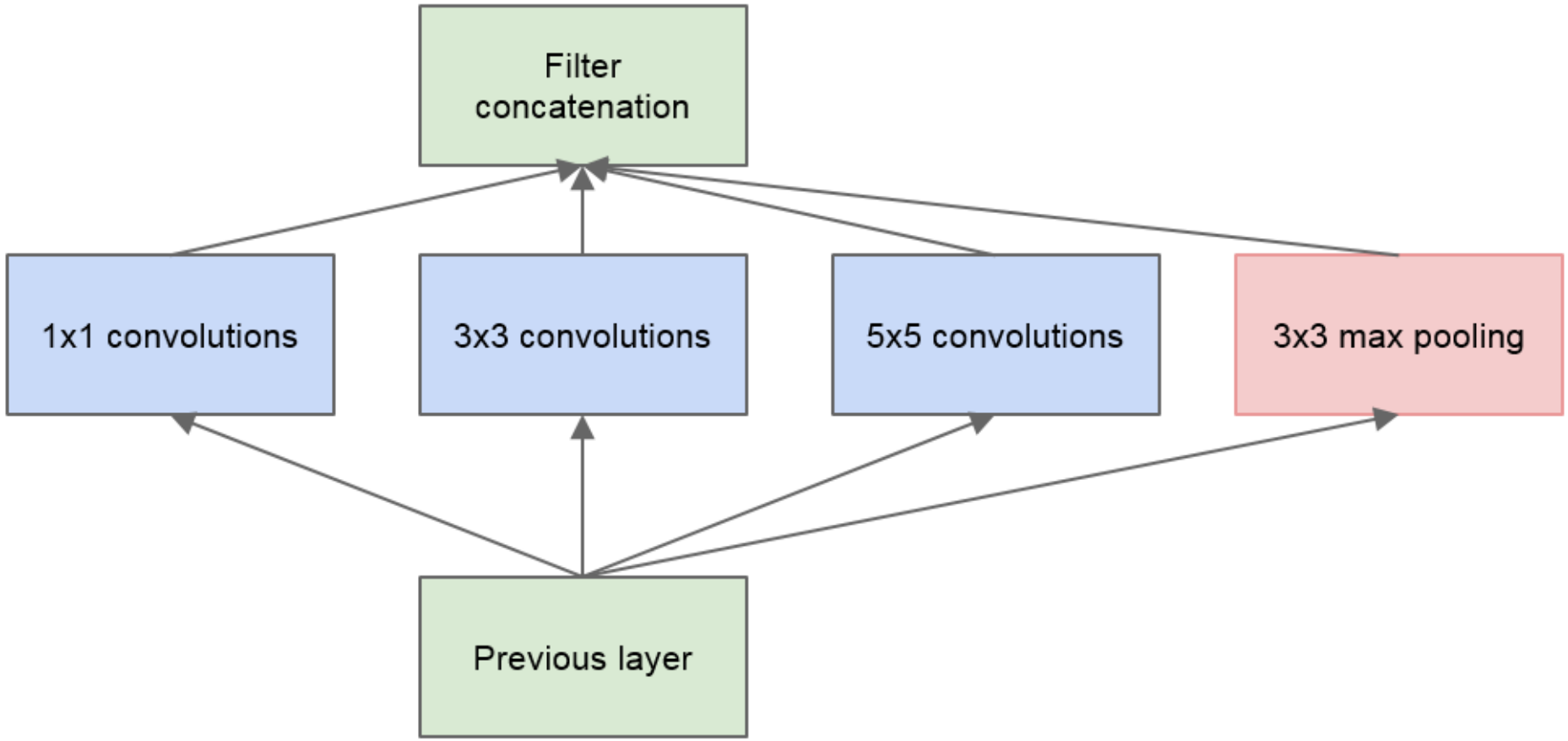
$$\mathbf{x}_\ell = H_\ell([\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}]),$$

Growth rate(k)

- Control the number of feature maps

DenseNet ($k = 12$)	40	1.0M
DenseNet ($k = 12$)	100	7.0M
DenseNet ($k = 24$)	100	27.2M
DenseNet-BC ($k = 12$)	100	0.8M
DenseNet-BC ($k = 24$)	250	15.3M
DenseNet-BC ($k = 40$)	190	25.6M

Bottleneck layers.



Compression

- Improve model compactness
- Descent the parameters
- Transition layers $0 < \theta < 1$

Code

- https://github.com/gpleiss/efficient_densenet_pytorch (pytorch)
- https://github.com/YixuanLi/densenet-tensorflow (tf)