DENSELY CONNECTED CONVOLUTIONAL NETWORKS

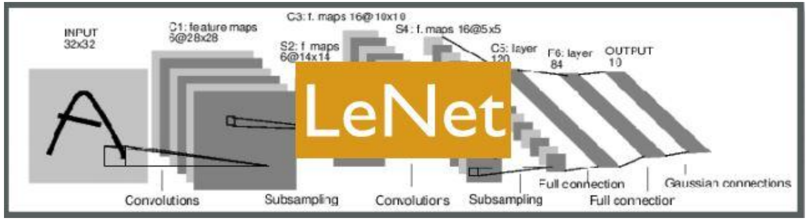
Authors: Gao Huang*, Zhuang Liu*, Laurens van der Maaten, Kilian Q. Weinberger

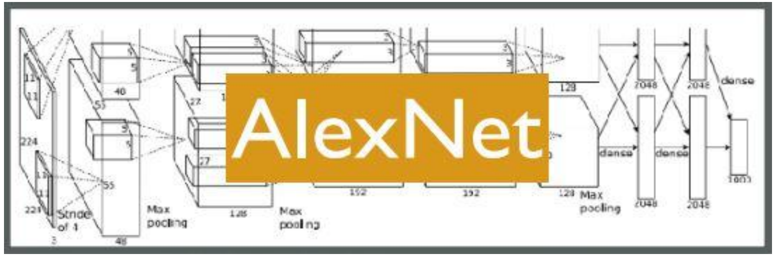
Oral Presentation in CVPR 2017,
Won Best Paper Award

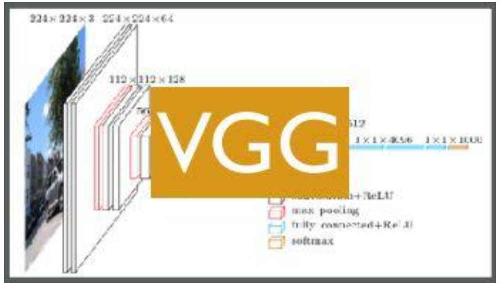
Presented by:
Kartik Thakral
P19CS205

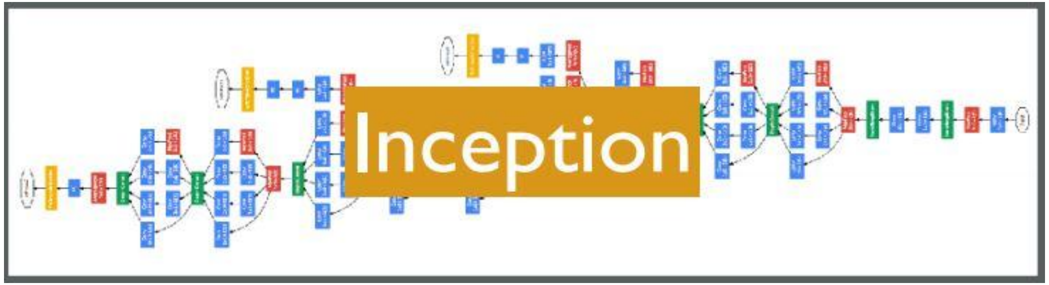
Contents of The Presentation

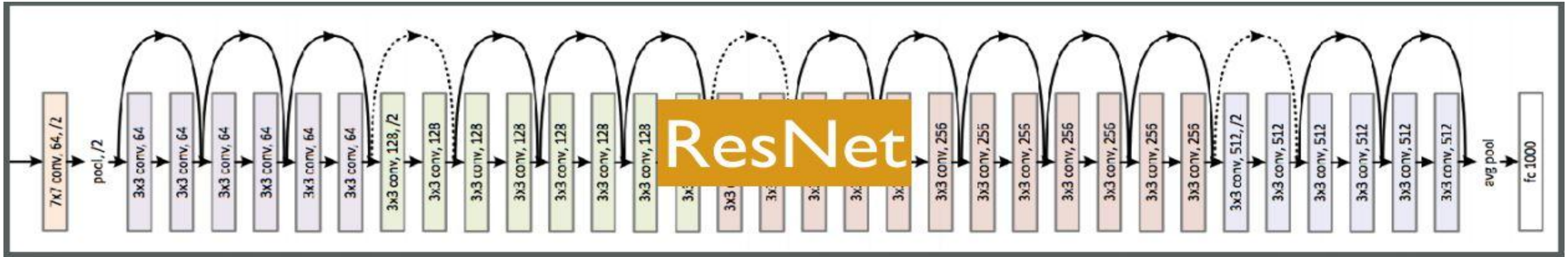
Existing CNN Architectures (Related Work)





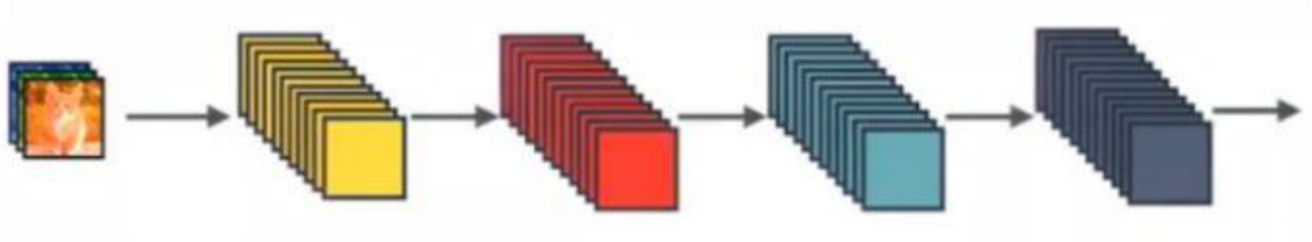






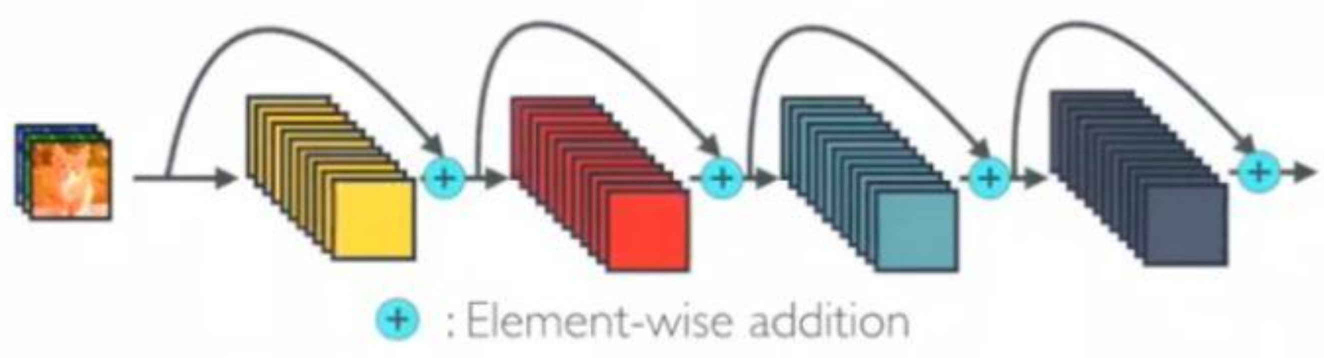
Dense Blocks

Dense Blocks



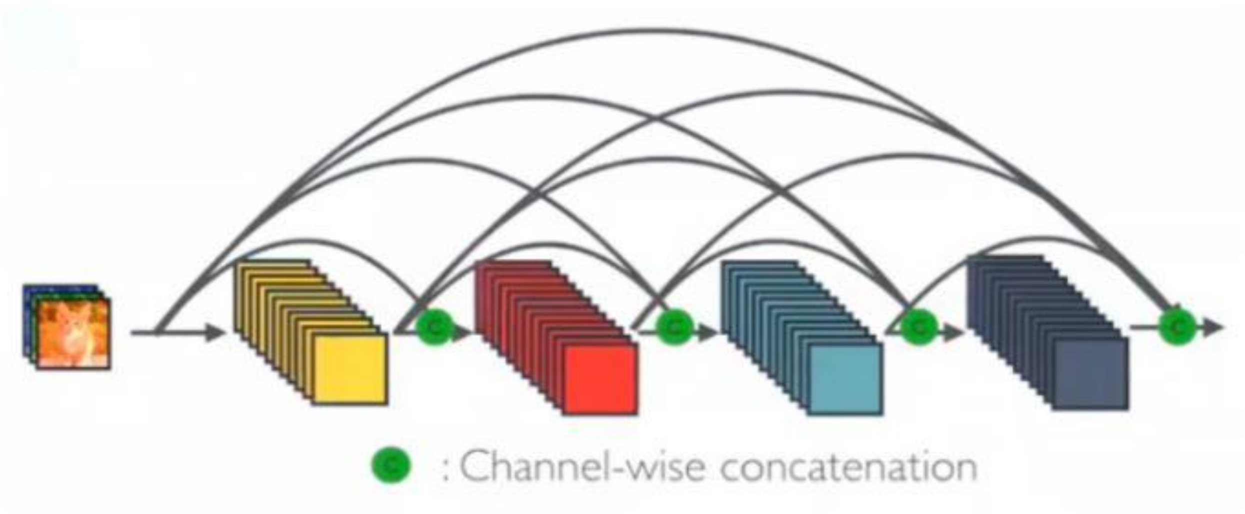
Standard convolution

Dense Block (contd.)



Element-wise addition in ResNet for gradient propagation.

Dense Block (contd.)



Dense - Connections

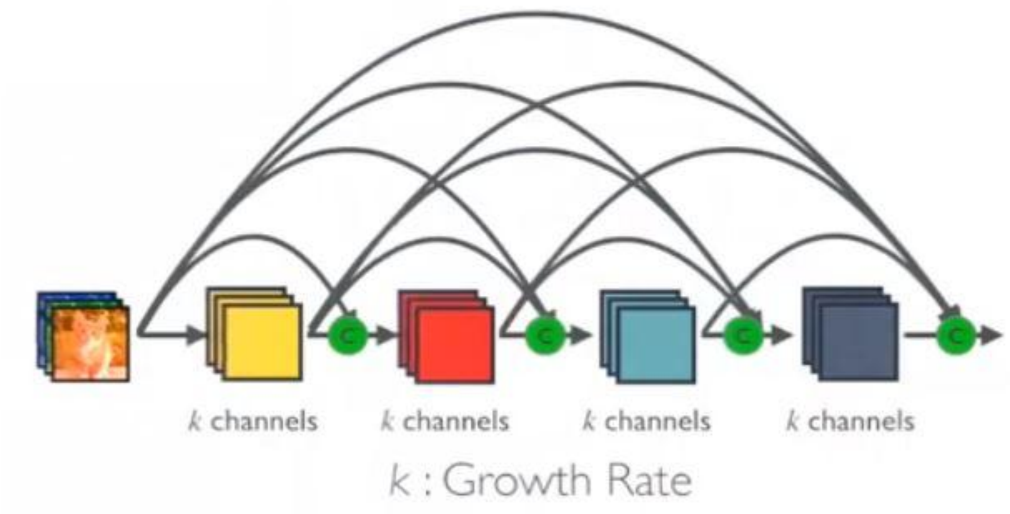
Features from each layer is passed to each subsequent layer

Dense Block (contd.)

Concatenation of feature maps during Forward Propagation

Dense block in action

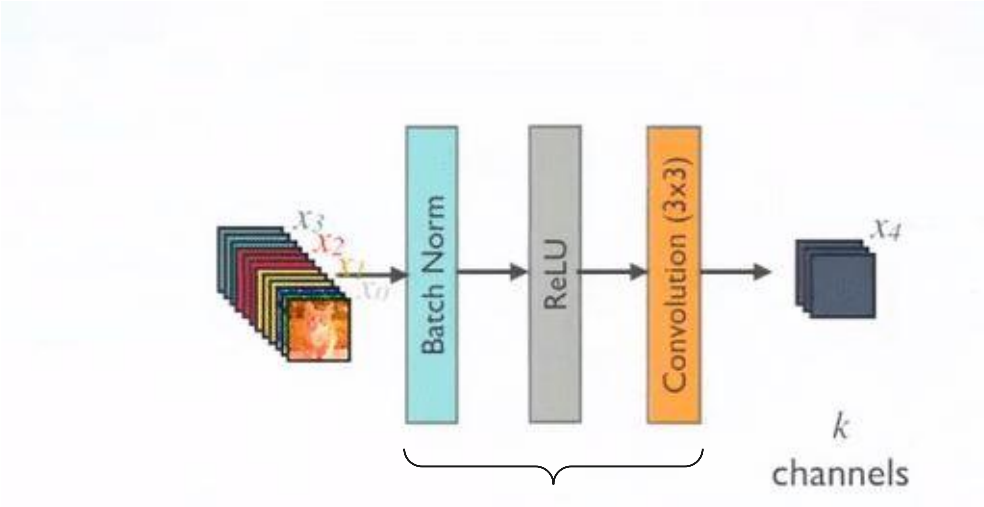
Dense block



Since each layer receives feature maps from all preceding layers, network can be **thinner** and **compact**, i.e. number of channels can be fewer. The growth rate *k* is the additional number of channels for each layer. Through this mechanism, we get higher **computational** and **memory** efficiency.

DenseNet Architecture

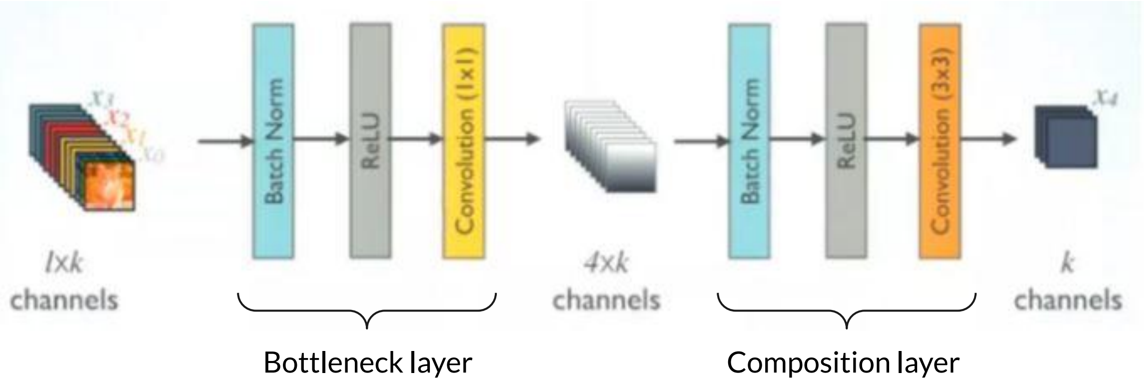
Composition Layers



Composition layer

In each composition layer, Batch Norm (BN) and ReLU, then 3×3 Convolution are done with output feature maps of *k* channels.

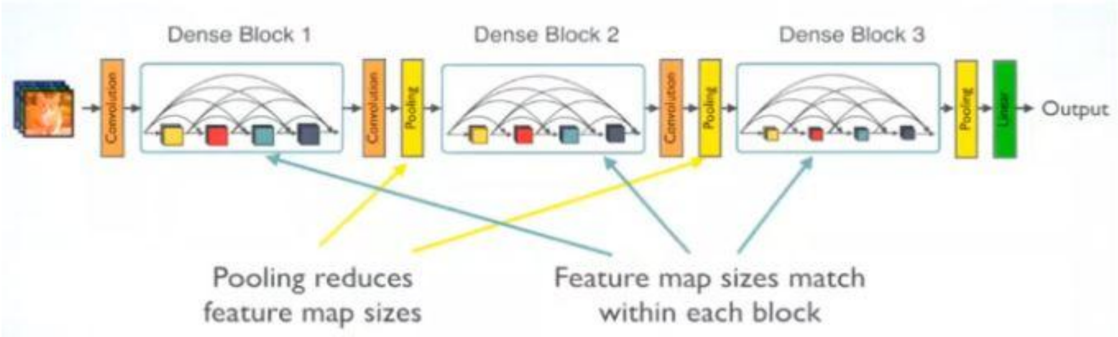
Bottleneck Layers



To reduce the model complexity and size, BN-ReLU-1×1 Conv is done before BN-ReLU-3×3 Conv.

1×1 Conv followed by 2×2 average pooling are used as the transition layers between two contiguous dense blocks.

DenseNet Architecture

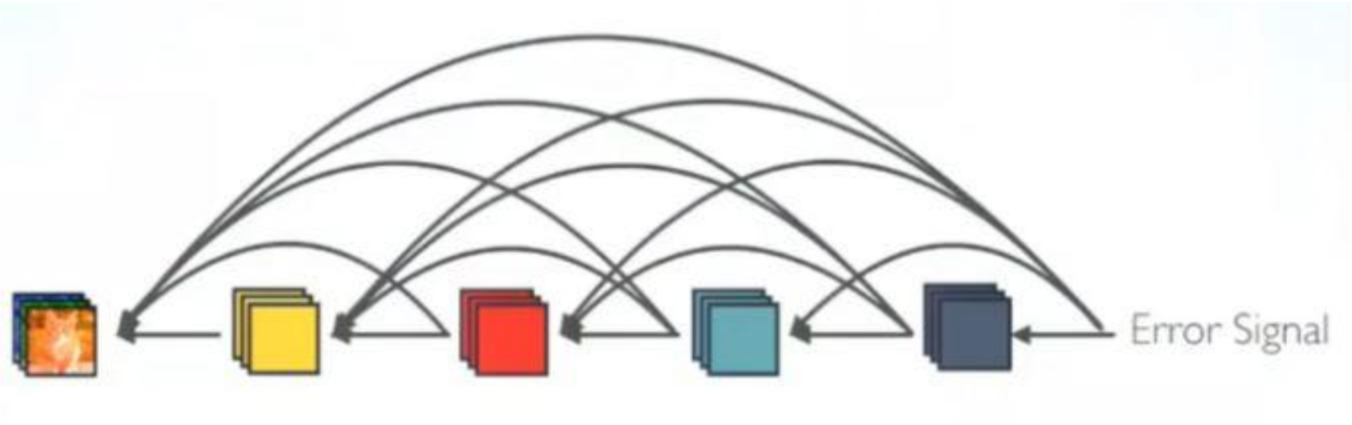


Feature map sizes are the same within the dense block so that they can be concatenated together easily.

At the end of the last dense block, a global average pooling is performed and then a softmax classifier is attached.

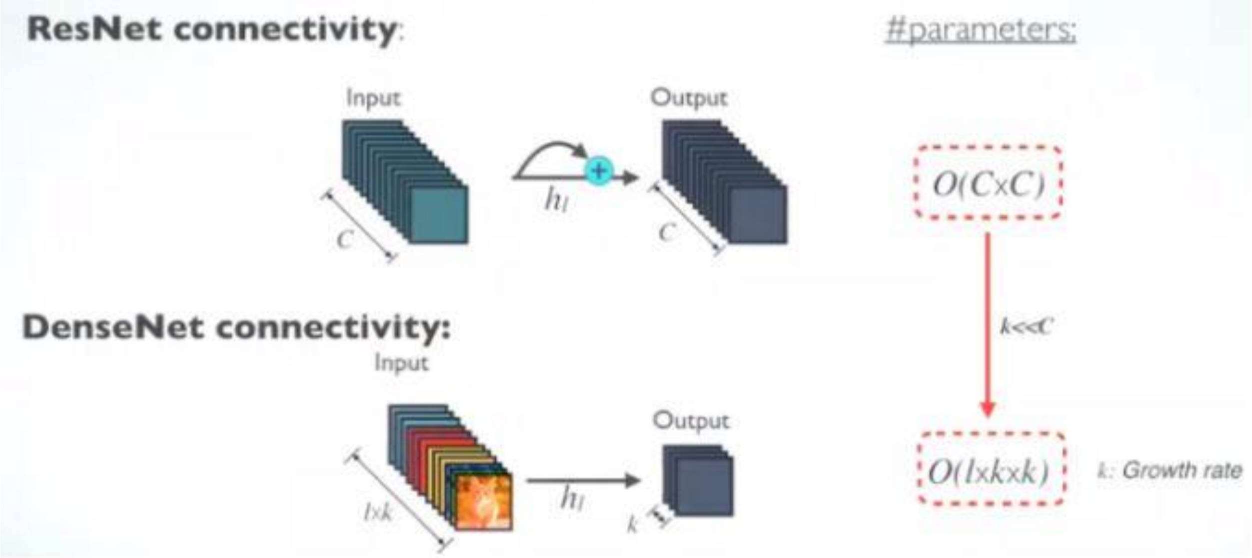
Advantages of the DenseNet architecture

Better Gradient Flow (Implicit Deep Supervision)



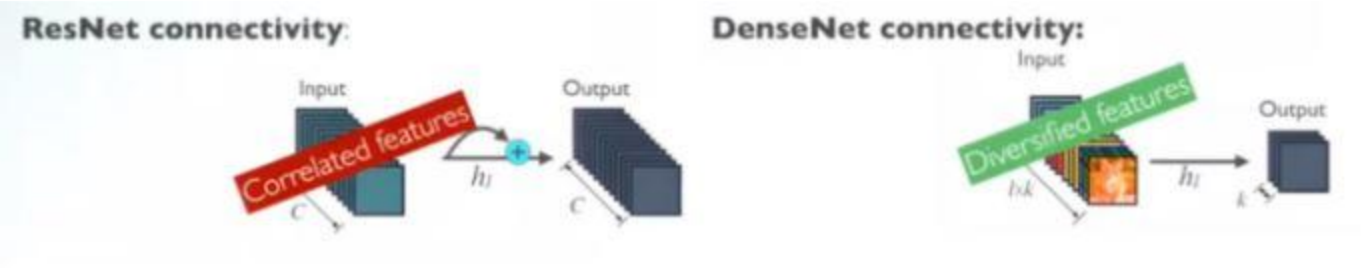
The error signal can be easily propagated to earlier layers more directly. This is a kind of **implicit deep supervision** as earlier layers can get direct supervision from the final classification layer.

Parameter & Computational Efficiency



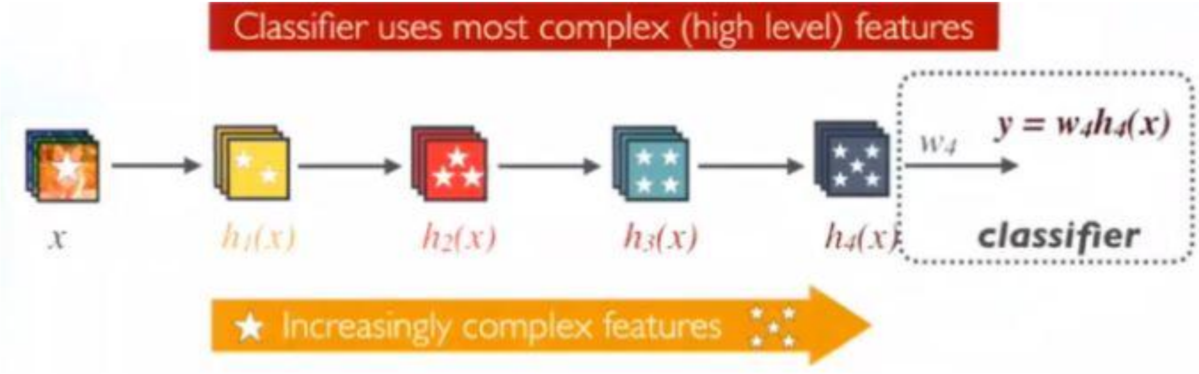
In ResNet number of parameters at each layer is proportional to C×C because features are learnt at each layer whereas in case of DenseNet it is proportional to *l×k×k (where k<<C)* because features from previous layers are reused at each dense layer.

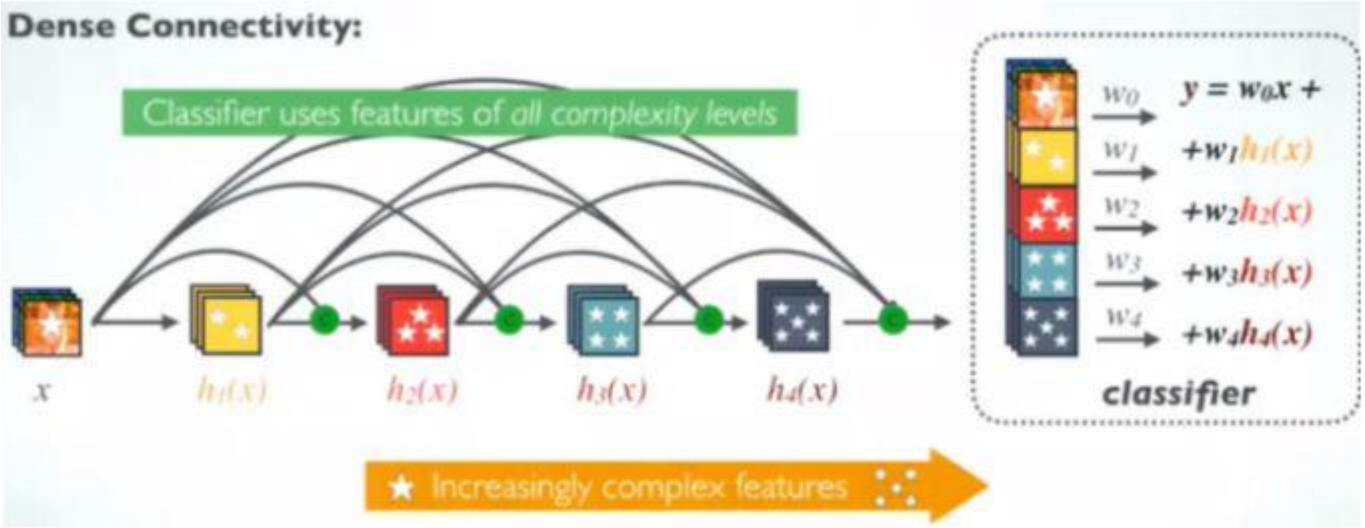
More Diversified Features



Since each layer in DenseNet receive all preceding layers as input, more diversified features and tends to have richer patterns.

Maintains Low Complexity Features



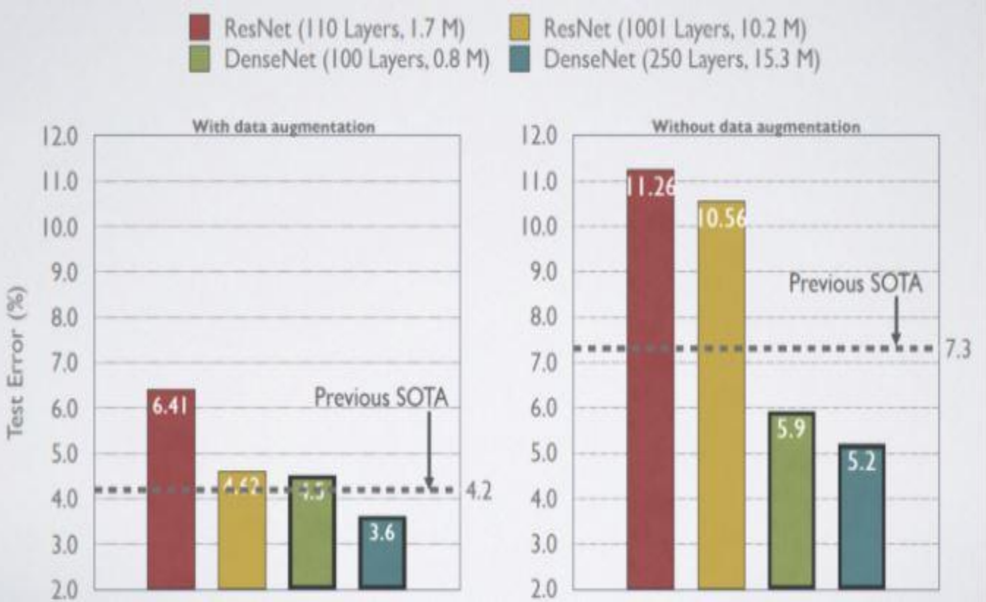


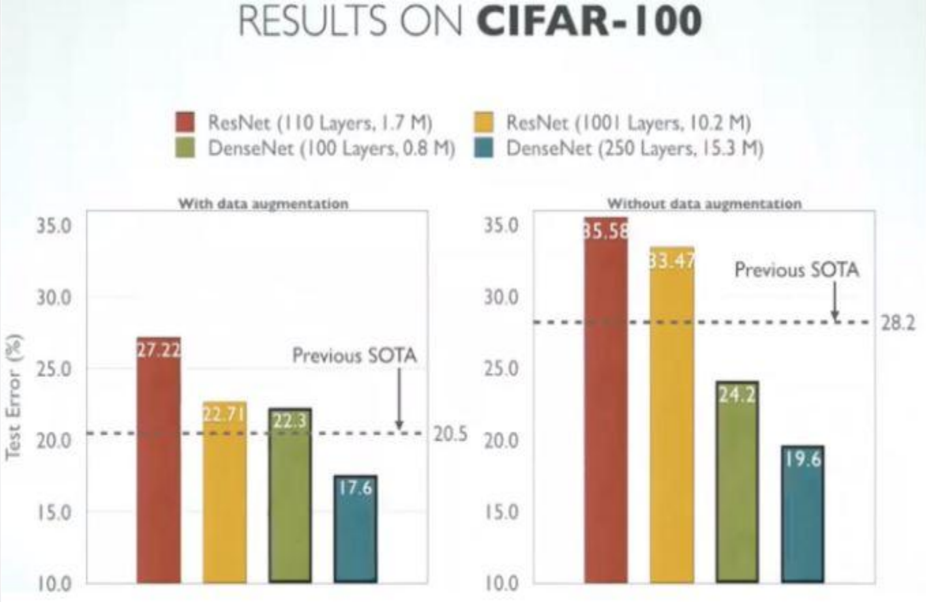
Results

Proposed DenseNet Architectures

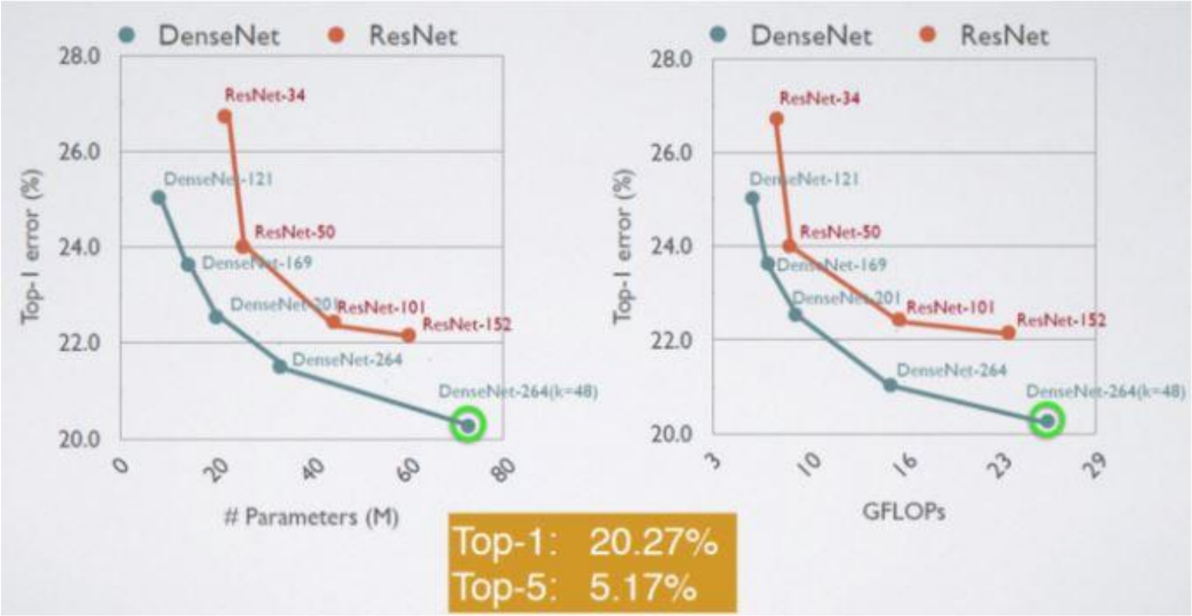
1. Naive DenseNet
   - Simple architecture with growth rate *k*.
2. DenseNet-B
   - DenseNet architecture with Bottleneck layers
   - This model reduced the model complexity and size w.r.t number of parameters.
3. DenseNet-C
   - DenseNet architecture with Compression.
   - If a dense block contains *m* feature-maps, The transition layer generate θm output feature maps, where $0 < \theta \leq 1$ is referred to as the compression factor.
4. DenseNet-BC
   - A combination of above two.
   - Achieved state-of-the-art.

Results on Cifar-10 and Cifar-100 Dataset





Results on ImageNet Dataset



DenseNet-264 (k=48) got the best results of 20.27% Top-1 error and 5.17% Top-5 error.

Analysis and Consolidation

- Features extracted by very early layers are directly used by deeper layers throughout the same dense block.
- Referred as “Collective knowledge”.
- Feature size needs to be same implement Dense Block so, pooling layer cannot be a part of dense layer.
- Each DenseNet architecture has 3-Dense blocks with variable number of layers in each dense block (depending of the architecture).

Problems

- No mathematical support in the paper.
- Why only 3 dense blocks.

References:

- Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).
- https://www.youtube.com/watch?v=-W6y8xnd--U&t=321s
- https://pdfs.semanticscholar.org/c3d9/26a85d85a83126f405ad40ff453611148c15.pdf
- https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803

Thank you

Any questions?