

Densely Connected Convolutional Networks (DenseNet)

Problem: gradient explosion/vanish of very deep CNN

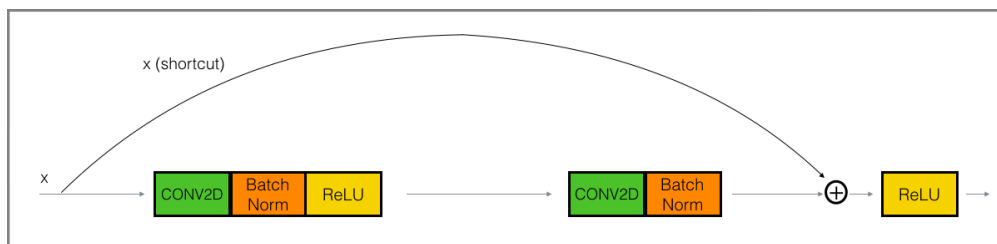
Previous CNN: connection between adjacent layers

ResNet: Skip connection between identity blocks

DenseNet: Dense connection between every pair of layers

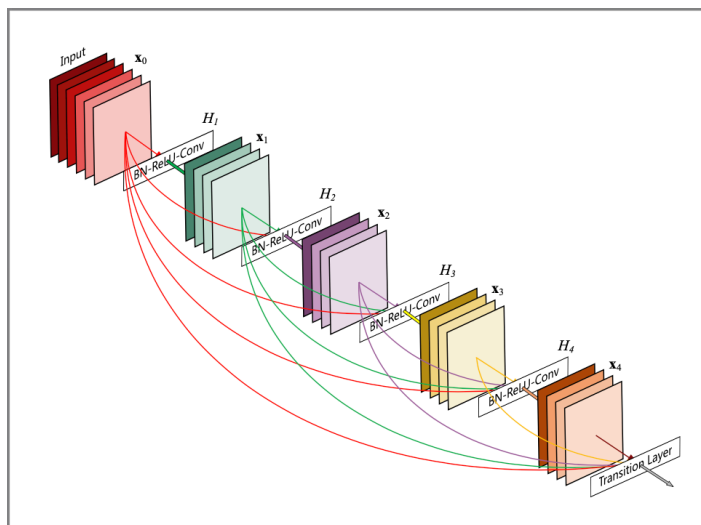
→ Can build very deep CNN without performance degradation

1. ResNet and DenseNet



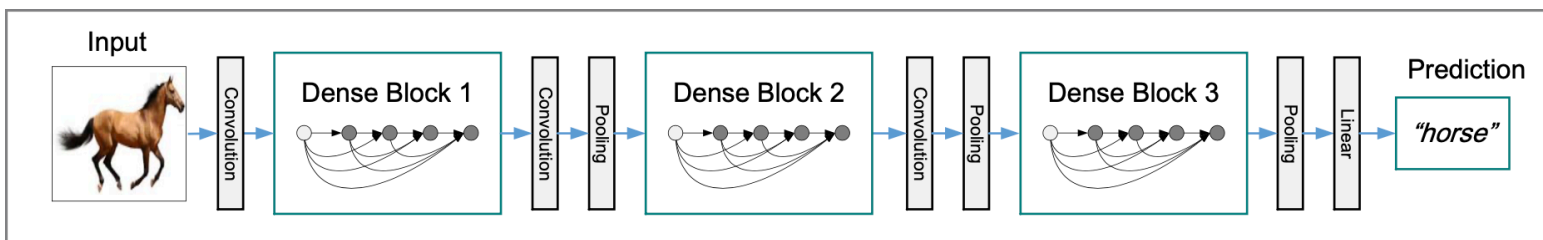
- Additive identity transformation (Impeded network)

Identity Block of ResNet



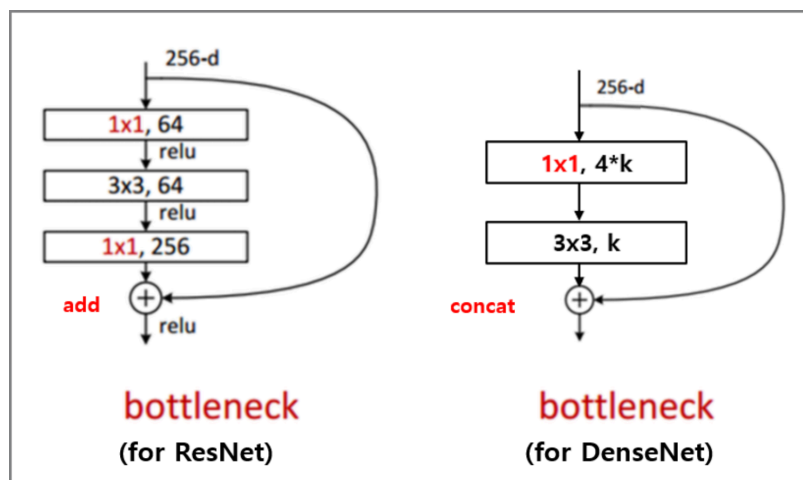
- Concatenate features, presuming every layers' feature-maps
- Feature Reuse
- Remove redundant training
- Fewer parameters
- Improved flow of info. and gradients
- Regularization effect

DenseNet



2. Architecture

- Dense connectivity: $X_l = H_l([X_0, X_1, \dots, X_{l-1}])$
- Growth rate (k): the number of feature maps that H_l produces for each layer
(X_1, \dots, X_{l-1})
 - Can be very narrow ($k = 12$)
- Composite function: BN \rightarrow ReLU \rightarrow 1×1 Conv. \rightarrow BN \rightarrow ReLU \rightarrow 3×3 Conv.
- Bottleneck layer:



- Unlike ResNet, does not reduce the number of feature maps for computational efficiency
- 1×1 Conv. Produces $4k$ feature maps.
- Transition layer: between dense blocks.
 - BN \rightarrow 1×1 Conv. \rightarrow 2×2 average pooling
 - 1×1 Conv. : reduce the number of feature maps
 - 2×2 average pooling: reduce the size of feature maps into half
- Compression: In transition layer, reduce the number of feature maps by hyperparameter θ
 - Set $\theta = 0.5$ (automatically set the number of 1×1 Conv. filter to $\lfloor \theta m \rfloor$)

3. Further Studies

- Bottleneck layer에서 전통적 접근에서 벗어나 growth rate을 이용한 접근이 참신했다. 하지만 왜 $4k$ (오히려 필터의 개수를 늘리는 것이 아닌가?) 개의 feature map을 만들어야하는지에 대한 설명이 부족하다. 나는 이해가 가지 않았다... 추후 다른 논문에서도 이 문제점을 언급하였다고 한다.
 - (<https://arxiv.org/abs/1804.06882>)
- 이론적으로는 작은 k 값 ($k = 12$), 즉 파라미터 개수가 작아도 충분한 성능을 보인다고 했다. 실험결과에 의하면, 실제로 비슷한 성능을 보이는 다른 모델보다 훨씬 적은 파라미터를 사용하는 것을 확인할 수 있다. 하지만, “높은” 성능을 보이려면 결국 $k = 24$, $\text{params} \geq 15M$ 이상이 필요하고, 다른 모델과 비슷한 정도의 파라미터 개수가 필요한 것으로 보인다.
- <https://github.com/liuzhuang13/DenseNet>