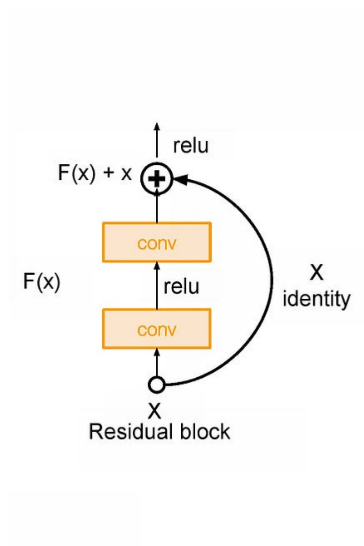
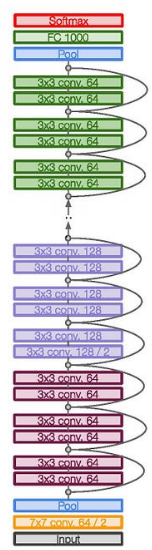
Transfer Learning Image Classification Algorithms – Assignment Week 1

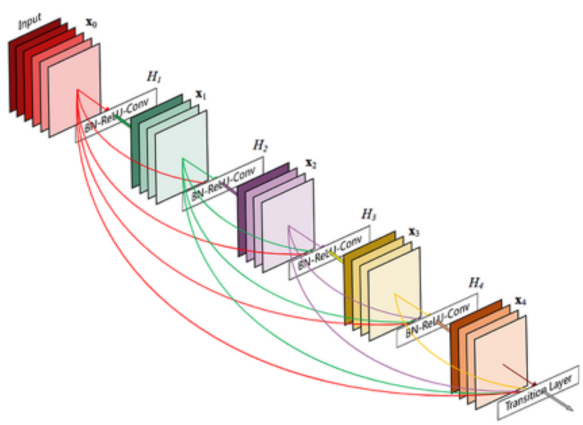
Dimas Wahyu Saputro

1. ResNet
    a. Pros
        i. Excellent performance in deep networks with 100+ layers
        ii. Easy to optimize due to skip connections
        iii. ResNet secured 1st Position in ILSVRC and COCO 2015 competition with just error rate of 3.6% of error rate. (Better than Human Performance !!!)
    b. Cons
        i. Training a ResNet requires a lot of computations (10x more than AlexNet), which means more training time and energy required
        ii. May require more memory and storage
        iii.
    c. High Level Structure
        i. The main base element is the residual block.
        ii. These layers put on top of each other and every layer try to learn some underlying mapping of the desired function and instead of having these blocks, we try and fit a residual mapping.
        iii. Stacked residual blocks

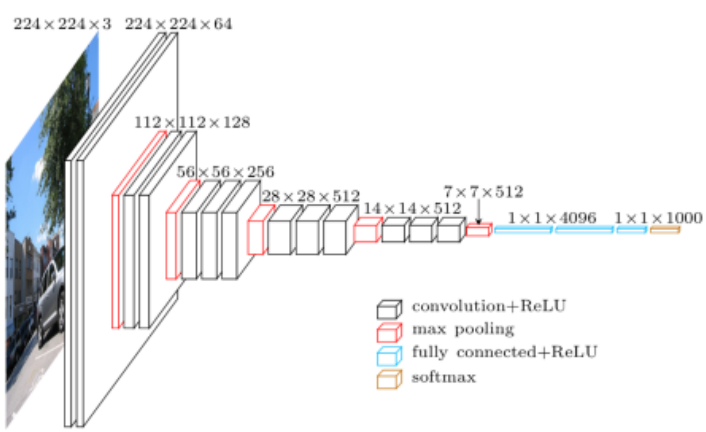


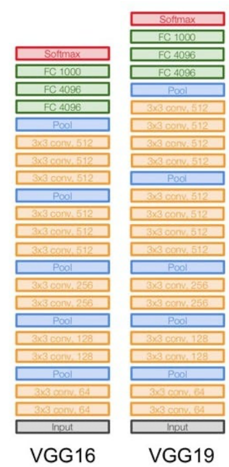


2. DenseNet
    a. Pros
        i. Extreme reuse of residuals creates deep supervision because every layer receives more supervision from the previous layer and thus loss function will react accordingly and due to this methodology, it makes it a more powerful network.
        ii. Strong feature reuse through dense connections
        iii. Parameter efficiency by reducing redundancy
    b. Cons
        i. High memory consumption due to dense connectivity
        ii. Increased computational complexity
    c. High Level Structure
        i. DenseNet is composed of Dense blocks. The layers are densely connected together.
        ii. Each layer gets the input from previous layers output feature maps
        iii. Consist two block: Dense block (Batch Normalization, ReLU Activation, 3x3 Convolution) and Transition Layer (sum of residual will be performed, made of BN, 1x1 Conv, Average Pooling)
        iv. Dense blocks with skip connections
        v. Transition layers for downscaling

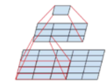


3. VGG
    a. Pros
        i. VGGNet was the runner up of the ImageNet Large Scale Visual Recognition Challenge(ILSVRC) classification the benchmark in 2014.
        ii. Very simple structure of 3 x 3 convs with the periodic pooling all the way through the network
        iii. Easy to understand and implement
    b. Cons
        i. Large number of parameters
        ii. its deployment on even the most modest sized GPUs is a problem because of huge computational requirements, both in terms of memory and time
    c. High Level Structure
        i. Stacked convolutional layers
        ii. VGG 16 is 16 layer architecture with a pair of convolution layers, poolings layer and at the end fully connected layer.
        iii. VGG network is the idea of much deeper networks and with much smaller filters.
        iv. One key thing is that these models kept very small filters with 3 x 3 conv all the way, which is basically the smallest conv filter size that is looking at a little bit of the neighbouring pixels
        v. VGG used small filters because of fewer parameters and stack more of them instead of having larger filters (VGG has smaller filters with more depth instead of having large filters. It has ended up having the same effective receptive field as if you only have one 7 x 7 convolutional layers.)

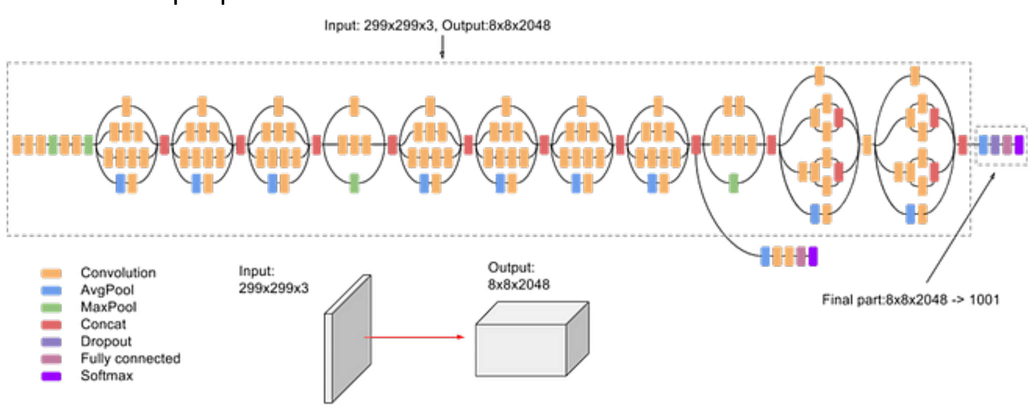


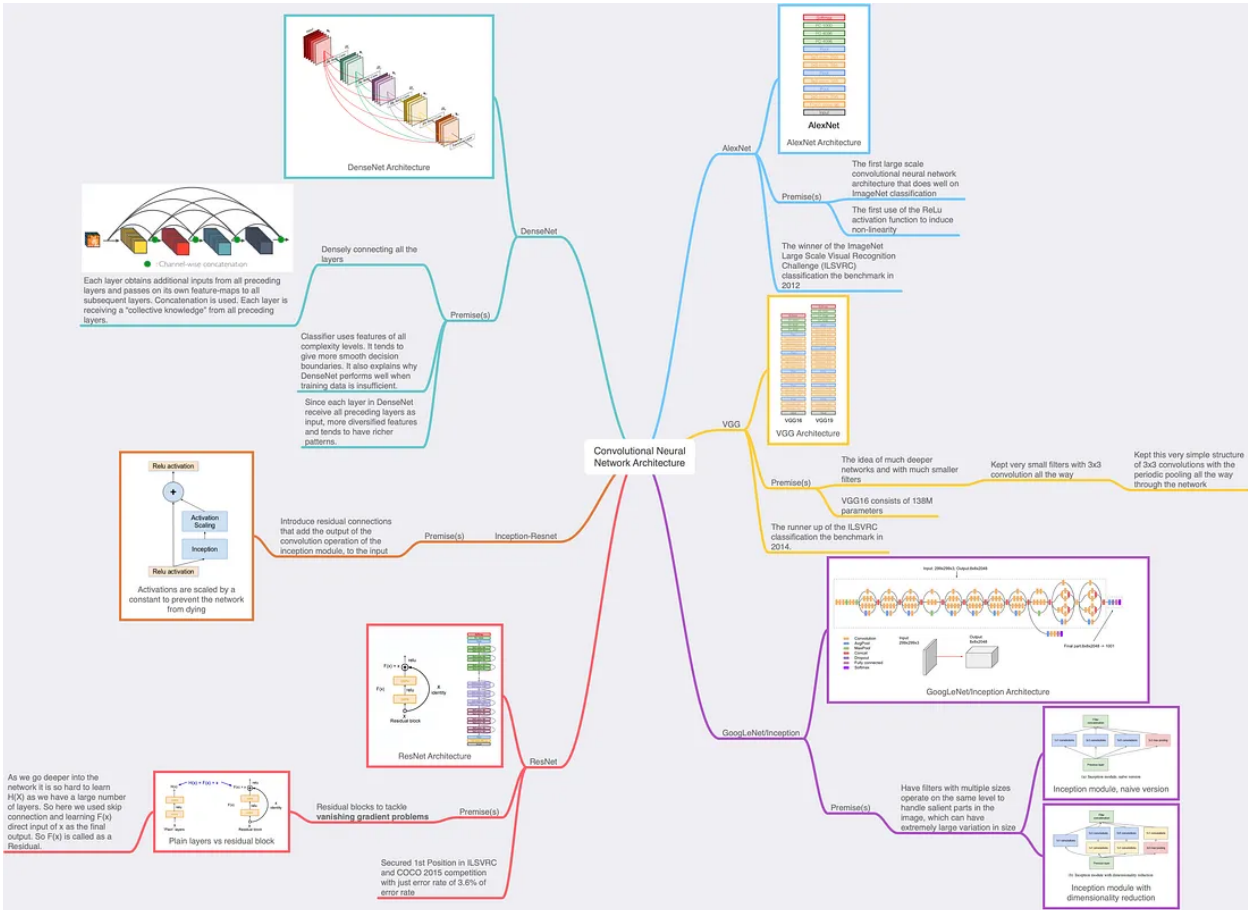


4. Inception
   a. Pros
      i. Inception work with Factorizing Convolutions. Factorizing Convolutions used to reduce the number of connections and parameters to learn. This will increase the speed and gives a good performance.



      ii. With 42 layers deep, the computation cost is only about 2.5 higher than that of GoogleNet and much more efficient than that of VGGNet
      iii. Efficient use of computational resources
   b. Cons
      i. Complex architecture design
      ii. Increased memory requirements
   c. High Level Structure
      i. Made up of symmetric and asymmetric building block, including convolutions, average pooling, max pooling, dropouts, and fully connected layer.
      ii. Batchnorm is used extensively
      iii. Loss is computed via softmax
      iv. Multiple parallel convolutional branches





Reference

1. https://towardsdatascience.com/architecture-comparison-of-alexnet-vggnet-resnet-inception-densenet-beb8b116866d
2. https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/
3. https://medium.com/@andikarachman/convolutional-neural-network-architecture-the-ultimate-summary-977c998f75d3
4. https://towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96