**Review of 2nd One**

**VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION**

In this paper, the authors looked at how the depth of a convolutional network affects its accuracy in large-scale picture recognition. Their primary addition was a thorough evaluation of increasing depth networks using an architecture with very small ( 3 3) convolution filters, demonstrating that increasing the depth to 16-19 weight layers results in a substantial improvement over prior-art configurations. These results was served as the foundation for their ImageNet Challenge 2014 submission, which earned them first and second place in the localisation and classification categories, respectively. They also demonstrate that their representations generalize well to other datasets, achieving state-of-the-art outcomes.

During training, Their ConvNets were fed a fixed-size 224 224 RGB picture. They only preprocess each pixel by subtracting the mean RGB value computed on the training group. The image was passed through a stack of convolutional (conv.) layers, with filters with a very tiny receptive field: 3 3 (the smallest size to catch the notions of left/right, up/down, and center). They also used (1x1) convolution filters in one of the setups, which could be thought of as a linear transformation of the input channels. (followed by non-linearity). The stride of the convolution was set at one pixel; the spatial padding of conv. The layer input was chosen in such a way that the spatial resolution was maintained after convolution, i.e. the padding was one pixel for three convolution layers. Five max-pooling layers followed some of the convolutional layers to perform spatial pooling (not all conv. layers were succeeded by max-pooling). Max-pooling was done with stride 2 over a (2x2) pixel frame.