

The paper represents a deep convolutional neural network method to classify high resolution image. The method uses eight convolutional layers with first five are convolutional and the rest are fully connected. But more explanation is expected that why the authors apply such architecture.

The net classifies 1000 class by maximizing the multinomial logistic regression. For the output of convolutional layers, ReLU nonlinearity is applied to reduce the training time. A overlapping pooling layer is used to summarized the output of neighborhoods neurons. The purpose is to find the most responsive nodes by max pooling. The major challenge in classification is to avoid overfitting. One method is to increase the labelled dataset. Transformed images are made from the original images by translations and horizontal reflections. Also PCA is applied on the RGB pixel values of the training set. The found principle components are added to each training image, which intensifies the important information in a image. Each hidden neuron has the 0.5 possibility to drop out. It avoids that multiple neuron contribute to same features and makes the model more robust.

The learning rate and gradient descent rate are both magic numbers. The advantage of CNN is that the model handles the most parameters without calculating approximation functions. Details are hidden in the neuron layers.