1. Stride is a way of decreasing parameters and reducing some of the side effects at the same time. In the paper, the authors simply assumed that the next layer's node has lots of overlaps with their neighbors by looking at the regions and they can manipulate the overlap by controlling the stride.

2. One of the drawbacks of the convolution step is the loss of information that might exist on the border of the image. Because they are only captured when the filter slides, they never have the chance to be seen. To resolve the issue, the author proposed to use zero-padding. The other benefit of zero padding is to manage the output size. This padding idea helps us to prevent network output size from shrinking with depth. Therefore, it is possible to have any number of deep convolutional networks.

3. Convolution layer is the most important layer in CNN. It takes most of the time within the network. Network performance also depends on the number of levels within the network. But on the other hand as the number of levels increases the time required to train and test the network.

4. Max-pool is one of the most common types of pooling methods. The main idea of pooling is down-sampling in order to reduce the complexity for further layers. It partitions the image to sub-region rectangles, and it only returns the maximum value of the inside of that sub-region. One of the most common sizes used in max-pooling is 2×2. When pooling is performed in the top-left 2×2 blocks, it moves 2 and focuses on the top-right part. This means that stride 2 is used in pooling. To avoid down-sampling, stride 1 was used. Since down-sampling does not preserve the position of the information, it should be applied only when the presence of information is important (rather than spatial information).

5. CNN has several advantages over traditional neural networks. The most beneficial aspect of CNNs is reducing the number of parameters in ANN. This achievement has prompted both researchers and developers to approach larger models in order to solve complex tasks, which was not possible with traditional neural networks. The most important assumption about problems that are solved by CNN should not have features which are spatially dependent. In other words, in a face detection application, we do not need to pay attention to where the faces are located in the images. The only concern is to detect them regardless of their position in the given images. Another important aspect of CNN, is to obtain abstract features when input propagates toward the deeper layers.