

Hence, the neurons in the visual cortex do not all sense the entire image at the same time, but they are activated by viewing a local area of the image via its local receptive field.

In Figure 35-4, the local receptive fields overlap to give a collective perspective on the entire image. Each neuron in the visual cortex reacts to a different type of visual information (e.g., lines with different orientations).

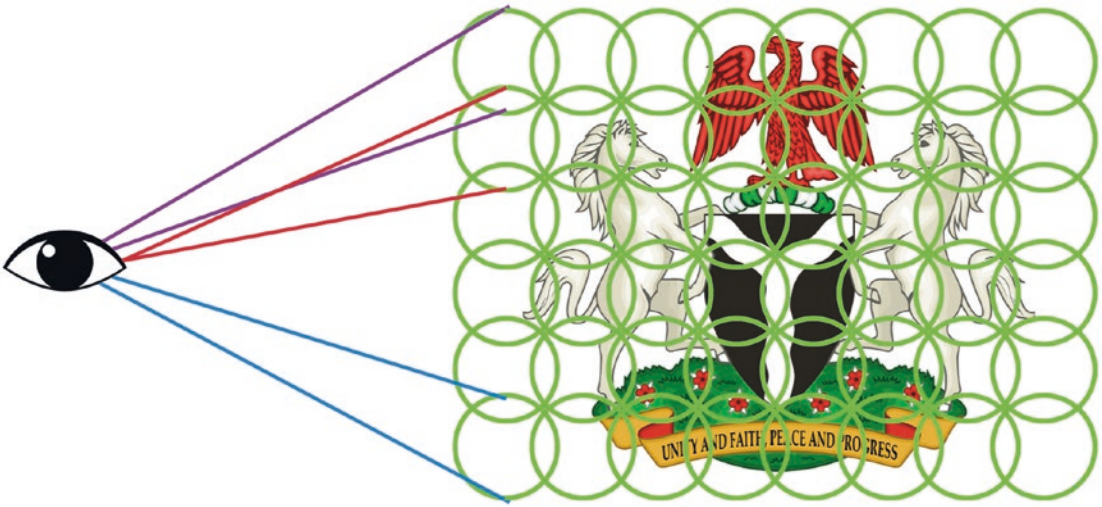


Figure 35-4. *Local receptive field*

Other neurons have large receptive fields that react to more complex visual patterns such as edges, regions, and so on. From here we get the idea that neurons with larger receptive field receive information from those with lower receptive fields as they progressively learn the visual information of the image.

Advantages of CNN over MLP

Suppose we have a 28 x 28 pixel set of image data, a feedforward neural network or multilayer perceptron will need 784 input weights plus a bias. By flattening an image as you would in MLP, we lose the spatial relationship of the pixels in the image.

CNN, on the other hand, can learn complex image features by preserving the spatial relationship between the image pixels. It does so by stacking convolutional layers whereby the neurons in the higher layers with a larger receptive field receive information

from neurons in the lower layers having a smaller receptive field. CNN learns a hierarchy of increasingly complex features from the input data as it flows through the network.

In CNN, the neurons (or filters) in the convolutional layer are not all connected to the pixels in the input image as we have in the dense multilayer perceptron. Hence, a CNN is also called a sparse neural network.

A distinct advantage of CNN over MLP is the reduced number of weights needed for training the network.

Convolutional neural networks are composed of three fundamental types of layers:

- Convolutional layer
- Pooling layer
- Fully connected layer

The Convolutional Layer

The convolution layer is made up of filters and feature maps. A filter is passed over the input image pixels to capture a specific set of features in a process called convolution (see Figure 35-5). The output of a filter is called a feature map.

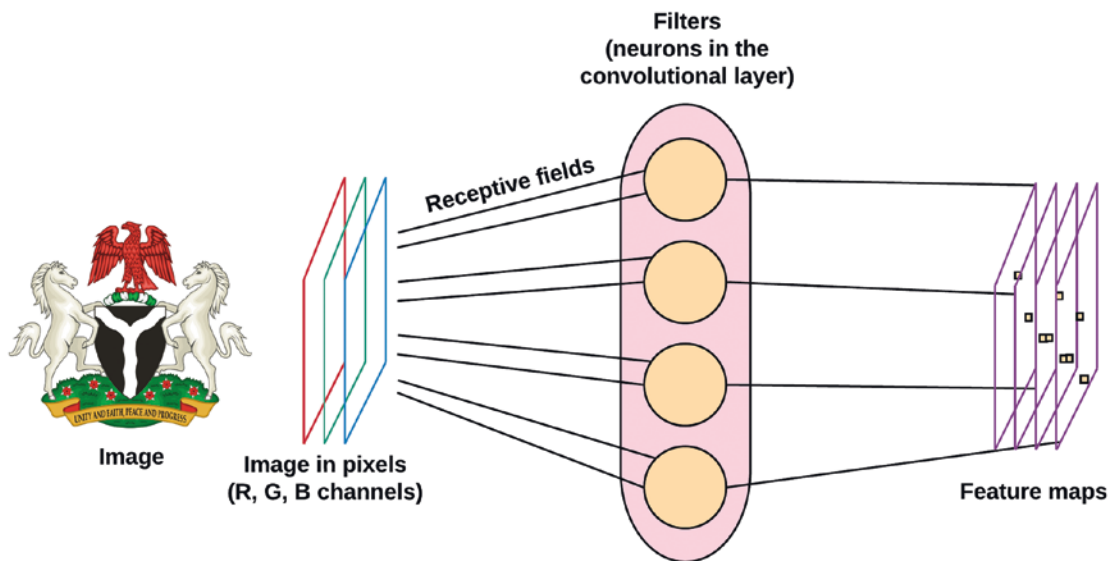


Figure 35-5. *The convolution process*