Assignment 4 solution

1. R: $\begin{bmatrix}9 & 9\\ 9 & 9\end{bmatrix}$ G: $\begin{bmatrix}18 & 18\\ 18 & 18\end{bmatrix}$ B: $\begin{bmatrix}18 & 18\\ 27 & 27\end{bmatrix}$ Final: $\begin{bmatrix}45 & 45\\ 54 & 54\end{bmatrix}$

2. R: $\begin{bmatrix}4 & 6 & 6 & 4\\ 6 & 9 & 9 & 6\\ 6 & 9 & 9 & 6\\ 4 & 6 & 6 & 4\end{bmatrix}$ G: $\begin{bmatrix}8 & 12 & 12 & 8\\ 12 & 18 & 18 & 12\\ 12 & 18 & 18 & 12\\ 8 & 12 & 12 & 8\end{bmatrix}$ B: $\begin{bmatrix}6 & 9 & 9 & 6\\ 12 & 18 & 18 & 12\\ 18 & 27 & 27 & 18\\ 14 & 21 & 21 & 14\end{bmatrix}$

   Final: $\begin{bmatrix}18 & 27 & 27 & 18\\ 30 & 45 & 45 & 30\\ 36 & 54 & 54 & 36\\ 26 & 39 & 39 & 26\end{bmatrix}$

3. R: $\begin{bmatrix}4 & 4\\ 4 & 4\end{bmatrix}$ G: $\begin{bmatrix}8 & 8\\ 8 & 8\end{bmatrix}$ B: $\begin{bmatrix}12 & 12\\ 8 & 8\end{bmatrix}$ Final: $\begin{bmatrix}24 & 24\\ 20 & 20\end{bmatrix}$

4. When we do the convolution, we product the filter with image. When filter resemble with the image, we expect a high response. The network is trying to find matches in the image.

5. When pooling between layers (or when using convolution with a stride greater than 1), the spatial dimensions are sampled and so we get an image pyramid with different spatial resolution at the different layers. In this way a fixed size convolution filter covers a larger spatial region in upper layers.

6. Increase the number of filters layer by layer.
   The purpose is to help in learning more levels of global abstract structures and shrink the feature space for input to the dense networks.

7. 126x126x16

8. 63x63x16

9. It can be reduced by the fewer number of filters.

10. The convolution layers are used to extract image pattern or template.

    Early convolution layers: extract simple pattern such as edge.

    Deeper convolution layers: extract more complex pattern.

11. R: $\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$ G: $\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$ B: $\begin{bmatrix}2 & 2\\ 4 & 4\end{bmatrix}$

12. Down sampling spatial dimension.

13. Purpose is for better generalization.

    It is most useful when the data is small.

14. The goal of transfer learning is to improve learning in the target task by leveraging knowledge from the source task.
    It is most useful when the target data is similar to the source data.

15. Avoid weights destruction by gradient from untrained fully connection layers.

16. 1) add customized layer on top of pretrained layer
    2) freeze the pretrained layer
    3) train the customized layer
    4) unfreeze the pretrained layer
    5) train the entire network

17. The inception blocks having filters with multiple sizes operate on the same level which can detect different size variation in the location of the information.

18. 1) With zero weights, the network computes the identity.
    2) Identity connections provide useful feedback throughout the network.
    3) While increasing network depth, it avoids negative outcomes.

19. 1) create a new model from existing model with new output
    2) use the model class instead of sequential class

    It can help us understand what convolution layers are doing.

20. Using gradient decent to find the input that will maximize the response of the filter. Or, using gradient descent to find the input that will minimize the loss at the filter.

    It can help us in interpretation filters and know what filters learn during training.

21. 1) Feed image to the network.
    2) Compute the gradient of selected output note with respect of each channel of target layer.
    3) Compute the average gradient of each channel weight.
    4) Add the activations of each channel weighted by average gradient magnitude.
    5) Super impose the activations on input image.

A higher weight is given to channels with higher gradients. And a higher gradient means the solution is more sensitive to this channel.

The purpose is to find out which parts of image contributed to the solution.