HARISH GANESAN

UB Person No: 50249122

Computer Vision and Image Processing
Homework 1

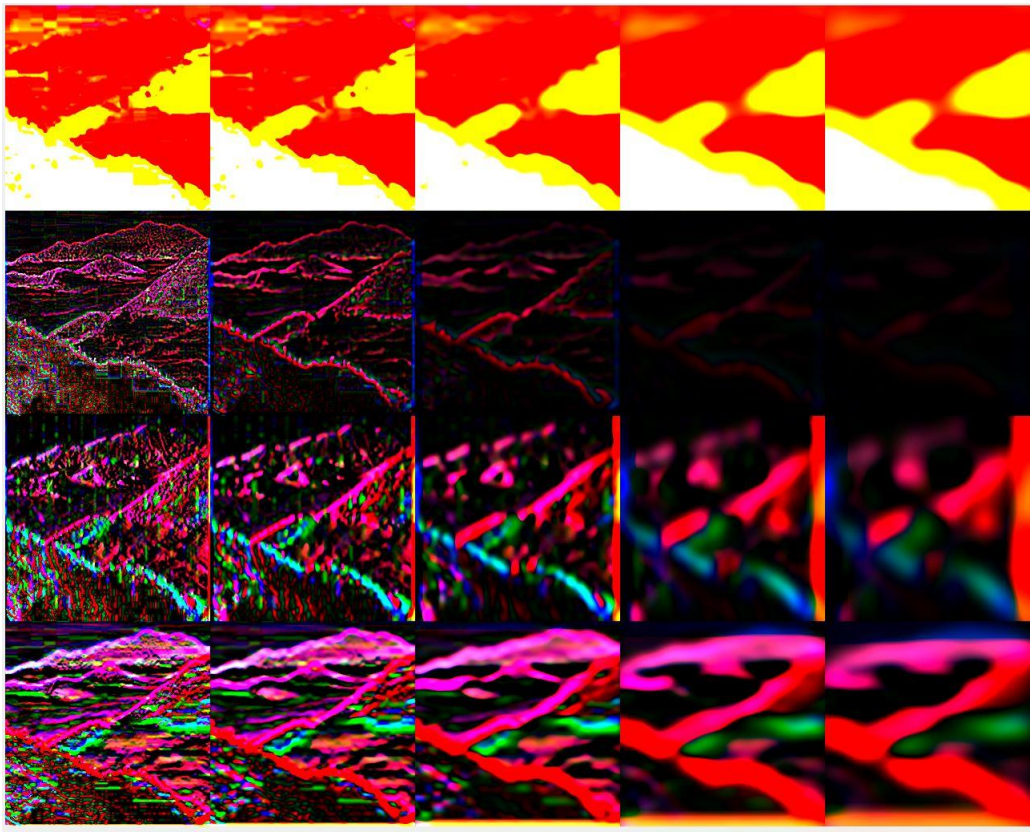
Spatial Pyramid Matching
and Scene Classification

1.0 There are 20 filters in total, and they are of broadly four categories:

- Gaussian – this type of filter is used to obtain a filtered image using a 2-D lowpass Gaussian smoothing kernel, it is used for blurring the image to remove the noise and yet preservse high frequency components
- Laplacian of Gaussian (LoG) – this filter is obtained by calculating the Laplacian of a Gaussian filter, it is used for edge detection as it highlights regions of rapid intensity change
- Oriented Gaussian in X Direction – this filter is used to detect any horizontal edges or lines in the images
- Oriented Gaussian in Y Direction – this filter is used to detect any vertical edges or lines in the images

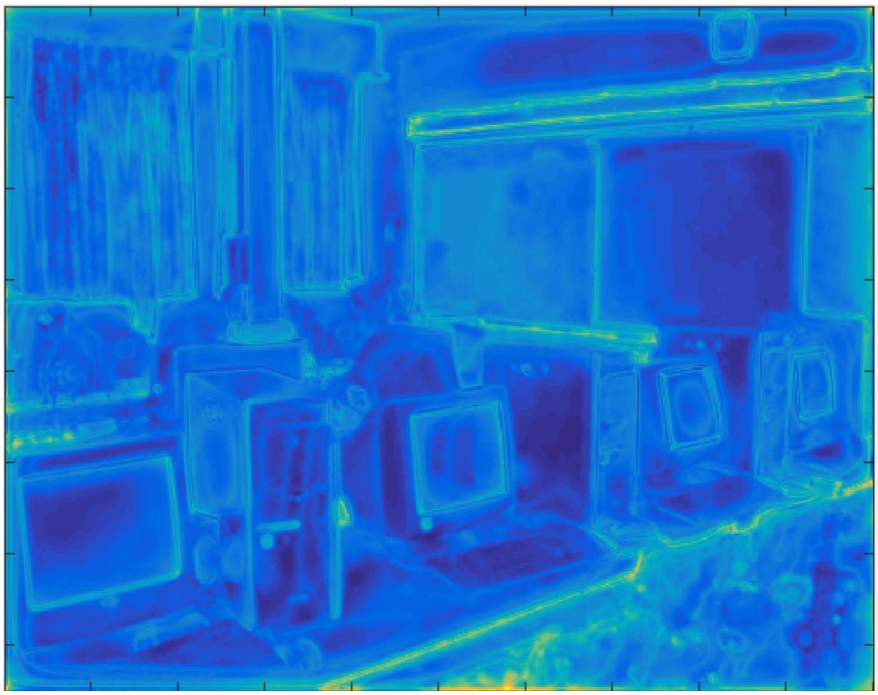
1.1 For the original image given below :

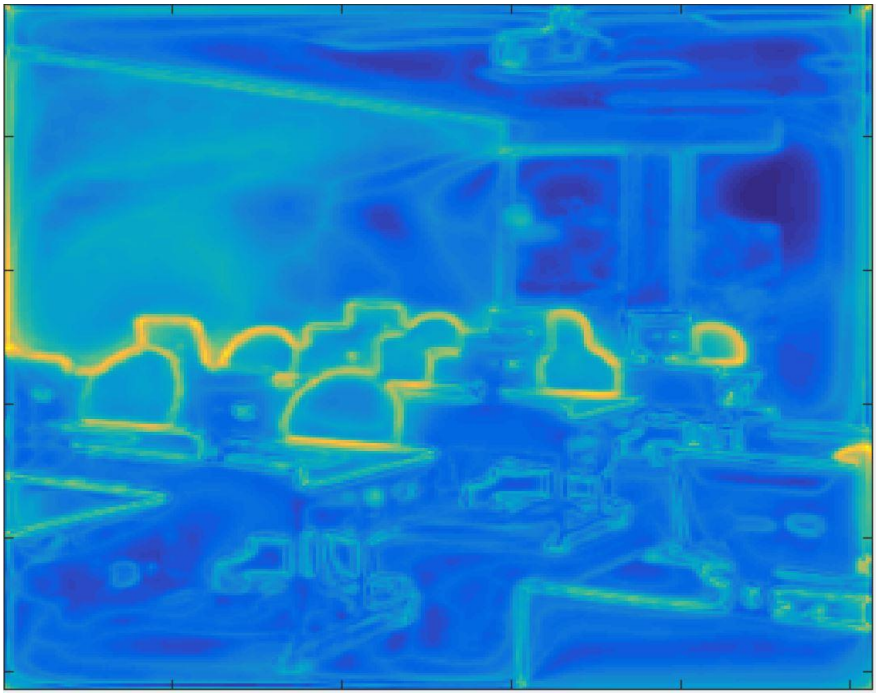
The montage of all the filters applied on the image is as follows :

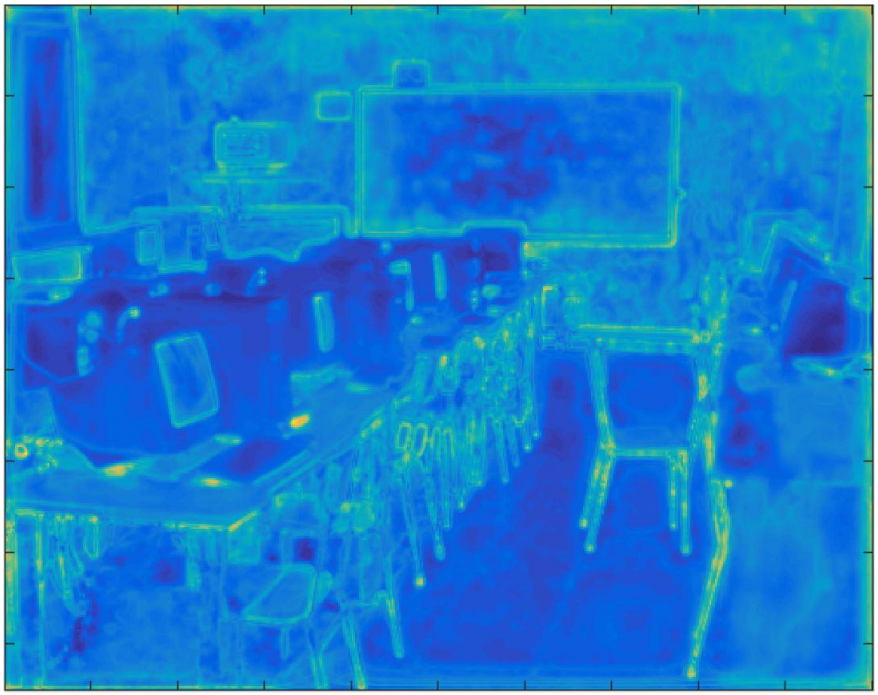


1.2 File provided in the .zip

1.3 Three images from the category ‘computer-room’ have been picked and their word-maps are shown below:







As we can see from the word maps shown above, the edges of the objects present in the image have been highlighted. This includes edges of two types – horizontal and vertical. It also determines where the colour is similar and where it is changing. It also differentiates light intensity in the different areas in the image.

2.1 – 2.4 present in the .zip file

2.5 The confusion matrix is as follows :

conf

8x8 double

	1	2	3	4	5	6	7	8
1	9	2	1	1	6	0	0	1
2	1	6	2	3	6	0	0	2
3	1	1	13	2	3	0	0	0
4	3	2	2	12	0	0	0	1
5	0	2	6	2	9	0	0	1
6	1	3	4	1	1	3	7	0
7	0	4	3	0	2	3	8	0
8	2	5	4	1	4	0	0	4

Upon summing the diagonal elements, the number of correct predictions = 9 + 6 + 13 + 12+ 9 + 3 + 8 + 4 = **64**.

The total number of test images are **160**.

Therefore the accuracy percentage is **(64 / 160**) * 100 = **40 %**

2.6 The three classes with the lowest accuracy are 6, 8 and 2, which are ‘mountain’, ‘tennis_court’ and ‘computer_room’ respectively.

- It is likely that the ‘mountain’ class is harder to classify because of its similarity to the ‘garden’ samples due to both the images containing a lot of the similar objects and colors and both being in a outdoor setting.
- The class ‘tennis_court’ is also being mis-classified as ‘garden’ because of the similar colous and due to there being not much variety in the training images.
- The samples in ‘computer_room’ are mostly being mis-classified as class ‘library’ as they are both in an indoor setting and share similar elements in the images.

2.7 Performance can be improved if the values of Alpha are raised from the current values of alpha = 100 and k = 250 to some larger values. It can also possibly be improved if we obtain the filter responses to the image once the random-pixels have been picked from the images. We can also try new filters such as ‘laplacian’ or ‘average’.