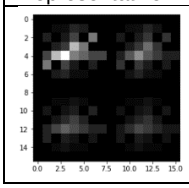
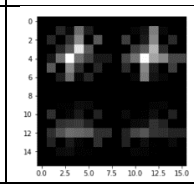
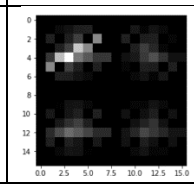
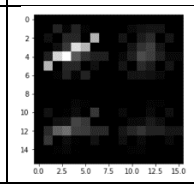
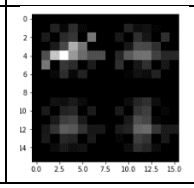
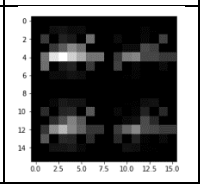
Bag of Words using Histogram of Oriented Gradients :

Methodology :

1. The train-test dataset is loaded and the RGB channel is converted to gray channel.
2. For each image in training dataset, the images are divided into patches. For creating patches we have use grid method and none of the patch is overlapping. The 32x32 image is divided into 4 patches of 16x16.
3. After this, for each patch we have done feature extraction using Histogram of Oriented Gradient (HOG). In HOG it uses gradient and direction of each pixel in image and then generate descriptor based on values of gradient and direction.
4. After feature extraction of each patch, it is stored in Bag and the bag is called bag of visual words or bag of features. After these to form visual vocabulary, we perform clustering on this bag and select the representatives. We have used k-means clustering and since there are 10 classes in given dataset and each image has 4 patches so the number of representatives is 40.
   This shows result after clustering.

Representative					
					

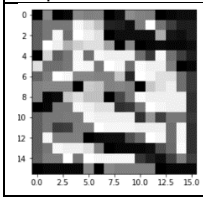
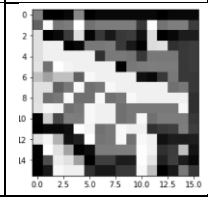
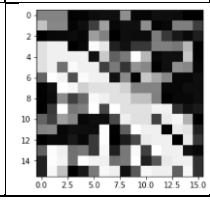
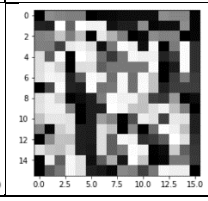
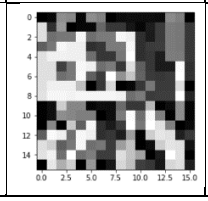
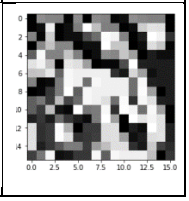
5. Now for image in training set, it will be represented as vector corresponding to visual vocabulary.
6. The test images are also represented as vector corresponding to visual vocabulary.
7. Now the gaussian classifier is used for prediction the labels of test images.
8. Using HOG as feature vector, we get the testing accuracy of 27% on dataset.

Bag of Words using Local Binary Pattern(LBP):

Methodology :

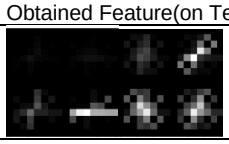
1. The train-test dataset is loaded and the RGB channel is converted to gray channel.
2. For each image in training dataset, the images are divided into patches. For creating patches we have use grid method and none of the patch is overlapping. The 32x32 image is divided into 4 patches of 16x16.
3. After this, for each patch we have done feature extraction using Local Binary Pattern(LBP). In LBP for each pixel value in patch, it is depends on its neighbours. The central pixel value is computed by comparing it's neighbour values and we get binary value and convert it into decimal and set that value of pixel.
4. After feature extraction of each patch, it is stored in Bag and the bag is called bag of visual words or bag of features. After these to form visual vocabulary, we perform clustering on this bag and select the representatives. We have used k-means clustering and since there are 10 classes in given dataset and each image has 4 patches so the number of representatives is 40.

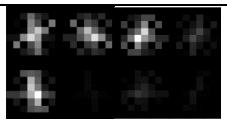
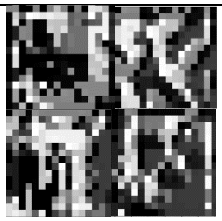
This shows result after clustering.

Representative					
					

1. Now for image in training set, it will be represented as vector corresponding to visual vocabulary.
2. The test images are also represented as vector corresponding to visual vocabulary.
3. Now the gaussian classifier is used for prediction the labels of test images.
4. Using HOG as feature vector, we get the testing accuracy of 27% on dataset.

Comparison of HOG and LBP:

Feature	Accuracy	Obtained Feature(on Test Image)
HOG	27%	

		
LBP	20%	

Observations :

The accuracy using HOG Features is more as compared to LBP. It is because the HOG uses gradient and magnitude of the image. The key points of the patches are taken into consideration because corner and edges of images are detected which are more informative. While in LBP it only provide texture information that is the pixel intensity of the patches. The LBP captures the local pattern while HOG captures the corners and edges.