BONUS ASSIGNMENT

SWAGATAM CHAKRABORTI (MT18146)

PROBLEM STATEMENT: To classify images of 20 different classes using machine learning techniques
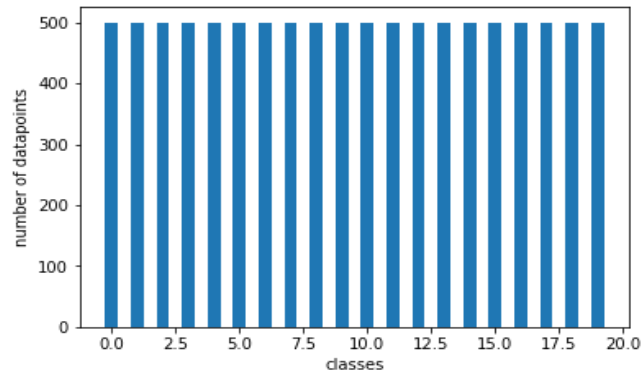
DATASET VISUALIZATION:

Figure 1: Datapoints per class

PREPROCESSING:

    i.   Gaussian blurring: In order to remove some extra noise from the images
    ii.   Image sharpening: In order to sharpen the object edge of the images

FEATURE EXTRACTION:

    i.   Images are converted to three different color models : a. bgr model b. hsv model
    ii.   Hog descriptor: to get the gradient information as feature
    iii.   Local binary pattern: to get the texture information as feature
    iv.   Laplacian filter: to get the regions of high intensity change as feature
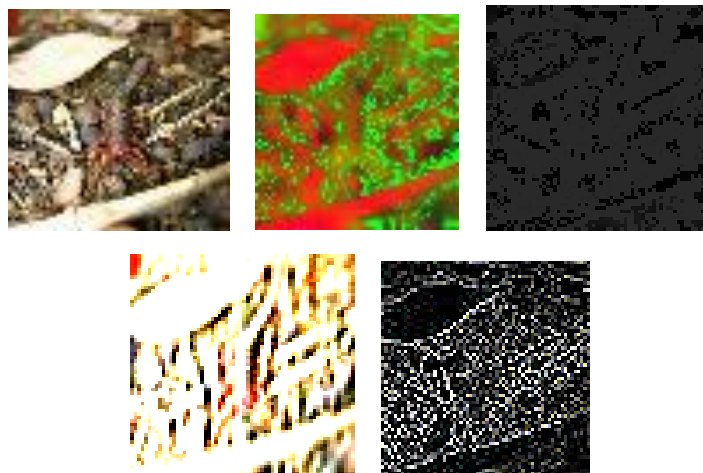   v.   Gabor filter:

Figure 2: Top-Down Left-Right : BGR model , HSV model, LBP , Gabor filter, Laplace filter

OUTLIERS REMOVAL:

Isolation random forest technique have been used for outliers removal with contamination factor 0.1. So total of 100 outliers g=have been removed from the initial dataset.
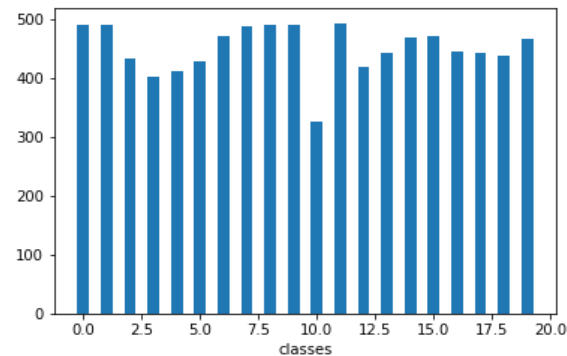

Figure 3: Class distribution after outliers removal

METHODOLOGY:

- Linear SVC have been used as the classifier. Created 5 different models of the linear SVC and trained against the dataset and the prediction of the test dataset have been made on the max vote.
- Bagging is implemented over MLP classifiers and Gaussian Naïve Bayes classifier over 20 iterations and max voting technique have used to predict the final output for each classifier.
- Finally the finally prediction is made based on the max vote of the above SVC, MPL classifier and Gaussian Naïve Bayes.

RESULTS AND ANALYSIS:

FEATURES	CLASSIFIERS	ACCURACY
hog + lbp	Gaussian Naïve Bayes	37
hog + lbp	Linear SVC	42.66
Hog + Laplacian	Linear SVC	45.32
Hog + lbp + gabor	Linear SVC	50.33
Hog + lbp + Laplacian + gabor	Linear SVC + NB + MPL(Bagging)	52.6

Table 1: Accuracy over different techniques