Bag of Color:

1. Install Python 3.x+, Scikit-learn, keras and all other required modules.

2. boc.py is the file that we use for training cluster histogram feature on label y. In the main section of the code put correct folder path to the dataset for both methods; makeVocabulary and makeDataset.

3. This will extract the feature of clusters and labels by itself by calling implemented required methods. You may change output file names of trained model pkl file, generated vocabulary pkl file, and train set of all extracted features in npy file

4. Default: "HISTO-CLASS_model.pkl" is used for trained model and "model_vocabulary.pkl" for generated vocabulary of features.

5. You may also change attribute of batch_size in makeVocabulary method that loads that number of images at one time, other K-Means parameters, Classifier, and nclusters attribute in makeDataset.

6. The usage of test file for this model is simple as you just need to provide directory of data set in data_dir variable in python code along with foldername of testing patches in validation_states variable

7. It will compute the predictions of all files in these folders and will print F1 score and IOU score of the model.