ML Assignment - 2

Section - C

Overview of dataframe:

Labels DataFrame Overview:

	filename	label
0	Image_1.jpg	sitting
1	Image_2.jpg	using_laptop
2	Image_3.jpg	hugging
3	Image_4.jpg	sleeping
4	Image_5.jpg	using_laptop

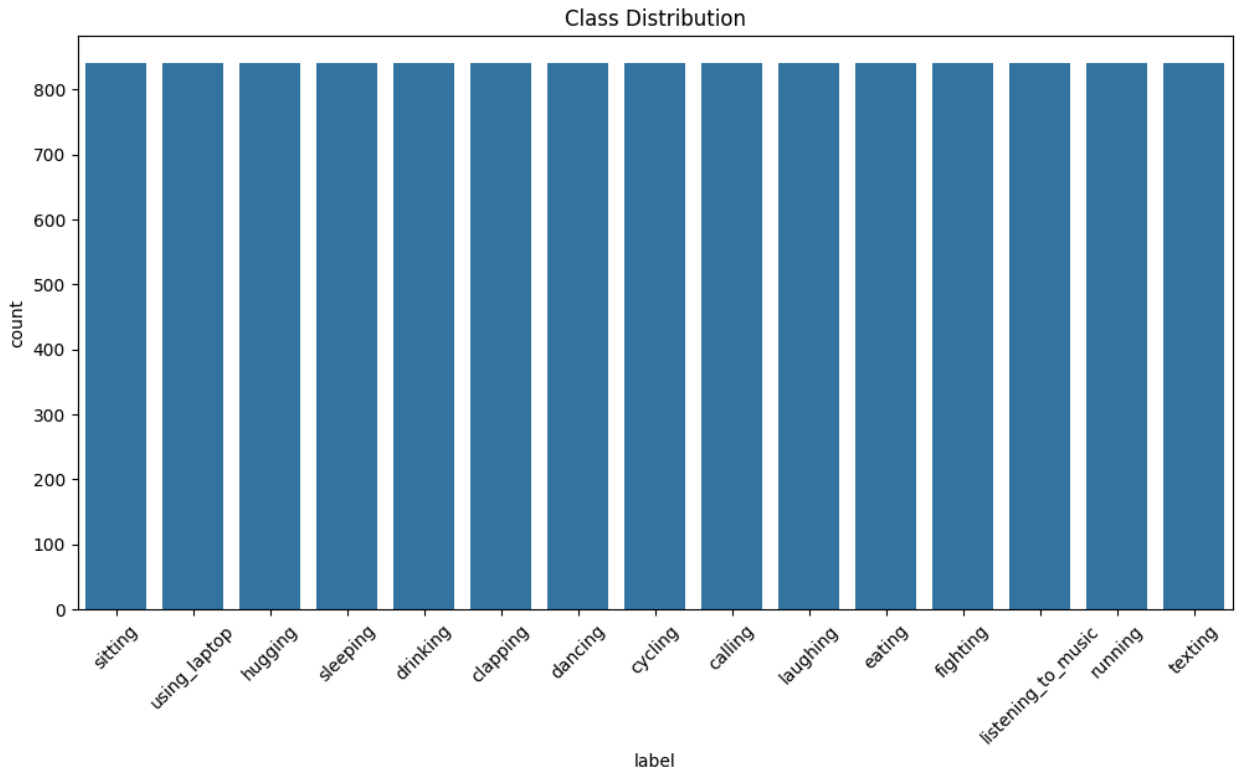
Class Distribution:

label	
sitting	840
using_laptop	840
hugging	840
sleeping	840
drinking	840
clapping	840
dancing	840
cycling	840
calling	840
laughing	840
eating	840
fighting	840
listening_to_music	840
running	840
texting	840

Name: count, dtype: int64
Mean Image Size: 260.38 x 196.57
Std Dev of Image Sizes: 39.92 x 35.28

Total number of images: 12600
Number of unique classes: 15

There are 12600 total number of images and 15 unique classes. All the classes contain same number of images.



All the classes are distributed equally with 840 images in each class.
There is no class imbalance between the classes.

Some sample images from the dataset.

Converted labels to numbers:
Class Dictionary: {'calling': 0, 'clapping': 1, 'cycling': 2, 'dancing': 3, 'drinking': 4, 'eating': 5, 'fighting': 6, 'hugging': 7, 'laughing': 8, 'listening_to_music': 9, 'running': 10, 'sitting': 11, 'sleeping': 12, 'texting': 13, 'using_laptop': 14}

HOG features:

Extracting HOG features: 12600it [07:02, 29.81it/s]

	0	1	2	3	4	5	6 \
0	0.023298	0.014453	0.027768	0.044335	0.100117	0.066850	0.260110
1	0.265062	0.178128	0.072811	0.000000	0.245108	0.096397	0.242874
2	0.314956	0.000000	0.070109	0.051723	0.314956	0.000000	0.000000

3 0.266134 0.109558 0.129011 0.012327 0.073463 0.034740 0.031583
4 0.208969 0.031647 0.056159 0.000000 0.191552 0.000000 0.002721

           7   8   9 ...   8091   8092   8093 \
0 3.331229e-01 0.185694 0.138057 ... 0.091618 0.006241 0.038311
1 1.617676e-01 0.146356 0.265062 ... 0.278305 0.278305 0.043674
2 1.569144e-15 0.000634 0.314956 ... 0.200699 0.006121 0.015988
3 2.896461e-02 0.035759 0.266134 ... 0.252731 0.252731 0.252731
4 7.930886e-02 0.080310 0.067451 ... 0.212578 0.072534 0.078154

      8094   8095   8096   8097   8098   8099   label
0 0.092327 0.151983 0.049297 0.035928 0.000000 0.029561 sleeping
1 0.038651 0.083517 0.026796 0.075051 0.011169 0.162180   cycling
2 0.070171 0.339350 0.153658 0.194351 0.006076 0.000000 sleeping
3 0.252731 0.194674 0.048064 0.111134 0.027942 0.130885 sleeping
4 0.376478 0.376478 0.113994 0.137372 0.268506 0.086373   texting

[5 rows x 8101 columns]

Using all there models:
Using device: cuda
Naive Bayes Accuracy: 0.2615
Decision Tree Accuracy: 0.1421
Random Forest Accuracy: 0.2492
Epoch [10/100], Loss: 2.1350
Epoch [20/100], Loss: 1.9240
Epoch [30/100], Loss: 1.7796
Epoch [40/100], Loss: 1.6700
Epoch [50/100], Loss: 1.5803
Epoch [60/100], Loss: 1.5033
Epoch [70/100], Loss: 1.4360
Epoch [80/100], Loss: 1.3763
Epoch [90/100], Loss: 1.3225
Epoch [100/100], Loss: 1.2742
Perceptron Accuracy: 0.2730

Best Model: Perceptron with Accuracy: 0.273015873015873

Random forest with all features:
RandomForest Test Accuracy: 0.3444

XGB boost with five features [HOG, Color histogram, LBV, Gabor] :
0.39404761904761904

Summary of feature selected:

This approach applies several feature extraction techniques to each image to capture various visual aspects. The Histogram of Oriented Gradients (HOG) extracts gradient-based features, capturing edge directions and shapes from the grayscale image. It focuses on local gradients in 16x16 pixel cells and effectively detect structures.

HSV histograms are computed after converting the image from RGB to the HSV color space. Separate histograms for the Hue, Saturation, and Value channels (each with 32 bins) are generated to capture the color distribution in the image. These histograms are then normalized to provide a compact colour representation.

Local Binary Patterns (LBP) are applied to the grayscale image to capture texture information by comparing pixel intensities with their neighbors. A 10-bin histogram of uniform LBP patterns is calculated and normalised, focusing on local texture variations.

Gabor filters are used to extract texture and edge information from multiple orientations (0°, 45°, 90°, 135°). For each orientation, the mean and standard deviation of the filtered image are computed, providing insights into texture at various angles.

These features (HOG, HSV histograms, LBP, and Gabor) are concatenated into a single feature vector for each image, effectively capturing texture, color, and shape information for image classification tasks.