WEEK 2

Date: 06.10.2023 – 13.10.2023

Project Team: Esma Zeynep Sayılı – Esra Cesur

Project Name: Neckline Classifier

In week 2, we started to examine and prepare the dataset.

Objectives:

1. Uploading the Dataset
2. Examinining the Dataset
3. Preparing the Dataset

Uploading the Dataset

Firsly, we gathered the paths to the directories containing images of each neckline type. We used the "listdir" function from the "os" module to obtain the filenames of the images from the path. Then we created an array to contain these images. In each loop, the corresponding image is read bu using OpenCV and assing and indexed in the array. We repeated the project for each neckline class.

```python
onlyfiles = [ f for f in listdir(path_round) if isfile(join(path_round,f)) ]
round = np.empty(len(onlyfiles), dtype=object)
for n in range(0, len(onlyfiles)):
    round[n] = cv2.imread( join(path_round,onlyfiles[n]) )
```

Examinining the Dataset

Sing matplotlib and cv2, we displayed the images and examined the dataset for each class.
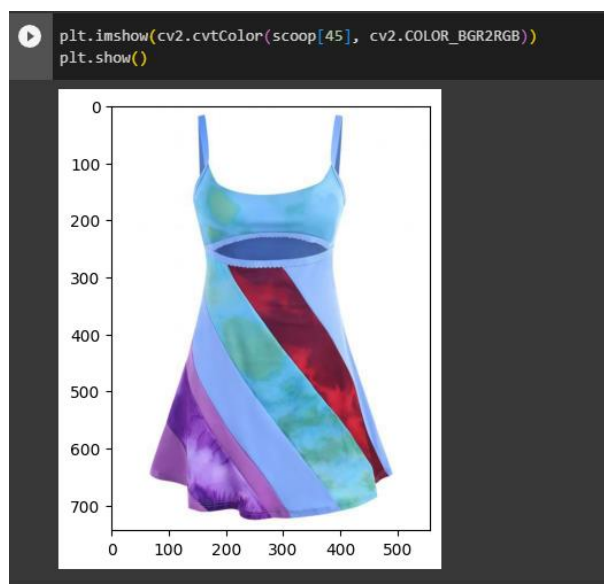
We gathered the shape imformation of the array and various images. There are 966 scoop, 621 round, 1298 v-neck neckline in out dataset.
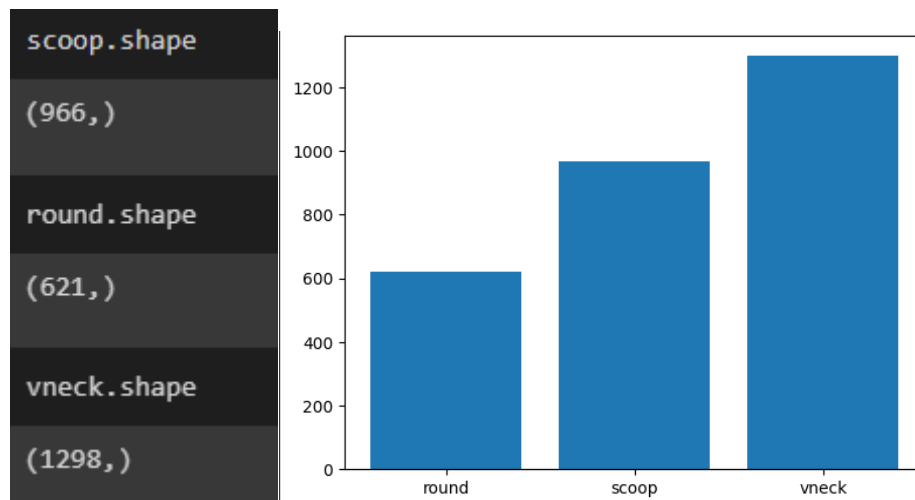

Preparing the Dataset

We applied a cropping process for the image. The reason for this process is to crop images to the necklines of the clothes. We only need the neckline of the images to train our model, so the excess parts should be cropped out. We also added an empty dimension to each image to assign the labels. In each iteration, the corresponding image is cropped according to the coordination information and labeled as its class.

```python
scoop_crop = np.empty((len(scoop), 2), dtype=object)
for i in range(0,len(scoop)):
    y = int(scoop[i].shape[1]/2)
    x = int(scoop[i].shape[0]/8)
    w = int(scoop[i].shape[0])-int(scoop[i].shape[0]/8)
    scoop_crop[i] = (scoop[i][0:y, x:w], 'scoop')
```

We can see that our array have two dimention now, 2nd one including the labels of the images and the images are cropped to their necklies.
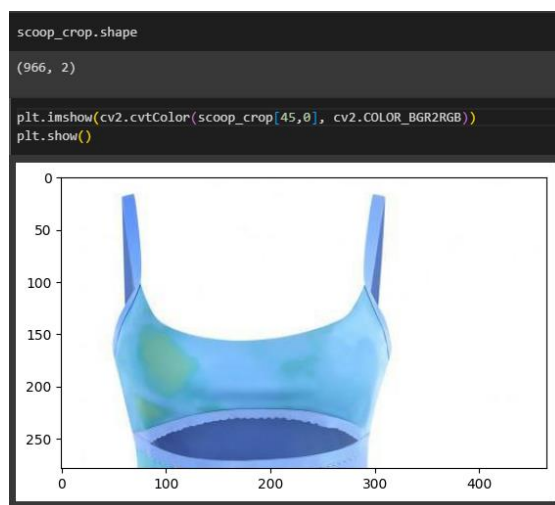
We combined each dataset of three classes together to created our main dataset. We applied shuffling process to create randomness in the data and avoid iteration of the same classes. Lastly we created a array for the labels to use in the training and testing processes.