CSE 535 Project: SmartHome Gesture Application Control – Part 2

Approach to the given problem

The first step is to extract middle frame of the 51 video recorded from the part 1. The code is provided and the only thing added is to extract the gesture name from the video name and assign the right label to all of them. Store both video name and labels to a dataframe.

The second step is extracting the feature to form the penultimate layer. The code is provided. One line of code to transfer the three channel RGB to one channel GRAY needs to be added. Then the numpy array of features are ready.

Do the same frame fame extraction to the test video as well. And Perform the feature extraction as well.

Compare each of the test video feature to the traindata features (penultimate layer) and find the index of the minimum. Use the index and the traindata feature dataframe to find the label accordingly.

Solution for the problem

It is very difficult to get a good prediction accuracy to this question. The issues as follow:

1. The tempo of the gesture action in the video may different between train and test video, so the middle frame cannot guarantee the best practice.
2. The model is pretrained model. To meet different training set, a modified model will be required.
3. It's better to have distinct hand gesture and background color while recording the video

So the solutions will be, try to find a white background for the train video, try to follow the same tempo as the expert gesture video, trim the model to meet with the requirements of the traindata.