In the second phase of the SmartHome Gesture Control Application, we focus on accurately identifying specific gestures from a collection of provided gesture videos. The process involves the utilization of the training set created in Part 1 to generate a comprehensive feature vector set for all videos. Here's a breakdown of the steps involved:

1.   Training Set Feature Extraction: a. Extract the middle frame for each video. b. Utilize a pre-trained CNN model to extract feature vectors for the middle frame images. c. Consolidate all the extracted vectors, associating each with its corresponding gesture.

2.   Testing Data Gesture Determination: a. Extract the middle frame for each test video. b. Employ the same CNN model to extract feature vectors for the middle frame images of the test videos. c. Determine the gesture for each test video by computing the cosine similarity against the testing vector set. d. Identify the vector with the minimum cosine difference and record its output label in the results.csv file.

Additionally, a robust mechanism for handling discrepancies has been implemented:

•   Testing Data Mutation:
•   In cases where a gesture video does not match the expected label, the vector data is tagged with the correct label and added to a larger vector set.
•   The Cosine similarity is rerun on this augmented vector set to enhance accuracy and accommodate potential misclassifications.