SmartHome Gesture Control Application Project Part 2 Report

Introduction

This report evaluates a gesture recognition application service designed to identify specific hand gestures from video files. The application leverages computer vision and machine learning techniques to analyze video data, extract relevant features, and classify these features into predefined gesture categories.

Problem Statement

The primary challenge was to develop a system capable of accurately recognizing hand gestures from a series of video files. This involves several key tasks:

- Extracting relevant frames from video data.
- Preprocessing these frames to be suitable for feature extraction.
- Extracting meaningful features from the processed frames.
- Classifying these features into gesture categories based on a pre-trained model.

Approach

My approach was divided into below steps:

- The first step involved organizing the video data into training and testing datasets. The training data was used to train the model on various gestures, while the testing data was used to evaluate the model's performance.
- I utilized the frameextractor module to select specific frames from each video file. This assumed that key gesture features are present throughout the video and can be adequately represented by a single frame.
- The handshape_feature_extractor module was implemented to extract meaningful features from the selected frames. This involved converting images into grayscale, normalizing pixel values, and applying image processing techniques to highlight gesture characteristics.
- I used TensorFlow to compare the extracted features against a pre-trained model using cosine similarity. This allowed us to classify the extracted features into predefined gesture categories.

Solution

The solution was implemented in Python, utilizing libraries such as OpenCV for image processing and TensorFlow for machine learning operations. The application's workflow is as follows:

- Preprocess the Data: Video files are processed to extract relevant frames.
- Feature Extraction: Selected frames undergo feature extraction to prepare for classification.
- Gesture Classification: Extracted features are compared to a pre-trained model to recognize gestures.

- Output Generation: The application outputs the recognition results, mapping video files to recognized gestures.