Project Report: Hand Gesture Recognition System

1. Introduction

The goal of this project is to develop a hand gesture recognition system that can detect American Sign Language (ASL) gestures using a webcam. The system collects hand gesture data, preprocesses it, and trains a machine learning model for real-time prediction. The steps include data collection, preprocessing, model training, and testing.

2. Tools and Technologies

- **Python**: Used for coding and implementing the system.
- **OpenCV**: For capturing video frames and handling image processing tasks.
- **MediaPipe**: Used for hand landmark detection, a key part of recognizing hand gestures.
- **scikit-learn**: Used for training the RandomForestClassifier model for gesture prediction.

3. Methodology

- **Data Collection (collect_images.py)**:
  This script captures images of hand gestures for training. It uses the webcam (index 0) to collect frames and stores them in a directory for each class of gesture (A to Z). A total of 100 images per class are collected. Each gesture class is saved as a separate folder within the 'data' directory.
- **Dataset Creation (create_dataset.py)**:
  The collected images are processed to extract hand landmarks using MediaPipe. These landmarks (normalized x and y coordinates) are used as features for training the model. The extracted features are stored in a pickle file (data.pickle), which includes both the feature data and labels (gesture classes).
- **Model Training (train_classifier.py)**:
  The train_classifier.py script loads the dataset from data.pickle, splits it into training and testing sets, and trains a RandomForestClassifier. The classifier predicts the gesture based on the extracted features. The model is saved in a pickle file (model.p).
- **Gesture Recognition (test_classifier.py)**:
  This script performs real-time gesture recognition. It uses the webcam to capture live video, extracts the hand landmarks using MediaPipe, and makes predictions using the trained model. The recognized gesture is displayed on the screen, along with a bounding box around the hand.

4. Results

- **Model Accuracy**:
  The RandomForest model achieved a classification accuracy of 100% on the test set, which demonstrates its ability to recognize hand gestures accurately.

```
PS E:\ankith\PROJECTS\SignLanguageDetector> python train_classifier.py
100.0% of samples were classified correctly !
```

- **Real-time Gesture Recognition**:
  The system works in real-time, with the webcam feed being processed for hand landmarks and predictions displayed instantly. The model accurately predicts the gesture corresponding to the hand position captured in each frame.

5. Challenges Faced

- **Camera Calibration**:
  Initially, there were issues with capturing frames from the webcam due to incorrect camera index values. This was resolved by changing the camera index.
- **Multi-Hand Recognition**:
  Initially identifying multiple hands within the same frame is complex, particularly when the hands overlap or are positioned closely together. Incorrect detection or confusion between hands can occur, especially in crowded scenes.

6. Future Work

- **Extend Dataset**:
  Currently, the dataset only includes 26 hand gestures (A to Z). Expanding the dataset to include more gestures can improve the model's capability to interpret sentences
- **Improved Model**:
  Exploring more complex models (e.g., Convolutional Neural Networks) might improve the accuracy, especially for noisy data or non-standard gestures.
- **Multi-Hand Recognition**:
  Future improvements could include recognizing gestures made by multiple hands simultaneously.

7. Conclusion

This project successfully demonstrates a hand gesture recognition system capable of identifying ASL letters using a webcam. By leveraging machine learning techniques and hand landmark detection, the system can process live video and make accurate predictions. The system can be extended to include more gestures and further optimized for better performance.