Simple Sign Language Detection Using Computer Vision

B200767CS Rhuthuvarna SP

November 11, 2023

1 Introduction

In a world that is increasingly interconnected and reliant on digital communication, the importance of inclusive and accessible communication methods cannot be overstated. One significant group facing unique communication challenges is individuals with hearing impairments. For them, traditional spoken language is often inadequate, making sign language a critical means of expression. To bridge this communication gap and empower individuals with hearing impairments, this project introduces a comprehensive Sign Language Detection System using Computer Vision. The primary objective is to develop a real-time solution capable of recognizing and interpreting sign language gestures, thereby enabling seamless communication for this community. Beyond technical innovation, the project embodies a broader commitment to inclusivity and accessibility, advocating for the rights and empowerment of individuals with hearing impairments.

2 Objective

The central objectives of this project are to develop a real-time sign language detection system capable of interpreting sign language gestures, thus enhancing communication for individuals with hearing impairments. By promoting inclusivity and accessibility through technology, the project not only seeks to break down communication barriers but also to empower individuals to engage in everyday conversations, access information, and fully participate in society. Beyond the technical aspects, it aspires to raise awareness about the challenges faced by this community and advocate for their rights, ultimately contributing to a more inclusive and empowered society where everyone's voice can be heard and understood.

3 Dataset

The heart of this project is the creation of a custom dataset for sign language gestures. The dataset includes hand images for four American sign language letters: A, B, C, and D. The data collection process involves using datacollection.py, which captures images of these gestures from the webcam. These images are then saved for training and evaluation. The dataset plays a crucial role in training the gesture recognition model.

Figure 1: Sample images from the dataset: ASL sign for A and B

4 Methodolology

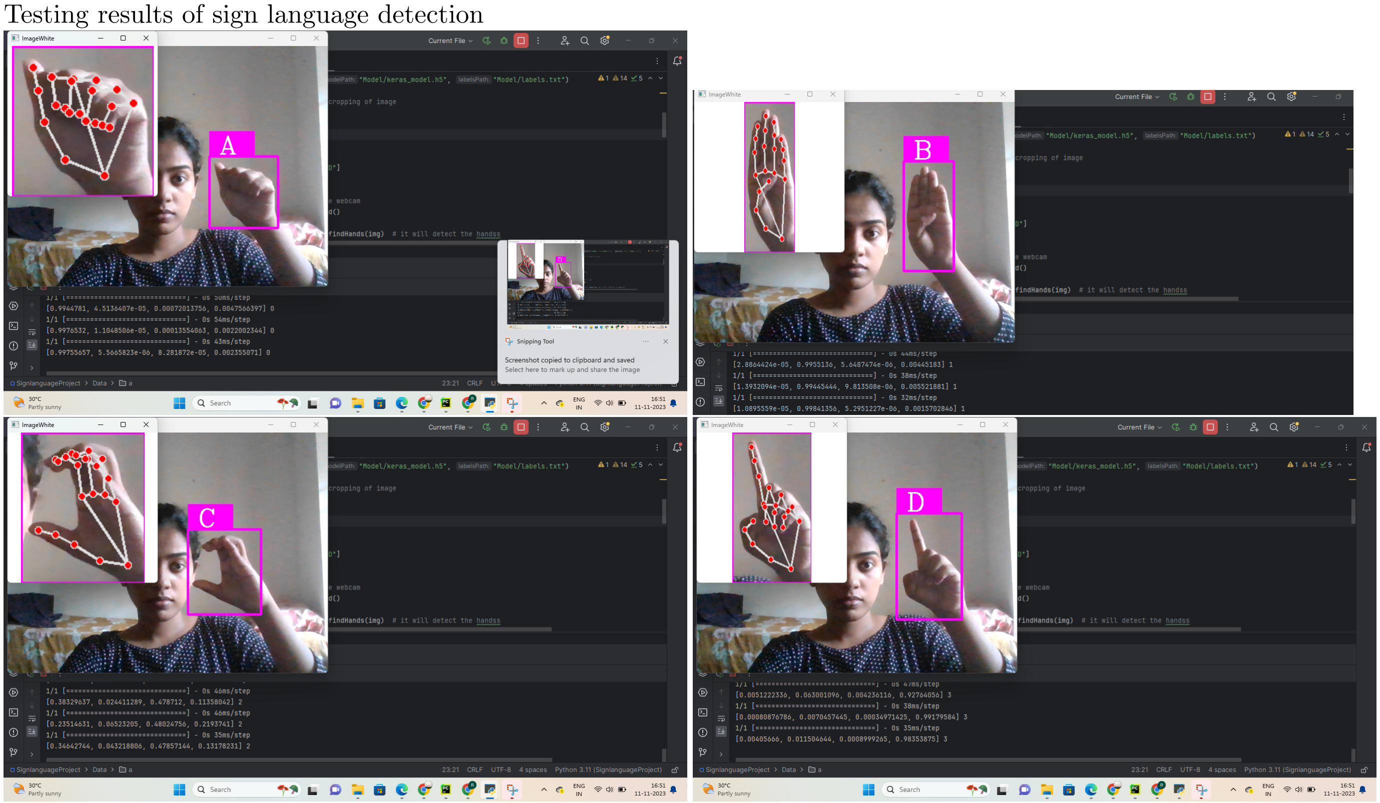
1. Webcam Initialization: We start by initializing the webcam using OpenCV's cv2.VideoCapture(0). This provides us with a video stream to capture frames.
2. Hand Detection: We use the HandDetector from the cvzone library for hand detection. The max-Hands parameter is set to 1, as we aim to detect a single hand. The system's ability to locate hands in real-time is crucial.
3. Image Cropping: When a hand is detected, we crop the image to focus on the hand region. We apply an offset to ensure that the entire hand is included in the cropped image.
4. Standardized Image Size: To maintain consistent image sizes, we create a white background of fixed dimensions (e.g., 300x300 pixels) and place the cropped hand image in the center. This standardization is essential for classification.
5. Sign Classification: Google's Teachable Machine is used to create labeled datasets for training the deep learning model. The Classifier object loads the trained model from "keras model.h5" and labels from "labels.txt." This model can identify ASL signs based on the cropped hand image.

5 Result

The developed system demonstrated real-time sign language detection capabilities, accurately recognizing and interpreting sign language gestures as they are performed.

5.1 output

Testing results of sign language detection



6 Conclusion

In conclusion, this project demonstrates the practical application of computer vision and deep learning in the context of ASL sign language interpretation. By combining hand tracking, image preprocessing, and classification, we've developed a system that makes ASL more accessible in the digital world. The use of Google's Teachable Machine showcases the potential of AI in creating inclusive solutions for the hearing-impaired community. This technology can greatly improve communication and accessibility.

6.1 Future work

- Extend the project to include real-time sign language translation to spoken language or text.
- Continue refining the deep learning model used for sign classification. Gather a more extensive and diverse dataset to improve the model's accuracy and its ability to recognize a broader range of ASL signs.
- Modify the system to support the detection and classification of multiple hands simultaneously.