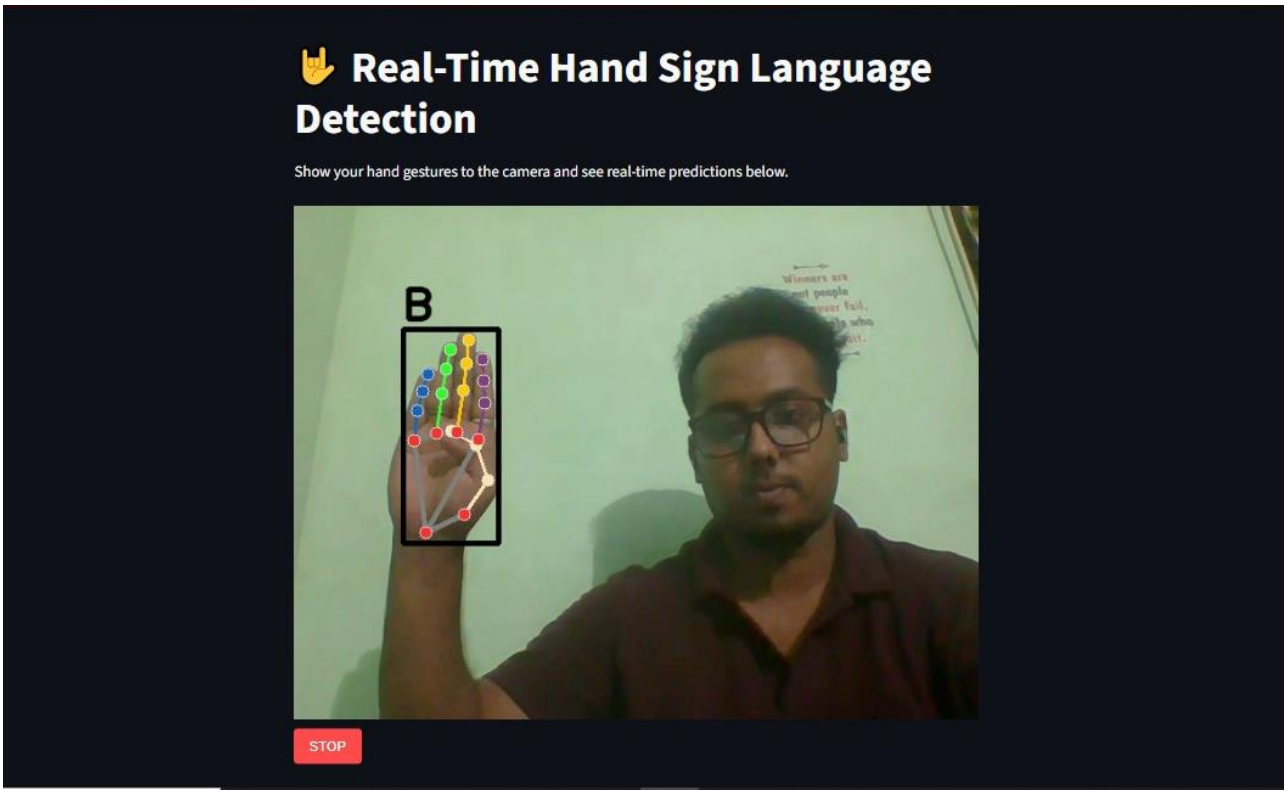
Hand Sign Language Detection using MediaPipe, OpenCV & Random Forest

"Bridging Communication Gaps with Machine Intelligence"

Author: Tanmay Dhar
Date: June 24, 2025



GitHub: [Click Here]

Project Drive: [Click Here]

1. Objective

The objective of this project is to develop a real-time system that can recognize American Sign Language (ASL) hand gestures using webcam video feed. The system leverages MediaPipe for hand landmark extraction and a Random Forest classifier for gesture recognition. The goal is to aid communication for the hearing and speech impaired by providing a non-verbal input interface.

2. Tools & Technologies Used

- Python 3.10
- MediaPipe (Hand Landmark Detection)
- OpenCV (Real-Time Image Processing)
- scikit-learn (Machine Learning)
- Streamlit (Web-based Interface)
- NumPy, Matplotlib, Seaborn

3. Dataset

A custom dataset was created by capturing hand gesture images using a webcam. It contains 100 images for each of the 37 classes (A-Z, 0–9, and Empty). MediaPipe was used to extract 21 hand landmarks per image, resulting in 42 feature values.

4. Model

A Random Forest Classifier was trained on the extracted features. The model was evaluated on a test split and achieved 100% accuracy. Feature vectors were generated using normalized landmark positions.

5. Project Phases

Phase 1: Data Preparation:
- Captured images using webcam (100 per class)
- Extracted 21 hand landmarks using MediaPipe
- Normalized and serialized data

Phase 2: Model Building & Training:
-Trained a Random Forest Classifier on the processed features

Phase 3: Evaluation:
- Accuracy: 100%
- Visualized Confusion Matrix
- Classification Report with precision, recall, and f1-score

Phase 4: Real-Time Detection with OpenCV:
-Live webcam feed processed in real-time
- MediaPipe landmarks extracted and passed to trained model
- Predictions shown on screen

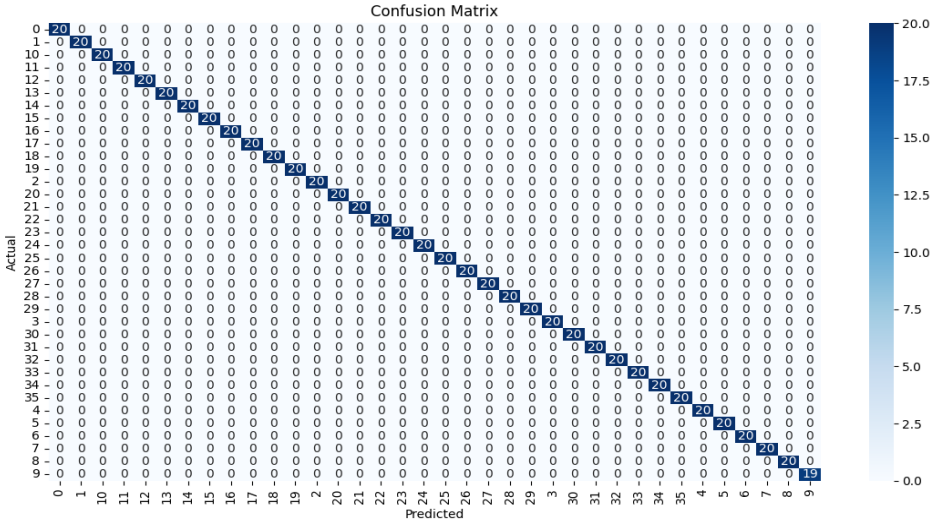
Phase 5: Deployment:
-Web-based deployment using Streamlit with webcam support

6. Results

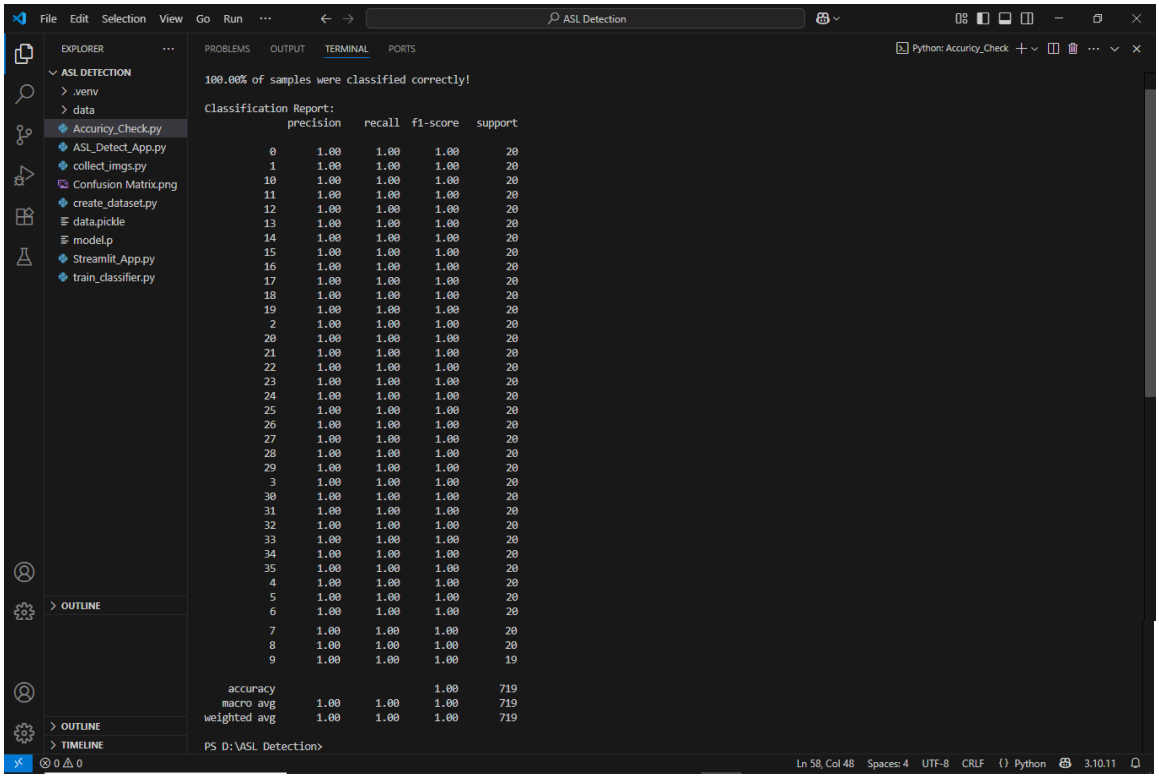
The system demonstrated perfect classification on the test data with 100% accuracy. Real-time prediction was tested through both a desktop application and a Streamlit-based web interface.

7. Screenshots

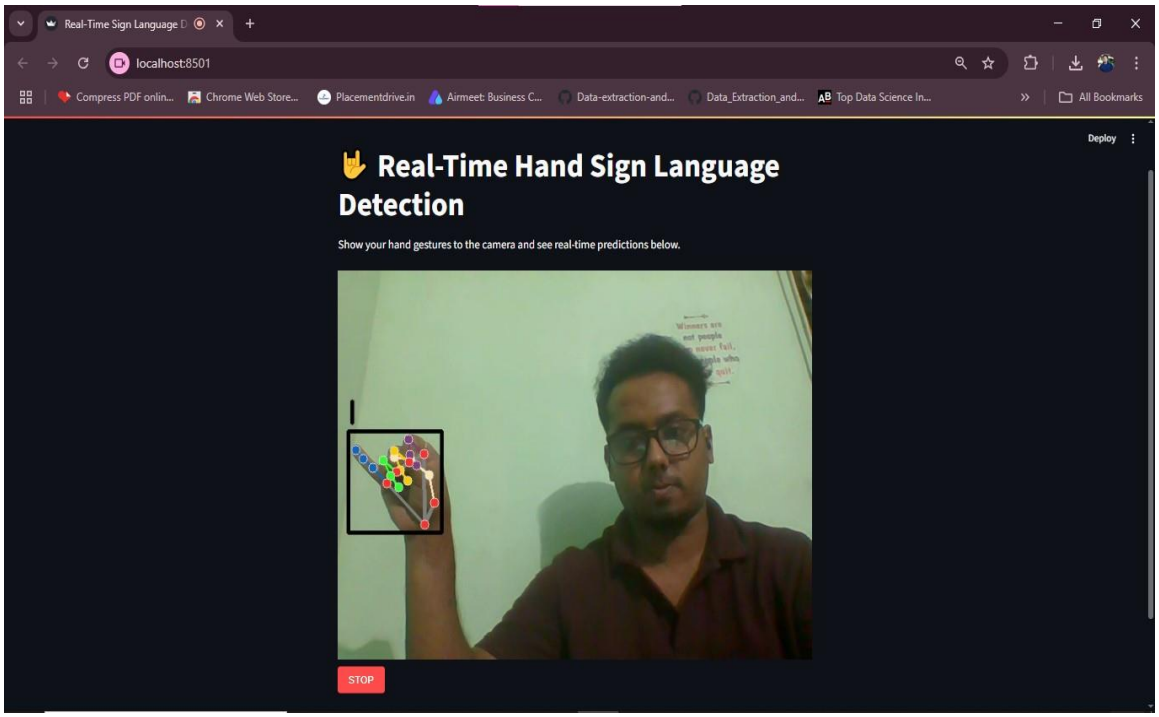
Confusion Matrix:

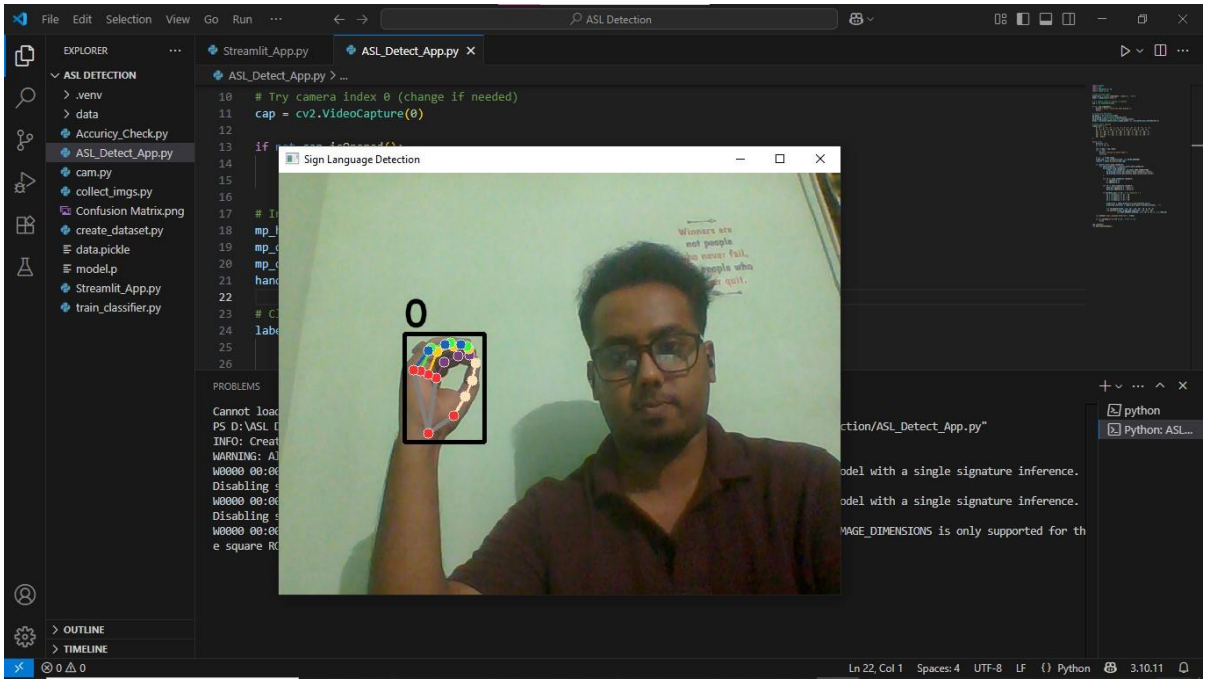


Classification Report Output:



Real-time Detection Demo:





8. Conclusion

The hand sign language detection system achieved perfect recognition on test data and works efficiently in real-time. The use of MediaPipe and a traditional machine learning model made the system lightweight and highly accurate. In the future, this system can be expanded to interpret sequences of gestures, enabling full sentence recognition.

9. Future Scope

- Expand the system to support sentence-level recognition using gesture sequences.
- Integrate Natural Language Processing (NLP) to convert signs into grammatically correct sentences.
- Train on more diverse datasets with different backgrounds, lighting, and hand types for better generalization.
- Deploy on mobile platforms using TensorFlow Lite or MediaPipe on-device ML for accessibility.
- Add support for different sign languages (e.g., ISL, BSL) for multilingual sign recognition.
- Implement voice output to make the system a complete communication aid for the hearing impaired.
- Enable gesture learning feature to allow custom sign training by the user.