



SIGN LANGUAGE DETECTION USING ACTION RECOGNITION WITH VOICE OUTPUT

DEBASMITA DHAR
PARTH PATEL
DHRUV KUMAR
SAMANGYA NAYAK
SHASHWAT MISHRA
ARPITA DATTA

2105117
21052512
21052513
21052526
21053356
21053382



Contents

- INTRODUCTION
- System Overview
- MediaPipe for Hand Tracking
- System Execution Flow
- Model Evaluation
- Algorithm Workflow Diagram
- INNOVATIVE APPLICATIONS OF SIGN LANGUAGE DETECTION
- Audio Output Integration
- Key Challenges
- Limitation
- Future Work



INTRODUCTION



Sign language is a crucial mode of communication for individuals with hearing or speech impairments. However, it is not universally understood, creating a gap in effective communication. This project presents a real-time system that detects sign language gestures using computer vision and machine learning. It translates these gestures into text and voice output, making communication more accessible and inclusive.

System Overview

This project aims to recognize sign language gestures using real-time video input and convert them into both text and voice output. By combining computer vision, machine learning, and web technologies, the system enables smoother communication for individuals with hearing or speech impairments. The model detects hand gestures through a webcam, processes them using a trained classifier, and provides immediate visual and audio feedback. A user-friendly website was developed to interact with the system seamlessly.

MediaPipe for Hand Tracking



MediaPipe is an open-source framework developed by Google for building multimodal machine-learning pipelines. In this project, real-time hand tracking was used to detect 21 hand landmarks with high accuracy. These landmarks were crucial for extracting gesture features, which were then passed to the machine learning model for sign recognition. MediaPipe's lightweight and fast performance made it ideal for seamless webcam-based gesture detection.





System Execution FLOW



Data Acquisition:

Collected diverse gesture videos from various users, lighting, and angles.



Evaluation

- Assessed model using accuracy, precision, recall, and F1-score.



Preprocessing:

Applied frame segmentation, noise reduction, and resizing for consistency.



Feature Extraction

Used optical flow and CNNs to extract motion and spatial features.



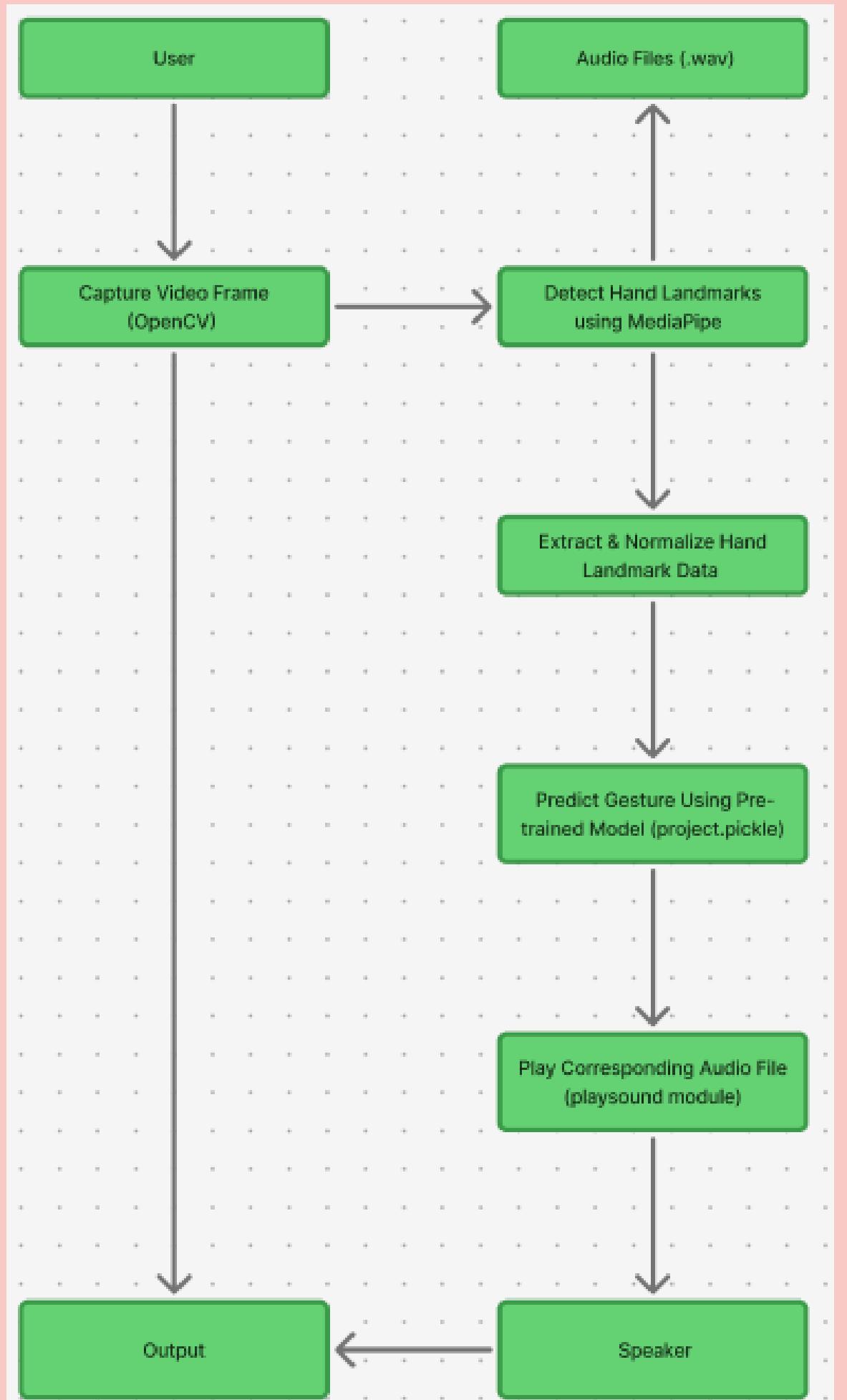
Real-time Deployment

- Integrated the model with a live video feed and audio output via Flask.

Model Evaluation

- Recall: Measures how well the model detects actual gestures correctly
- F1-Score: Balances precision and recall to evaluate overall accuracy
- Latency: Assessed the time delay between gesture input and output response





Algorithm Workflow Diagram



Website Overview

“HOME” INTERFACE

- Buttons to view the project report and presentation.
- Includes an “About the Project” section explaining its purpose and goals.
- Simple and clear navigation bar with links to Home, Model, and Team pages

The screenshot shows the homepage of a website titled "Sign Language Detection" with the subtitle "Breaking barriers with technology 🙌". The header includes a logo for "SIGN LANGUAGE DETECTION" and navigation links for "Home", "Try the Model", and "Meet the Team". The main content area features a section titled "About the Project" which describes the use of machine learning for real-time sign language recognition. Below this is a "Did You Know?" box containing a quote: "Sign language: Where hands do the talking, and the world listens!" followed by two hand-emoji. At the bottom, there is a "Project Resources" section with buttons for "View Project PPT" and "View Project Report".

Sign Language Detection
Breaking barriers with technology 🙌

Home Try the Model Meet the Team

About the Project

This project uses machine learning to recognize and interpret sign language gestures in real-time. Our goal is to make communication more inclusive and accessible for everyone.

Did You Know?

"Sign language: Where hands do the talking, and the world listens!" 🙌

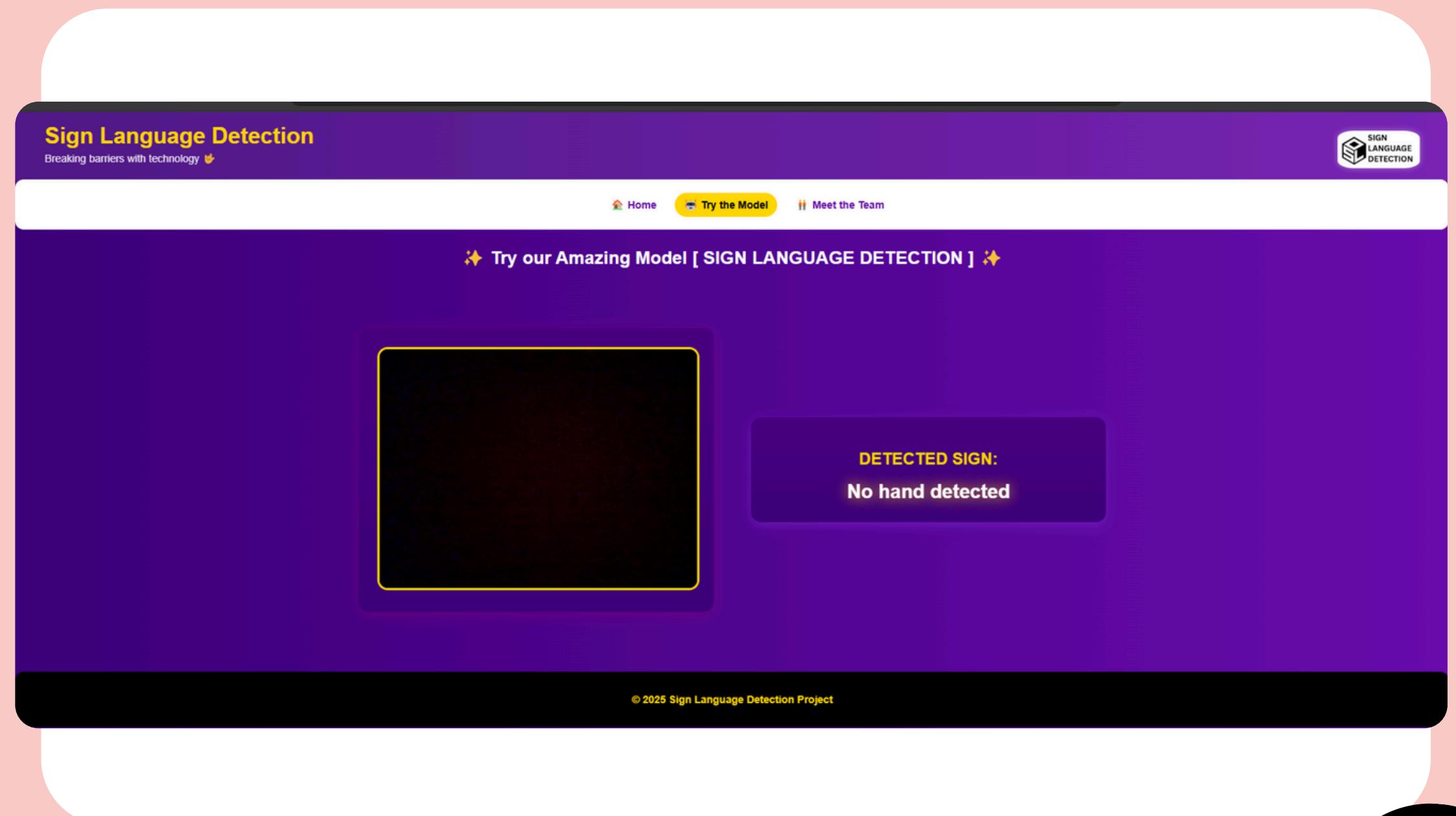
Project Resources

[View Project PPT](#) [View Project Report](#)

Website Overview

“Try the Model” INTERFACE

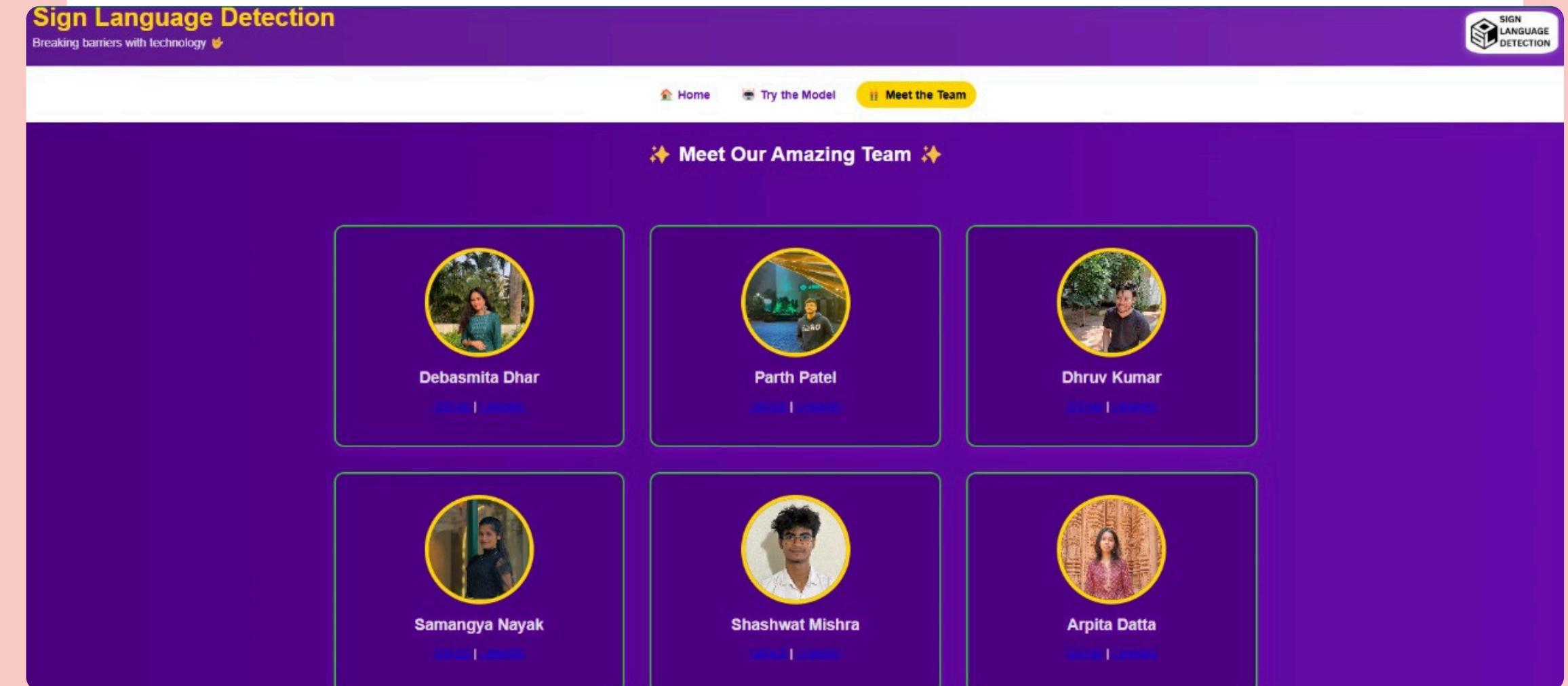
- Accessible via the “🤖 Try the Model” tab in the navigation bar
- Displays a live webcam feed for real-time gesture detection
- Sends video frames to the backend for prediction
- Displays the detected sign on the screen
- Includes voice output for the recognized gesture



Website Overview

“Meet The Team” INTERFACE

- Clicking a member displays their contribution.
- Includes direct links to each member's GitHub and LinkedIn profiles.
- Responsive layout for both desktop and mobile views
- Helps users connect with the people behind the project



INNOVATIVE APPLICATIONS OF SIGN LANGUAGE DETECTION



EDUCATIONAL TOOLS

Sign language detection can enhance teaching methods for learners of all ages.

COMMUNICATION AIDS

These technologies assist the deaf and hard of hearing in everyday interactions.

CUSTOMER SERVICE ENHANCEMENTS

Sign language detection can improve customer interactions, making services more accessible.



Audio Output Integration

- Integrated using Python's playsound, text-to-speech libraries
- Recognized gestures are converted into corresponding spoken words
- Enhances accessibility for users with hearing or speech impairments
- Voice playback is managed with cooldown logic to avoid repetition
- Makes the system more interactive and inclusive

Key Challenges

- Integrated using Python's playsound or text-to-speech libraries
- Recognized gestures are converted into corresponding spoken words
- Enhances accessibility for users with hearing or speech impairments
- Ensures immediate auditory feedback after gesture detection
- Voice playback is managed with cooldown logic to avoid repetition
- Makes the system more interactive and inclusive

Limitation

- Currently supports only a limited set of gestures
- Accuracy may decrease under poor lighting or cluttered backgrounds
- Struggles with overlapping or very fast hand movements
- Works best with a single hand in the frame; multi-hand gestures can be misinterpreted
- Audio output may not sound natural for all gestures due to basic text-to-speech
- Not yet optimized for all mobile devices or screen sizes
- Limited language support in voice output (currently single-language)



FUTURE WORK

- Add full words, phrases & regional sign language support
- Enable multilingual voice output
- Improve speed & reduce system latency
- Build a mobile app version
- Deploy the system on the cloud
- Add gesture-to-text chat & UI customization
- Integrate an AI chatbot for real-time responses
- Use in public spaces like metros, airports & govt offices



THANK you

