Sign Language Recognition System Using Machine Learning

Pratiksha Ranavare[1], Sanika Mhatugade[2], Rajeshwari Kumbhar[3], Mayuri Patil[4]

[1,2,3,4]*Computer Engineering*

[1,2,3,4]*K.P. Patil. Institute of Technology*

Abstract—**Sign language is the primary mode of communication for people with hearing and speech impairments. This paper presents a real-time Sign Language Recognition System that converts hand gestures into text and speech using computer vision and machine learning techniques. The proposed system uses a webcam to capture hand gestures, Media Pipe for detecting hand landmarks, and machine learning algorithms such as Random Forest and Support Vector Machine for gesture classification. Recognized gestures are converted into text and further transformed into speech using a text-to-speech module. The system is low cost, easy to use, and works in real time without requiring special sensors or gloves, making it suitable for practical communication assistance.**

Index Terms—*Sign Language Recognition; Machine Learning; Media Pipe; Computer Vision; Hand Gesture Recognition*

I. INTRODUCTION

Sign language is the main way for deaf or hard-of-hearing people to communicate. Unfortunately, most people don't know it, which makes communication difficult. Thanks to advances in AI and computer vision, we can now create systems that can understand sign language and turn it into text or speech in real time.

This project aims to build such a system using:

- Python
- OpenCV
- MediaPipe
- Machine Learning models (Random Forest / SVM)
- Text-to-Speech (pyttsx3 library)

The idea is to make a simple, affordable system that can recognize hand gestures without special gloves or sensors.

II. LITERATURE REVIEW

Earlier sign recognition systems used glove-based sensors, which were expensive and uncomfortable. Later, researchers started using image processing and deep learning instead.

Some key improvements include:

- CNN models for gesture recognition
- MediaPipe for fast hand tracking
- RNN/LSTM for understanding moving signs
- Support for different sign languages (ASL, ISL, BSL)

Recent research focuses on lightweight models that run even on normal CPUs, making them more practical.

Our project also follows this approach using MediaPipe with a simple ML model for fast and accurate results.

III. METHODOLOGY

The system works through several steps:

1. Data Collection
   Capture hand gesture images from a webcam
   Extract 21 hand landmark points
   Save them as CSV files for training

2. Preprocessing
   Normalize coordinates
   Remove unwanted noise
   Prepare data for model training

3. Feature Extraction
Each hand has 21 landmarks (x, y, z) → 63 features
These features are used to train the ML model
4. Model Training
Use Random Forest or SVM
Train and test the model with labeled data

Save the trained model
5. Real-Time Recognition
Webcam detects hand
MediaPipe extracts landmarks
ML model predicts the sign
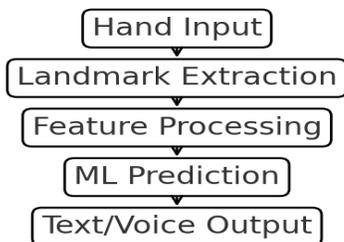Display text and convert it to speech

IV. SYSTEM ARCHITECTURE
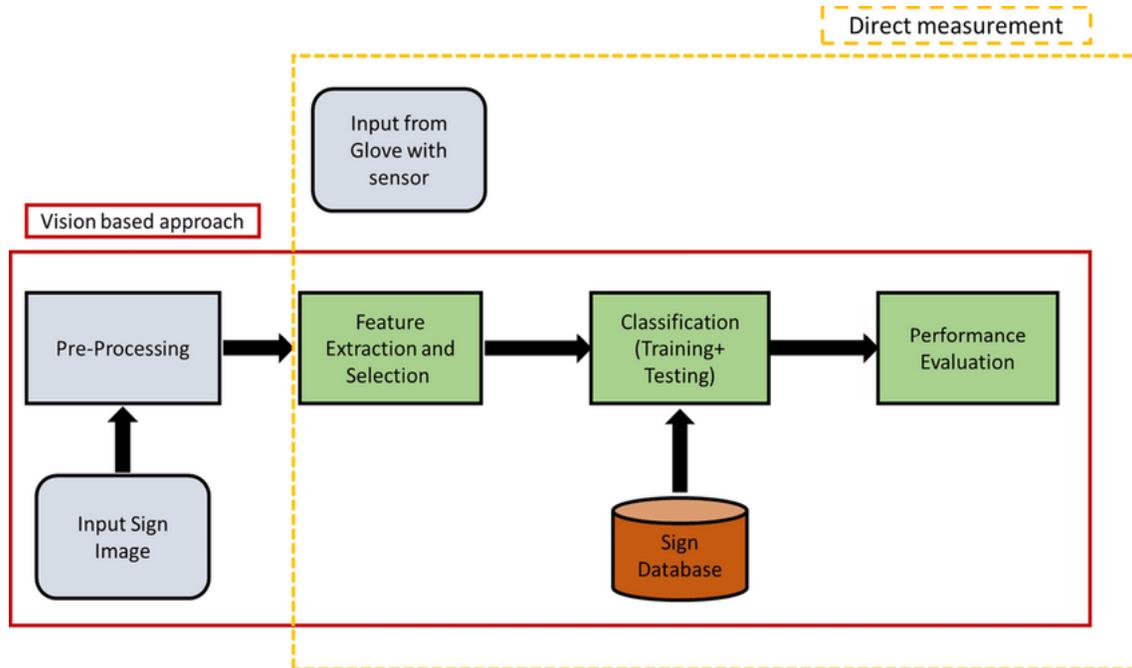
COMPONENTS:
1. Camera
2. Hand Detection Module
3. Landmark Extraction
4. ML Classifier
5. Text Output
6. Voice Output

V. WORKFLOW DIAGRAM

STEPS:
1. Start
2. Capture real-time video
3. Detect hand region
4. Extract 21 MediaPipe landmarks
5. ML model prediction
6. Display text
7. Convert to speech
8. End

VI. RESULT

1. Works in real time at 15–25 FPS
2. High accuracy for alphabet signs (A–Z)
3. Instant voice output
4. Successfully recognizes words like HELLO, YES, NO, THANK YOU
5. Works smoothly on normal laptops without GPU
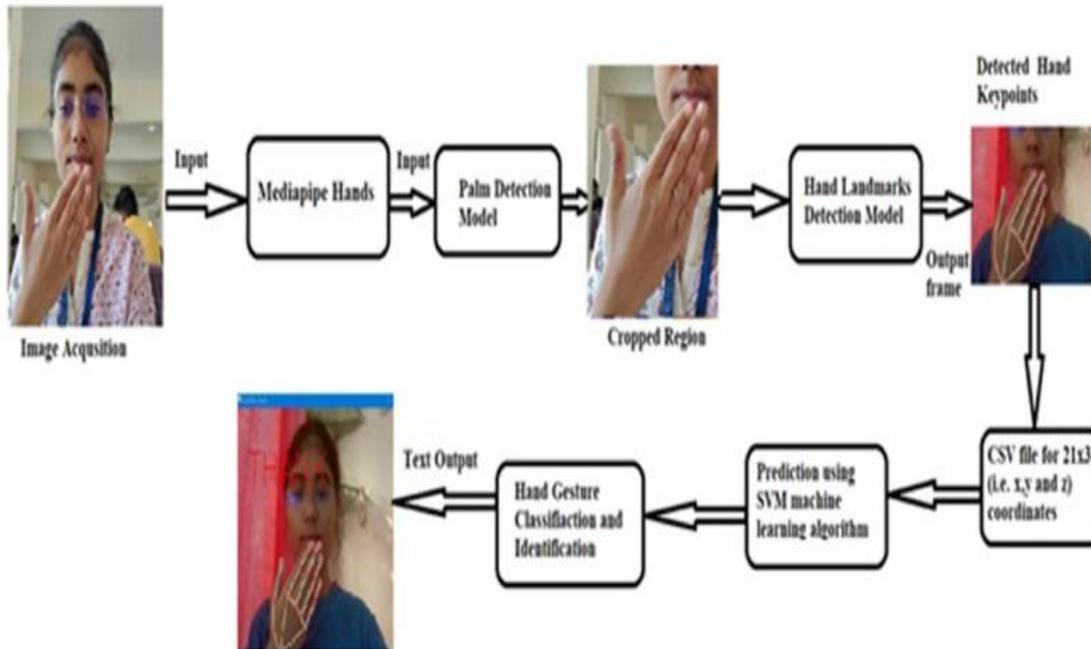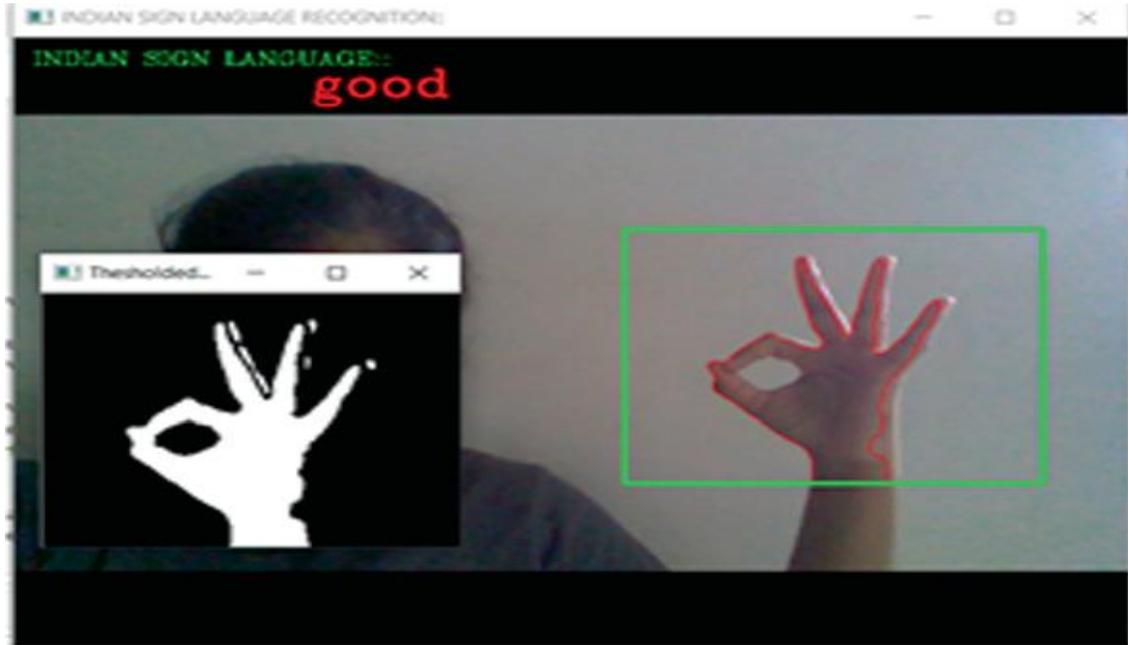
VII. LIMITATIONS

- Moving (dynamic) signs are not yet supported
- Lighting affects accuracy
- Cluttered backgrounds reduce performance
- Works best with one hand

VIII. FUTURE SCOPE

- Add support for dynamic gestures
- Use advanced models like CNNs or Transformers
- Create a mobile app version
- Add better voice assistant features
- Support multiple languages
- Recognize two-hand ISL sentences

IX. CONCLUSION

This project proves that a real-time sign language recognition system can be built using Python, Media Pipe, and ML models. It accurately detects hand gestures and turns them into both text and voice. The system is affordable and helps improve communication between hearing and non-hearing individuals. With more improvements, it can become a complete sign-to-speech translator.

REFERENCES

[1] Python Official Documentation
https://www.python.org/doc/
[2] Kaggle Sign Language MNIST Dataset
https://www.kaggle.com/datamunge/sign-language-mnist
[3] YouTube Tutorials on OpenCV and Media pipe
[4] GitHub Repositories on Real-time Gesture Recognition
[5] Research paper.
https://doi.org/10.48550/arXiv.2204.03328
https://www.atlantis-press.com/proceedings/icamida-22/125986284
https://www.mdpi.com/1424- 8220/23/18/7970