ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 8, Issue 3, March 2025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Hand Gesture Recognition System for Deaf and Mute People

A. Mane[1], H. Mundhe[2], S.Supekar[3], T.Thorat[4], S.Wadekar[5]

Department of Computer Engineering, AISSMS Polytechnic, Shivajinagar, Pune, Maharashtra, India

ABSTRACT: Communication is an essential part of human interaction, but deaf and mute individuals often struggle to express themselves in environments where sign language is not widely understood. This project aims to bridge the communication gap by developing a **Hand Gesture Recognition System** that converts sign language gestures into text and speech in real time.

By leveraging **computer vision and deep learning**, the system captures hand gestures through a webcam, processes them using **OpenCV**, and classifies them with a **Convolutional Neural Network (CNN) trained in Keras**. The web-based platform ensures accessibility from any device with a camera, making it useful in **education, workplaces, and public spaces**.

This project contributes to **assistive technology** by promoting **inclusivity and independence**, helping individuals with hearing and speech impairments communicate effortlessly in their daily interactions.

KEYWORDS:
- Hand Gesture Recognition
- Sign Language Translation
- Deep Learning
- Computer Vision
- OpenCV
- CNN (Convolutional Neural Network)
- Real-Time Processing
- Inclusivity
- Assistive Technology
- Speech-to-Text Conversion

I. INTRODUCTION

Imagine not being able to speak or hear and struggling to communicate with the world around you. For many deaf and mute individuals, this is a daily challenge, especially when others don't understand sign language. Our Hand Gesture Recognition System helps bridge this gap by recognizing hand gestures in real-time and converting them into text and speech.

Using a webcam, the system captures hand movements, and with the help of computer vision and deep learning, it accurately translates gestures into words. A CNN model (Keras) ensures precise recognition, while text-to-speech technology makes conversations smoother. Since it's a web-based solution, anyone with a camera can use it—whether at school, work, or in public spaces.

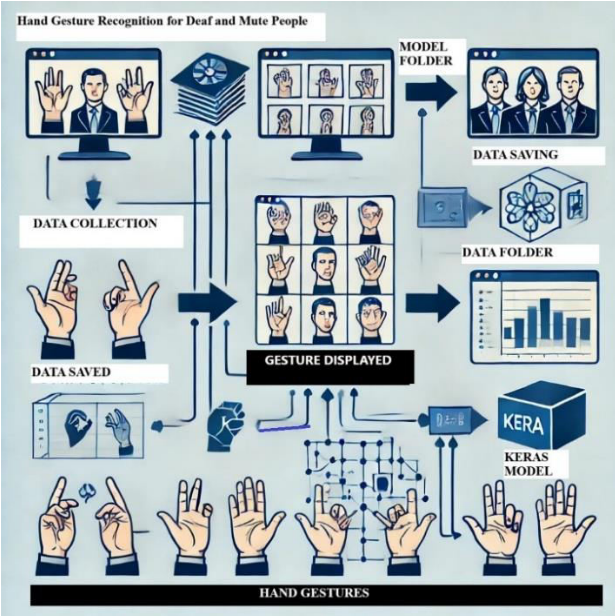
By combining AI and sign language recognition, this system empowers individuals with hearing and speech impairments, making communication more inclusive, independent, and effortless.

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803032

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



- **Importance**

Communication is a basic human need, but for **deaf and mute individuals**, it can be a daily struggle—especially when others don't understand **sign language**. This system **breaks barriers** by converting **hand gestures into text and speech**, making interactions **easier and more inclusive**.

With **real-time recognition**, people can express themselves freely in **schools, workplaces, and public spaces**. This technology not only enhances **independence** but also **raises awareness** about sign language, fostering a **more connected and understanding society**.

- **IMPLIMENTATION**

The **Hand Gesture Recognition System** is built using computer vision and deep learning to recognize sign language gestures in real time. Here's how it works:

1. Data Collection:
   - Images of different hand gestures are captured using a webcam.
   - The dataset is categorized into different gesture classes (e.g., Hello, Thank You, Yes, No).
2. Model Training:
   - A Convolutional Neural Network (CNN) is trained using Keras and TensorFlow.
   - The model learns to recognize and classify different hand gestures based on image features.
3. Gesture Recognition in Real Time:
   - The webcam captures a live video feed, which is processed using OpenCV.
   - The trained CNN model predicts the gesture in real time.
   - The recognized gesture is converted into text and then into speech using a text-to-speech (TTS) engine like pyttsx3.
4. Web-Based Interface:
   - A Flask-based web application allows users to interact with the system.
   - The UI displays the recognized gesture and provides voice output for better communication.

II. LITERATURE SURVEY

Gesture Recognition Using Deep Learning

Researchers have implemented Convolutional Neural Networks (CNNs) to enhance hand gesture recognition accuracy. Techniques such as background subtraction, skin color segmentation, and adaptive thresholding have been employed to improve performance under varying lighting conditions.

Real-Time Sign Language Recognition Using Computer Vision

Studies have explored landmark detection and feature extraction to differentiate between similar hand gestures. Methods like MediaPipe hand tracking and temporal modeling have significantly improved recognition accuracy in real-time applications.

Applications of AI in Assistive Technology

The application of AI in assistive technology has enabled solutions such as gesture-to-text translation and sign language recognition systems. Research indicates that deep learning models trained on large datasets can generalize well across multiple sign languages.

III. SYSTEM ARCHITECTURE

1. Input Layer (**Data Collection**)

- A webcam captures hand gestures in real time.
- Image data is preprocessed (resized, normalized) for consistency.
- The dataset consists of multiple gesture categories (e.g., Hello, Thank You, Yes, No).

2. Processing Layer (**Model Training**)

- Uses OpenCV and MediaPipe to detect and extract hand features.
- A Convolutional Neural Network (CNN), trained with Keras and TensorFlow, classifies the gestures.
- The trained model predicts the recognized gesture with high accuracy.

3. Output Layer (**Text & Speech Conversion**)

- Recognized gestures are converted into text and displayed on the web interface.
- Text-to-Speech (TTS) engine (pyttsx3) converts the text into speech for audio output.

4. Web-Based Interface

- A Flask web application provides an accessible platform.
- Users can see the recognized gestures in real-time on a user-friendly UI.
- The system can run on any device with a camera and internet connection.

5. System Components

- Hardware: Webcam, Computer with a processor capable of running deep learning models.
- Software & Libraries: Python, OpenCV, MediaPipe, Keras, TensorFlow, Flask, pyttsx3.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

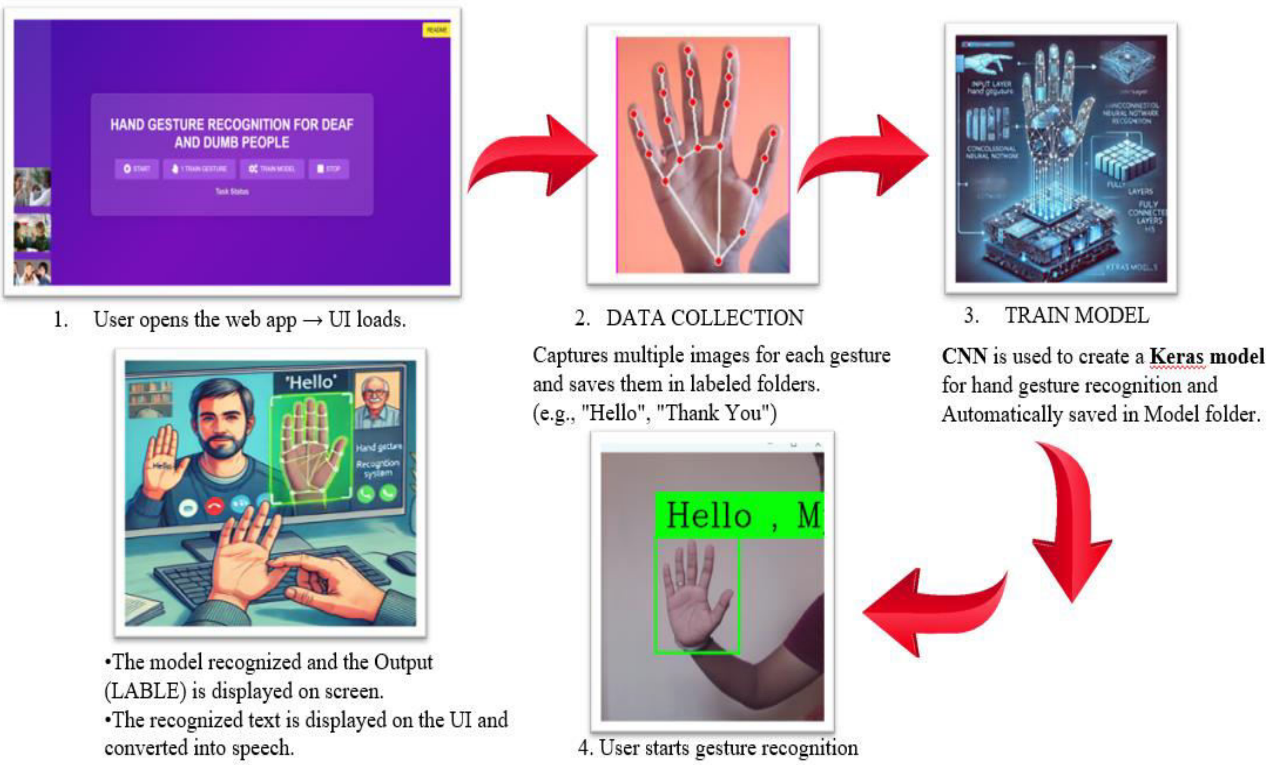
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **Directory structure**

```
Sign-Language-Detection-Main/
├── Data/
│   ├── Hello/
│   │   ├── image1.jpg
│   │   ├── image2.jpg
│   ├── Nice/
│   │   ├── image1.jpg
│   │   ├── image2.jpg
│   └── ... (More gesture categories)
│
├── Model/
│   ├── keras_model.h5   # Trained Keras model
│   ├── label.txt        # Labels for the gestures
│
├── static/              # Static assets ( Images For Page)
│   ├── images/
│   │   ├── logo.png
│
├── templates/           # HTML Templates for Flask
│   ├── index.html
├── app.py               # Flask backend
├── datacollection.py    # Script for collecting gesture images
├── train_model.py       # Script to train the model
├── test.py              # Script to recognize gestures in real time
├── readme.md            # Project documentation
```

IV. FLOW OF WORKING



Flow of Working: Hand Gesture Recognition System

1. **User Opens the Web App**
   - The UI loads in the browser, displaying buttons to start recognition, train gestures, or stop processes.
2. **Data Collection**
   - The system captures multiple images of different gestures using a webcam.
   - These images are stored in labeled folders (e.g., "Hello," "Thank You") for training.
3. **Train Model**
   - A CNN-based Keras model is trained on the collected gesture images.
   - The trained model is automatically saved in the "Model" folder for future recognition.
4. **User Starts Gesture Recognition**
   - The user performs a gesture in front of the webcam.
   - The model recognizes the gesture, displays the output (label) on the screen, and converts it into text and speech for communication.

Where it is used

This web app can be used in various real-world scenarios to help deaf and mute individuals communicate effectively:

1. Virtual Meetings & Video Calls – Enables seamless interaction by detecting gestures and converting them into text or speech.
2. Education & Learning – Helps students with hearing impairments communicate in classrooms and online learning platforms.
3. Public Services & Offices – Assists in customer service, government offices, and hospitals where communication barriers exist.
4. Daily Conversations – Enhances communication in social settings, making interactions smoother and more inclusive.
5. Assistive Technology – Can be integrated into mobile apps, kiosks, and smart devices for greater accessibility.

By making gesture-based communication more accessible, this app empowers individuals and promotes inclusivity in various aspects of life.

V. PROJECT CODE STRUCTURE

1. **Backend (Flask)**

app.py (Main Flask Server)

- Handles HTTP requests and serves the web UI.
- Loads the trained model (keras_model.h5) and label mappings (label.txt).
- Captures real-time webcam input and passes it to the model for gesture recognition.
- Converts recognized gestures into text and speech (using pyttsx3).
- Serves index.html for the user interface.

2. **Data Collection**

datacollection.py (Collect Gesture Data)

- Captures images of hand gestures using OpenCV.
- Saves images into the Data/ directory under labeled folders (e.g., Hello/, Nice/).
- Uses MediaPipe to detect and extract hand landmarks for better accuracy.

3. **Model Training**

train_model.py (Train the Gesture Recognition Model)

- Loads gesture images from the Data/ folder and preprocesses them.
- Uses a CNN (Convolutional Neural Network) model built with Keras.
- Trains the model on collected images and saves it as keras_model.h5.
- Generates label.txt, which stores the class names (gestures).

4. **Gesture Recognition**

test.py (Real-Time Gesture Recognition)

- Opens the webcam and detects hand gestures in real time.
- Loads keras_model.h5 and label.txt to classify gestures.
- Displays recognized text on the screen and converts it to speech.
- Uses OpenCV for camera input and MediaPipe for hand tracking.

5. **Web Interface**

templates/index.html (User Interface)

- Frontend UI for the web application.
- Provides buttons to start/stop recognition and training.
- Uses JavaScript to send webcam feed to app.py for real-time processing.

6. Static Files

static/images/ (Assets)

- Stores images used in the web app (e.g., logo.png).

7. **Model & Labels**

Model/keras_model.h5 (Trained Model)

- The trained deep learning model for gesture classification.

Model/label.txt (Gesture Labels)

- A list of class names corresponding to different gestures.

VI. CONCLUSION

This project successfully bridges the communication gap for deaf and mute individuals by recognizing hand gestures and converting them into text and speech. By integrating **computer vision, deep learning (CNN with Keras), and a Flask-based web interface**, the system provides an accessible and efficient solution for real-time gesture recognition. The use of **OpenCV and MediaPipe** ensures accurate hand tracking, while the trained model effectively classifies gestures for seamless communication. The web-based interface makes it **user-friendly and accessible from any device with a camera**, enabling usage in various settings such as education, workplaces, and public services.

This project **highlights the potential of AI in assistive technology** and can be further enhanced by expanding the gesture dataset, improving model accuracy, and integrating support for multiple sign languages. Future developments could include **mobile app integration** and **real-time translation for different languages**, making communication even more inclusive.

REFERENCES

[1] "HAND GESTURE RECOGINATION SYSTEM FOR DUMB AND DEAF PEOPLE" :- Arun Kumar R V (2024). Hand gesture recognition systems use computer vision to analyze live footage, enabling machines to interpret and respond to human gestures in real-time.

[2] "Hand Gesture Recognition and voice Conversion for Deaf and Dumb" :- Rupesh Prajapati (2018). This project aims to assist deaf and dumb individuals in overcoming communication barriers, promoting equal participation in society through effective visual communication methods.

[3] "Development of Real Time Hand Gesture Recognition for Dumb and Deaf using Machine Learning and Deep Learning" :- Tushar Shelke (2021). This project develops a video-based hand gesture recognition system using deep learning and machine learning to convert gestures into text and speech, enabling communication for deaf and dumb individuals.

[4] "Hand Gesture Recognition for Deaf and Dumb" :- Syed Raquib Shareef (2020). This study proposes a non-vision-based solution for hand gesture recognition, aiming to bridge the communication gap between speech-impaired individuals and others by converting gestures into spoken language.

[5] "Hand Gesture Recognition and Voice Conversion For Deaf and Dumb" :- Meera J (2024) . This study proposes a cost-efficient gesture recognition system using deep learning and image processing to help deaf and mute individuals communicate by converting hand gestures into speech, with a focus on hospitals and clinics.

[6] "Hand Gesture Recognition System for Deaf and Dumb People" :- Shubham Ade (2023). This research develops a sign language recognition system using deep learning and computer vision to detect and translate English alphabets and numbers into text, helping deaf and dumb individuals communicate with others.

[7] "Hand Gesture Recognition System for Deaf and Dumb People" :- Pankaj Barde (2023). This project develops a user-friendly Android application using Kotlin that

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803032

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

recognizes hand gestures through the device's camera, enabling people with disabilities to control devices and perform actions without touch input.

[8] "Hand Gesture Recognition and Voice Conversion System for Deaf and Dumb People" :- Mr.P.V.Krishna Rao (2019). This project develops a glove-based system for deaf and mute communication, using flex sensors and accelerometers to detect gestures and convert them into text or voice output through a text-to-speech module.

[9] "Hand Gesture Recognition System for Dumb and Deaf People" :- N Deepa (2024). This project develops an advanced gesture recognition system using computer vision to bridge communication gaps for individuals with hearing and speech impairments, enabling more expressive interaction with the wider community.

[10] "Hand Gesture Recognition Using CNN" :- Harikrishna H (2024). This project develops an advanced gesture recognition system using computer vision to bridge communication gaps for individuals with hearing and speech impairments, enabling more expressive interaction with the wider community.