SIGN TO SPEECH CONVERSION AND HOME AUTOMATION SYSTEM

Dr. Madhavi Patil
Asst. Prof, E&TC, SSWCOE, Solapur

Ms. Pooja Biradar
B. Tech E&TC

Ms. Ankita Dindore
B. Tech E&TC

Ms. Rupali Angule
B. Tech E&TC

Ms. Saraswati Limbute
B. Tech E&TC

Abstract
The growing demand for inclusive technologies has driven research into systems that bridge the communication gap between differently-abled individuals and the wider community. This paper presents an integrated system that interprets sign language into speech and simultaneously enables control over smart home appliances. The system utilizes computer vision for gesture recognition using convolutional neural networks (CNNs), speech synthesis for voice output, and IoT-based architecture for home automation. The proposed system enhances accessibility, independence, and inclusivity for individuals with speech and hearing impairments.

Keywords*:* Sign language, speech synthesis, gesture recognition, home automation, computer vision, IoT, CNN.

Introduction
Communication barriers for individuals with speech and hearing impairments significantly hinder social and professional integration. Traditional sign language is effective among trained individuals but limits interaction with the general population. Moreover, such individuals often face challenges in controlling smart appliances due to conventional voice- or touch-based systems. This research addresses these issues by combining a sign-to-speech translator with a gesture-controlled home automation system.

Related Work
Several approaches have been proposed for gesture recognition, including glove-based sensors, accelerometers, and vision-based systems. Glove-based methods offer high accuracy but lack user-friendliness due to hardware dependency. Recent advancements in deep learning have improved the accuracy of vision-based recognition. Similarly, home automation via gesture control is explored in

limited studies, primarily using IoT with voice assistants, which are ineffective for non-verbal users. This work aims to bridge both gaps in a unified system.

The problem statement for the sign to speech conversion project involves developing a system that can accurately interpret sign language gestures and convert them into spoken words or text. The goal is to bridge the communication gap between individuals who are deaf or hard of hearing and those who do not understand sign language. The project aims to create a reliable and efficient solution that can accurately recognize and interpret a wide range of sign language gestures, enabling effective communication between individuals with hearing impairments and the general population. The Sign to Speech Conversion project involves developing a system that can accurately translate sign language gestures into spoken words. Sign language is a visual language used by individuals with hearing impairments to communicate. While sign language is widely used within the deaf community, there is a communication gap between individuals who use sign language and those who do not understand it. The goal of this project is to bridge this gap by creating a system that can interpret sign language gestures and convert them into spoken language in real time. Key Challenges: • Gesture Recognition: Developing a reliable and accurate gesture recognition system that can identify and interpret various sign language gestures. • Real-time Processing: Ensuring that the system can process sign language gestures in real time to enable fluid and instantaneous communication. • Adaptability: Designing the system to handle a wide range of sign language variations, as sign languages can differ across regions and even among individuals. • Noise and Background Interference: Overcoming challenges related to background noise, varying lighting conditions, and other visual distractions that may affect the accuracy of gesture recognition. • Vocabulary Expansion: Expanding the system's vocabulary to include a broad range of signs and gestures commonly used in sign language communication. • User Experience: Developing a user-friendly interf

3. System Architecture

Overview

The system consists of three core modules:
- Sign Language Recognition using CNN
- Speech Synthesis
- IoT-based Home Automation

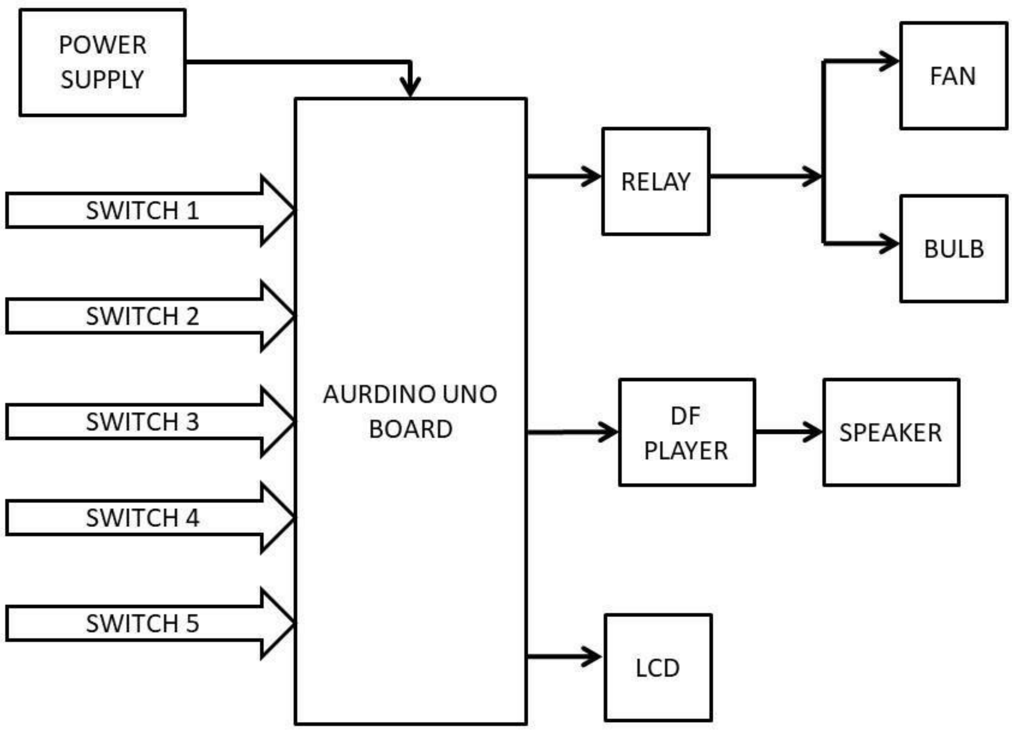
Hardware Components

- Raspberry Pi or Arduino microcontroller
- Webcam for gesture input
- Relay modules for appliance control
- Wi-Fi module (ESP8266/NodeMCU)
- IoT platform (Blynk/MQTT)

Software Components

- Python for image processing and AI
- OpenCV for hand tracking
- TensorFlow/Keras for gesture classification

- pyttsx3 or gTTS for text-to-speech
- MQTT for IoT device communication



In the above figure show that we use five input switch that switches connected to ATmega 328P pin number is d1 to d5 and VCC, GND are connected to d6 & d7. The ATmega 328p follow to instruction or program and it control the all-input pins. The ATmega 328p process to check the line by and instruction and process it. The controller gives output to test instruction. The Microcontroller output is connected to DF player, LCD, & Relay module. ATmega 328p is the processing part of the system. Switches are connected to the ADC of the ATmega 328p Microcontroller. It includes 5 switches. There is one output for each symbol, a text message shown on the LCD screen and a sound heard through the loud speaker. DF Player is a text-to-speech module; this module converts

Algorithm for proposed system

Step 1: Start.

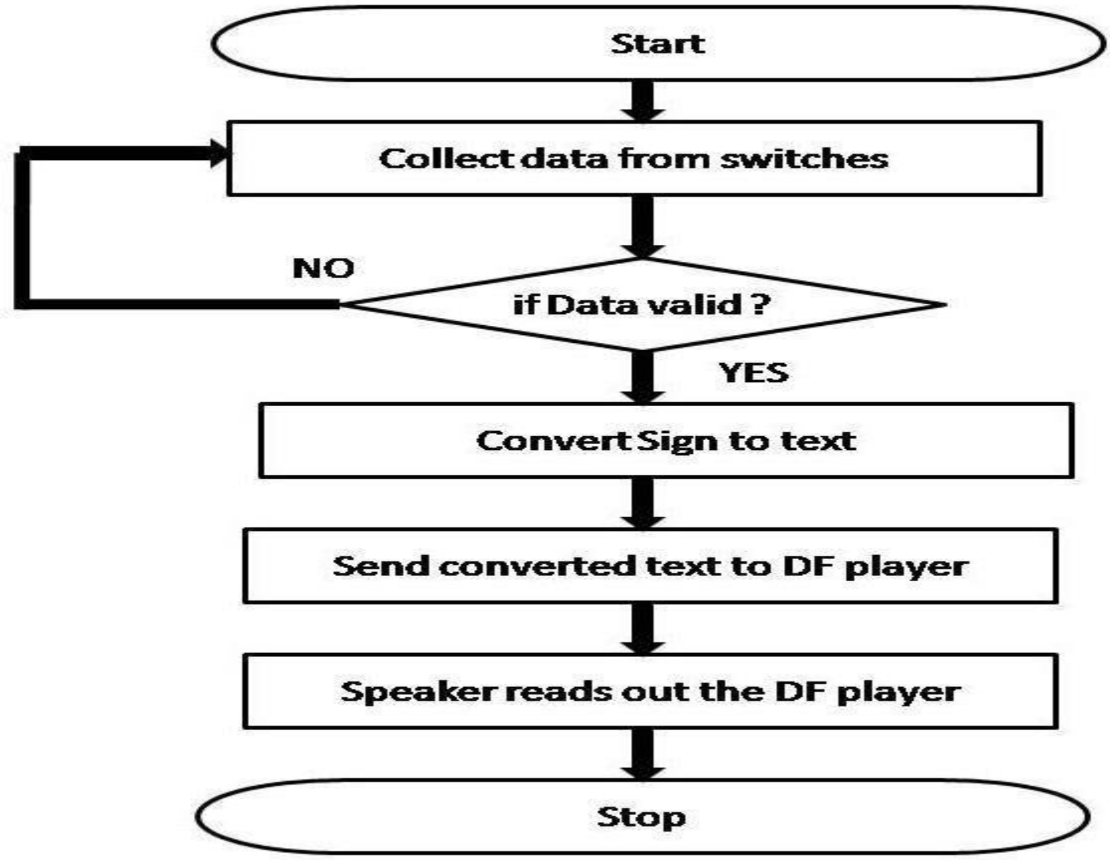
Step 2: Collect data from all the Switches.

Step 3: Is there any meaning to values? If yes then sends conversion of sign language to text.

Step 4: Send converted text to DF player.

Step 5: Speaker reads out the DF player.

Step6:Stop



Methodology

4.1 Data Acquisition and Preprocessing

Hand gesture datasets (e.g., ASL alphabet) were collected and augmented. Images were normalized, resized (e.g., 64x64 pixels), and preprocessed using skin segmentation and background subtraction techniques.

Gesture Recognition using CNN

A CNN model with multiple convolutional, pooling, and fully connected layers was trained to recognize signs. A softmax classifier was used for prediction.

Speech Conversion

Recognized text was converted to speech using TTS libraries like pyttsx3, enabling audible communication for the deaf or mute user.

Home Automation

Specific gestures (e.g., “ON”, “OFF”) were mapped to appliance control commands. These were sent via MQTT to control home devices like lights, fans, and appliances.

Results and Evaluation

The model was tested with real-time hand movements. Performance was consistent under various lighting conditions after training with augmented datasets.

Discussion

The integration of sign-to-speech conversion with gesture-controlled automation opens a new dimension for inclusive design. Challenges include:

- Differentiating similar hand signs
- Background noise in vision systems
- Latency in wireless control

Potential improvements include multi-angle recognition, glove-free 3D hand tracking, and voice feedback for command confirmations.

Conclusion and Future Work

This research demonstrates a cost-effective and scalable system that interprets sign language and controls home devices, enhancing independence and accessibility for differently-abled individuals. Future work includes multilingual sign support, emotion detection, and integration with wearable technology for mobile applications.

References

1. Starner, T., et al. “Real-time American Sign Language recognition using desk and wearable computer-based video.” IEEE TPAMI, 1998.
2. Molchanov, P., et al. “Hand gesture recognition with 3D convolutional neural networks.” CVPR, 2015.
3. Rautaray, S. S., & Agrawal, A. “Vision based hand gesture recognition for human computer interaction: A survey.” Artificial Intelligence Review, 2015.
4. Kaur, H., & Kaur, R. “Home automation system using hand gestures.” IJRTE, 2019.
5.Sign language recognition systems based on feature extraction from color gloves, sensor gloves, color markings on the hands or shadow-based information or depth data are developed by the researcher.
6. For recognition of continuous or dynamic sign language in real time or from videos researchers suggested various methods of features extraction and classification such as motion tracking, HMM classifier, SIFT algorithm, PCA, ANN, Kinet, Eigen values, DBN, etc.
7. Hand-talk Gloves with Flex Sensor: A Review AmbikaGujrati1, Kartigya Singh2, Khushboo3, Lovika Soral4, Mrs. Ambikapathy5.
8. L. Bretzner and T. Lindeberg (1998), ―Relative orientation from extended sequences of sparse point and line correspondences using the affine trifocal tensor, ‖ in Proc. 5th Eur. Conf. Computer Vision, Berlin, Germany,1406, Lecture Notes in Computer Science, Springer Verlag.