International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Sign language Detection and Recognization using Deep Learning

Prof. Archana Ghuge[1], Gaikwad Yash[2], Rajbhar Abhishek[3], Pawar Shantanu[4], Shetty Siddhant[5]

**1 Assistant Professor, Department of Information Technology, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India*
**2,3,4,5 Students, Department of Information Technology, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India*

Abstract - In our modern society, effective communication is crucial for social interaction, personal growth, and professional success. Yet, individuals with hearing and speech impairments often encounter significant obstacles in expressing themselves and understanding others, leading to communication barriers and social isolation. Sign language serves as a vital means of communication for the deaf and dumb community, allowing them to convey thoughts, ideas, and emotions through gestures and expressions.

This research paper introduces a groundbreaking "Sign Language Detection and Recognition using Deep Learning" system to address these challenges. Utilizing Convolutional Neural Networks (CNN) and image processing techniques like Histogram of Oriented Gradients , our system accurately detects and interprets sign language gestures in real-time. Additionally, it features a two-way communication interface that facilitates seamless interaction between individuals with different communication modalities, enabling inclusive communication among deaf, dumb, and normal individuals.

Through a comprehensive methodology encompassing data collection, preprocessing, model training, and system integration, our research aims to advance technology-driven solutions that promote inclusivity, equality, and dignity for individuals with hearing and speech impairments. By empowering effective communication and full participation in social, educational, and professional contexts, our research endeavors to create a more inclusive and accessible world for all members of society.

Key Words: Sign Language Detection, Deep Learning, Speech Recognition, CNN.

1.INTRODUCTION

In a world where communication is paramount, individuals with hearing and speech impairments face unique challenges in expressing themselves and interacting with others. Sign language serves as a vital mode of communication for the deaf and dumb community, offering a means to convey thoughts, ideas, and emotions through gestures, expressions, and movements.

Our research paper, titled "Sign Language Detection and Recognition using Deep Learning," aims to address the communication barriers faced by individuals with hearing and speech impairments through the integration of advanced technologies. Leveraging the power of deep learning, specifically Convolutional Neural Networks (CNN), our research endeavors to develop a robust system capable of detecting and interpreting sign language gestures in real-time.

Sign languages are diverse and varied, with different regions and cultures having their own unique forms of sign language, such as American Sign Language (ASL), Indian Sign Language (ISL), Bangladesh Sign Language (BSL), and Malaysian Sign Language (MSL). These languages are not merely collections of arbitrary gestures but are richly structured systems with grammatical rules and syntax, enabling nuanced expression and communication.

The inability to comprehend sign language poses significant challenges for both individuals with hearing and speech impairments and those who interact with them.
Miscommunication or delayed interpretation can lead to frustration, isolation, and limited opportunities for personal and professional growth.

Our research seeks to bridge this communication gap by harnessing the capabilities of deep learning technologies to accurately detect and interpret sign language gestures in real-time. By providing individuals with hearing and speech impairments with the ability to express themselves fluently through sign language, we aim to empower them to actively participate in various aspects of life with confidence and independence.

Central to our research is the development of a two-way communication interface that facilitates seamless interaction between individuals with different communication modalities. By converting sign language gestures into text or speech and vice versa, our system enables inclusive communication between deaf, dumb, and normal individuals, fostering understanding and empathy within society.

In this research paper, we present a comprehensive overview of our methodology, including data collection, preprocessing, model training, and system integration. We also provide insights into the potential impact of our research on enhancing communication accessibility and empowering individuals with hearing and speech impairments to engage more fully in social, educational, and professional contexts. Through our research, we aspire to contribute to the advancement of technology-driven solutions that promote inclusivity, equality, and dignity for all members of society, regardless of their communication abilities.

2. RELATED WORK

In this section we have written literature survey wherein we have collected papers from various ieee conferences and journals which are mentioned as below:

1. Paper Name: A Review on Smart Gloves to Convert Sign to Speech for Mute Community Author Name: Khan Sohel rana, Syed Faiyaz Ahmed, Shaik Sameer Description: The mute community all over the globe facing many problems while communicating. The normal and dumb people can communicate only in one way i.e. sign language, but many times communicating with normal persons they noticed difficulty. Therefore, there always exists communication barrier. This communication barrier is seen because a speech impaired person uses gesture to commune with common human being which is not suitable. We are implementing this project to reduce the barrier between dumb and normal person. This device design is based on the embedded system. Flex sensor and Node MCU are the key components.
2. Paper Name: A Gesture-to-Emotional Speech Conversion by Combining Gesture Recognition and Facial Expression Recognition
   Author Name: Nan Song, Hongwu Yang* Description: This paper proposes a facial expression integrated sign language to emotional speech conversion method to solve the communication problems between healthy people and speech disorders. Firstly, the characteristics of sign language and the features of facial expression are obtained by a deep neural network (DNN) model. Secondly, a support vector machine (SVM) are trained to classify the sign language and facial expression for recognizing the text of sign language and emotional tags of facial expression. At the same time, a hidden Markov model-based Mandarin-Tibetan bilingual emotional speech synthesizer is trained by speaker adaptive training with a Mandarin emotional speech corpus. Finally, the Mandarin or Tibetan emotional speech is synthesized from the recognized text of sign language and emotional tags. The objective tests show that the recognition rate for static sign language respectively. Subjective evaluation demonstrates that syn the sized emotional speech can get 4.0 of the emotional mean opinion score. The pleasure-arousal-dominance (PAD) tree dimensional emotion model is employed to evaluate the PAD values for both facial expression and synthesized emotional speech.
   Results show that the PAD values of facial expression are close to the PAD values of synthesized emotional speech. This means that the synthesized emotional speech can express the emotions of facial expression.
3. Paper Name: Hidden Markovmodel-based Sign Language to Speech Conversion System in TAMIL Author Name: Aiswarya V, Naren Raju N, Johanan Joy Singh S, Nagarajan T, Vijayalakshmi P Description: Quick-eared and articulately speaking people convey their ideas, thoughts, and experiences by vocally interacting with the people around them. The difficulty in achieving the same level of communicat ion is high in the case of the deaf and mute population as they express their emotions through sign language. An ease of communication between the former and the latter is necessary to make the latter an integral part of the society. The aim of this work is to develop a system for recognizing the sign language, which will aid in making this necessity a reality. In the proposed work an accelerometer-gyroscope sensor-based hand gesture recognition module is developed to recognize different hand gestures that are converted to Tamil phrases and an HMM based text-to-speech synthesizer is built to convert the corresponding text to synthetic speech.
4. Paper Name: A Translator for American Sign Language to Text and Speech
   Author Name: Vi N.T. Truong, Chuan-Kai Yang, Quoc-Viet Tran
   Description: In the year 2001, Viola and Jones's study is a milestone in developing an algorithm capable of detecting human faces in real time. The original technique was only used for the face detection, but many researchers have applied it for the detection of many other objects such as eyes, mouths, car's number plates and traffic signs.
   Amongst them, the hand signs are also detected successfully. This paper proposed a system that can automatically detect static hand signs of alphabets in American Sign Language (ASL). To do that, we adopted the two combined concepts AdaBoost and Haar-like classifiers. In this work, to increase the accuracy of the system, we use a huge database for training process, and it generates impressive results. The translator was implemented and trained using a data set of 28000 samples of hand sign images, 1000 images for each hand sign of Positive training images in different scales, illumination, and the data set of 11100 samples of Negative images. All the Positive images were taken by the Logitech Webcam and the frames size were set on the VGA standard 640x480 resolution. Experiments show that our system can recognize all signs with a precision of 98.7. Input of this system is live video and output is the text and speech.

5. Paper Name: Design and Implementation of A Sign-to Speech/Text System for Deaf and Dumb People Author Name: Dalal Abdulla, Shahrazad Abdulla and Rameesa Manaf, Anwar H. Jarndal Description: This paper presents an approach for designing and implementing a smart glove for deaf and dumb people. There have been several re searches done in order to find an easier way for non-vocal people to communicate with vocal people and express themselves to the hearing world. Developments have been made in sign language but mainly in American Sign Language. This research aims to develop a sign to Arabic language translator based on smart glove interfaced wirelessly with microcontroller and text/voice presenting devices. An approach has been developed and programmed to display Arabic text. The whole system has been implemented, programmed, cased and tested with very good results.
6. Paper Name: On Design and Implementation of a Sign-to Speech/Text Syst
   Author Name: Anwar Jarndal, Ahmed Al-Maflehi Description: In this paper two approaches have been investigated for designing and implementing a sign to speech/text translator. The approaches have been developed and implemented to display a dual language (Arabic and English) text and voice. In the first part of this paper, a vision-based system is developed and demonstrated. In the second part, a glove-based system is designed and implemented. The second system is based on wireless interfaced glove, microcontroller and presenting devices to translate the sign to Arabic/English language.
7. Paper Name: Talking Hands– An Indian Sign Language to Speech Translating Gloves
   Author Name: S Yarisha Heera1, Madhuri K Murthy1, Sravanti V S1
   Description: According to the recent statistics about 7.5population is hearing impaired and Indian Sign Language is the only mode of communication used by them. In this paper we have presented an approach that gives a technique for improving Sign Language Recognition system. In the proposed method; we will be using sensors which are incorporated on a glove to detect the gestures and convert it to speech with the help of a Bluetooth module and an cing artificial speech which provides an environment similar to daily communication which is hard to achieve for speech impaired people.
8. Paper Name: Sign Languages to Speech Conversion Prototype using the SVM Classifier Author Name: Malli Mahesh Chandra, Rajkumar S, Lakshmi Sutha Kumar
   Description: Around 70 million people in this world are mute people. There are children who suffer from Nonverbal Autism. Communication between the people with speech impairment and normal people is very difficult. Normally people with speech impairment use Sign Language to communicate with others. Not each and every person can understand Sign Language. In this paper, a prototype is proposed to give speech output for the Sign Language gestures to bridge the communication gap between the people with speech impairment and normal people. Arduino Nano micro controller is used to collect data from these sensors and sends it to the PC via Bluetooth.
   The PC processes the data sent by the Arduino and runs a Machine Learning Algorithm to classify the Sign Language gestures and predicts the word associated with each gesture. Support Vector Machine (SVM) is used for classification. This prototype is very compact and can recognize both American Sign Language (ASL) and Indian Sign Language (ISL). This prototype not only gives speech to the mute people but also makes them multi linguists.

3.METHODOLOGY

1. Data Collection: The first step in our methodology involves the collection of a diverse and comprehensive dataset comprising sign language gestures from various sign languages. We utilize publicly available datasets as well as collect custom datasets through video recordings of native signers performing a wide range of gestures and expressions.

2. Data Preprocessing: Once the dataset is assembled, we preprocess the data to enhance its quality and suitability for training the deep learning model. This includes tasks such as resizing, normalization, and augmentation to ensure uniformity and variability in the dataset.

3. Feature Extraction: In this phase, we employ image processing techniques, particularly Histogram of Oriented Gradients (HOG), to extract relevant features from the preprocessed images. HOG captures the distribution of gradient orientations in localized regions of the image, providing discriminative features for subsequent classification.

4. Model Architecture Selection: A Convolutional Neural Network (CNN) architecture is chosen and customized for sign language detection and recognition. The architecture is tailored to handle sign language complexities and extract high-level features from input images.

5. Training the CNN Model: The CNN model is trained using supervised learning techniques on the preprocessed dataset. It learns to map input sign language images to corresponding text labels or speech representations, facilitating accurate gesture interpretation.

6. Integration with Real-time Input Sources: The trained CNN model is integrated with real-time input sources such as web cameras and microphones. This enables seamless processing of sign language gestures and spoken language inputs, providing immediate feedback.

7. Two-way Communication Interface: An intuitive user interface is developed to enable two-way communication between deaf, dumb, and normal individuals. Users can input sign language gestures, text, or speech and receive corresponding outputs, fostering effective communication.

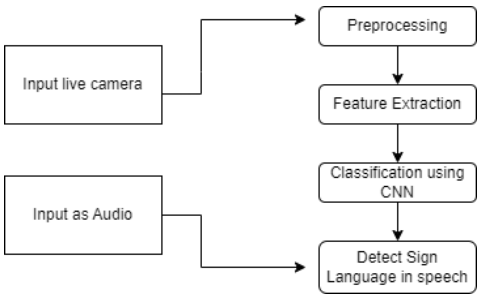
8. Voice Output Generation: Voice output generation capabilities are incorporated to provide auditory feedback to users. This feature enhances accessibility for individuals with hearing impairments, enabling spoken interpretations of sign language inputs.

9. Evaluation and Validation: The performance of the sign language detection and recognition system is rigorously evaluated. Key metrics such as accuracy and latency are assessed through testing and validation procedures.

10. Deployment: Upon successful validation, the system is deployed for real-world usage. It is ready to facilitate seamless communication between individuals with different communication modalities, without incorporating feedback or iterative improvement mechanisms.

4. DESIGN OF SYSTEM

System Architecture:



1. Input Modality Selection:

- The system presents a user interface allowing the user to choose the communication mode for input:
  - **Sign Language:** User performs signs in front of a webcam.
  - **Text:** User types text into a designated text box.
  - **Speech:** User speaks into a microphone.

2. Sign Language Detection (if sign language is selected):

- **Capture Video:** The webcam captures real-time video frames of the user's hand gestures.
- **Preprocessing:** The captured video frames undergo preprocessing to enhance their suitability for feature extraction.
- **Feature Extraction:** Features that uniquely represent the hand pose and movement are extracted from the pre-processed frames.
- **Hand Segmentation:** The extracted HOG features are used to isolate the signer's hand from the background. This allows the system to focus on the relevant region for sign recognition.

3. Sign Language Recognition (if sign language is selected):

- **Convolutional Neural Network (CNN):** A pre-trained CNN model is employed to recognize the detected signs.
  - The CNN architecture will likely consist of convolutional layers that extract higher-level features from the HOG features, followed by pooling layers for dimensionality reduction and fully connected layers for classification.
  - The CNN will be trained on a large dataset of labeled sign language images or video frames. This dataset will map specific hand shapes and movements to their corresponding signs.
  - During recognition, the extracted features from the user's hand gestures are fed into the trained CNN.

- **Output:** The CNN outputs a classification result, indicating the most likely sign based on the processed features.

4. Text Processing (if text is selected):

- **Text Input:** The user types text into a designated text box.
- **Preprocessing:** The entered text undergoes preprocessing techniques to standardize it for further processing.

5. Speech Processing (if speech is selected):

- **Audio Capture:** The microphone captures audio input from the user's speech.
- **Speech Recognition:** Speech recognition techniques are applied to convert the captured audio into text. This could involve using a pre-trained speech recognition model that has been trained on a large speech dataset.
- **Feature Extraction (Optional):** Depending on the conversion approach, additional features might be extracted from the speech audio to facilitate sign language generation.

6. Text/Speech-to-Sign Language Conversion (if text or speech is selected):

- **Text/Speech Representation:** The processed text or speech features are used to represent the content for sign language generation.

7. Sign Language Generation:

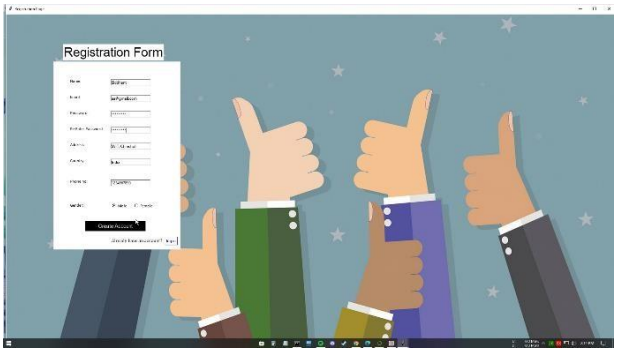
- **Sign Representation:** Based on the recognized sign, processed text, or speech features, the system generates a visual or dynamic representation of sign language. This could involve:
  - **Pre-recorded Videos and images:** A database of pre-recorded sign language videos or images corresponding to different signs can be used. The system would select and display the appropriate video or images based on the recognized sign or converted information.

5.IMPLEMENTATION

Results/Snapshot

1.Start Page

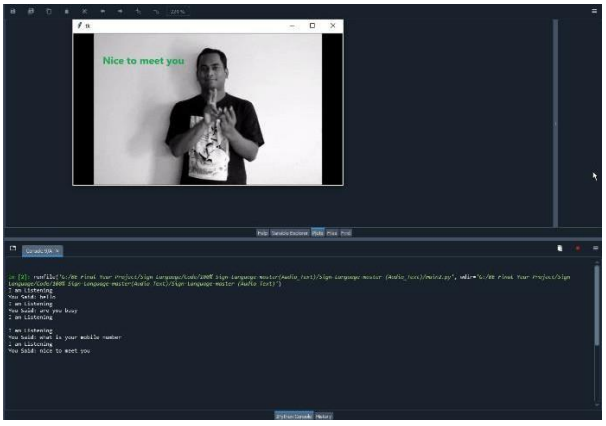
2.Registration Page



3.Login Page

4.Home Screen

5. Voice Recognition



6. Sign Recognition

7. Text Recognition

CONCLUSIONS

Sign Language serves as a vital means of communication for individuals who are deaf-mute, enabling them to express themselves and engage with the world. However, communicating with those who do not understand sign language can be challenging, leading to frustration and isolation.

Our proposed system aims to address this challenge by facilitating seamless communication between deaf-mute individuals and those unfamiliar with sign language. Through real-time translation of sign language into speech, our system eliminates the need for a translator, empowering direct communication.

Traditionally, deaf-mute individuals rely on translators to convey their message verbally. However, this method is impractical in spontaneous situations. Our system bridges this gap by providing immediate translation, allowing users to communicate independently.

Beyond translation, our system acts as the voice of the deaf-mute individual, enhancing their ability to express thoughts and emotions. Designed with user-friendliness in mind, our interface ensures accessibility for users of all technological backgrounds.

In conclusion, our project represents a significant stride towards inclusivity and empowerment for individuals with special needs. By bridging the communication gap, we aim to create a world where everyone can communicate freely and without barriers, fostering understanding and equality.

REFERENCES

1. Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak (Feb 2015), “Study of Sign Language Translation using Gesture Recognition”, International Journal of Advanced Research in Computer and Communication Engineering
2. Jayshree R. Pansare, Maya Ingle (2016),” Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram”, at 2016 International Conference on Image, Vision and Computing.
3. Arslan Arif, Syed Tahir Hussain Rizvi, Iqra Jawaid, Muhammad Adam Waleed, Techno-Talk: An American Sign Language (ASL) Translator, at CoDIT’16- April 6-8, 2016, Malta
4. Justin K. Chen, Debabrata Sengupta, Rukmani Ravi Sundaram,” Sign Language Gesture Recognition with Unsupervised Feature Learning”
5. Matheesha Fernando, Janaka Wijayanayaka,” Low cost approach for Real Time Sign Language Recognition” at 2013 IEEE 8th International Conference on Industrial and Information Systems, ICIIS 2013, Aug. 18-20, 2013, Sri
6. Aarthi M Vijayalakshmi, “SIGN LANGUAGE TO SPEECH CON VERSION”, at 2016 FIFTH INTERNATIONAL CONFERENCE ON RECENT TRENDS IN INFORMATION
7. Caixia Wu, Chong Pan, Yufeng Jin, Shengli Sun, and Guangyi Shi Shaoxing, “Improvement of Chinese Sign Language Translation System based on Collaboration of Arm and Finger Sensing Nodes”, At The 6th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems June 19-22, 2016,
8. Poonam Chavan1, Prof. Tushar Ghorpade2, Prof. Puja Padiya, “Indian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier” at 2016 Symposium on Colossal Data Analysis and Networking (CDAN)

9. Ms. Manisha D. Raut, Ms. Pallavi Dhok, Mr. Ketan Machhale, Ms. Jaspreet Manjeet Hora, “A System for Recognition of Indian Sign Language for Deaf People using Otsu’s Algorithm”, International Research Journal of Engineering and Technology (IRJET)

10. C. Swapna, Shabnam Shaikh “Literature survey on hand gesture recognition for video processing” International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSS 3. van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)